A Systems Level Approach to Identify and Validate Imprinted Genes Involved in Parental Care and the Associated Neural Circuitry

Matthew James Higgs

Supervisors: Professor Anthony Isles and Professor Rosalind John

Submitted for consideration for the degree of Doctor of Philosophy

December 2022

School of Medicine, Cardiff University

Higgsm@cardiff.ac.uk

Student Number: 1812952

Table of Contents

List of Common Abbreviations

***Agrp**:* agouti related neuropeptide

ANOVA: Analysis of Variance

ARC: Arcuate Nucleus

Avp: arginine vasopressin

BORIS: Behavioural Observation Research Interactive Software

BnST: Bed nuclei of Stria Terminalis

Brs3: bombesin receptor subtype 3

***Calcr**:* calcitonin receptor

CNS: Central Nervous System

***C-Fos**:* fos proto-oncogene

DAPI: 4′,6-diamidino-2-phenylindole

Dio3: iodothyronine deiodinase 3

Dlk1: delta like non-canonical notch ligand 1

DMRs: Differentially methylated regions

DRN: Dorsal Raphe Nucleus

E(17.5): Embryonic Day (17.5)

ECM: Extracellular Matrix

ETOH: Ethanol

FC: Fold change

FFPE: Fixed Formalin Paraffin Embedded

FISH: Fluorescent In-situ hybridization

***Gal**:* galanin

GEO: General Expression Omnibus

Gh(rh): Growth hormone (releasing hormone)

GSEA: Gene-Set Enrichment Analysis

***Grb10**:* growth factor receptor bound protein 10

HBAR: Medulla-based Adrenergic

HBNOR: Noradrenaline Hindbrain Neurons

HD: High Definition

HRP: Horseradish Peroxidase

ICC: Intra-class Correlation Coefficient

ICRs: imprinting control regions

IG(s): Imprinted gene(s)

***Igf2(r)**:* insulin like growth factor 2 (receptor)

IGN: Imprinted gene network

IUGR: Intra-Uterine Growth Restriction

KO: Knock-out

Log2FC: Log2 Fold Change

lncRNA: Long non-coding RNA

***Magel2**:* MAGE family member L2

MBA: Mouse Brain Atlas

MCA: Mouse Cell Atlas

MeA: Medial Amygdala

MEG(3): maternally expressed gene (3)

MERFISH: Multiplexed error-robust FISH

Mest: mesoderm specific transcript

MiRs: Micro RNAs

mNEUR: Mature Neurons

(M)POA: (Medial) Preoptic Area

mRNA: Messenger RNA

NBF: Neutral Buffered Formalin

ncRNA: noncoding RNA

NendC: Neuroendocrine Cells

Nms: neuromedin S

ORA: Over-Representation Analysis

Oxt(r): oxytocin (receptor)

P(1): Postnatal Day (1)

PAG: Periaqueductal Gray

PBS: Phosphate buffered Saline

PEG (3/10/13): paternally expressed gene (3/10/13)

PDB: Pup directed behaviour

POA: Preoptic Area

POE: Parent of Origin Effect

Pomc: proopiomelanocortin

PVN: Paraventricular Nucleus

Prl(r): prolactin (receptor)

PWS: Prader-Willi Syndrome

(RT)qPCR: (Real-Time) Quantitative Polymerase Chain Reaction

ROI: Region of interest

RNA: Ribonucleic Acid

SCN: Suprachiasmatic Nucleus

scRNA-seq: Single Cell RNA Sequencing

SD: Standard Deviation

SNc: Substantia Nigra

SnoRNAs: Small non-coding RNAs

SON: Supraoptic Nucleus

SRA: Sequence Read Archive

SSC: Saline-Sodium Citrate

Sst: somatostatin

Th: tyrosine hydroxylase

TIDA: Tuberoinfundibular Dopamine Neurons

TM: *Tabula Muris*

TNDM: Transient Neonatal Diabetes Mellitus

USV: Ultrasonic Vocalisations

VTA: Ventral Tegmental Area

WT: Wildtype

List of Figures

List of Tables

Declarations

This thesis is the result of my own independent work, except where otherwise stated, and the views expressed are my own.

Acknowledgements of Assistance received

Professor Anthony Isles – General advice, guidance and discussion with regard to project design, assistance with perfusion, dissection and tissue collection.
Professor Rosalind John – General advice, guidance and discussion with regard to project direction
Dr. Matt Hill – Specific advice and direction on analysing single-cell RNA sequencing data
Dr. Hannah Tyson – General guidance on animal husbandry, behavioural testing
Dr. Natalie Wellard – General guidance on use of cryostat and general administrative help
Katie Sedgwick – General guidance on use of microtome and mouse perfusion
Alice Chibnall – General guidance on RNAscope® and use of the AxioScan Z1
Anna Webberley – Acting as second scorer for behavioural videos and assisting with initial RNAscope® experiments

Materials/Data obtained from a technical service provider

Mary Lyons Centre (MRC Harwell) – Breeding and providing all genetically-altered mouse cohorts used in this thesis. Additionally for genotyping tail samples taken from the offspring of these mice.

Cardiff University Bioimaging Hub (Histology) – Processing, embedding and sectioning of all brain samples used for RNAscope® in this thesis.

Biorender.com – For providing the platform to create almost all Figures in this thesis.

Data/materials provided by someone else

Single-Cell RNA sequencing data analysed in this thesis were acquired from publicly available resources, but the original authors and specific details of these datasets can be found in Chapter 2 in Table 2.1

Published work associated with this thesis:

Most of the work presented in Chapter 3 has been published – **Higgs, M. J**., Hill, M. J., John, R. M., & Isles, A. R. (2022). Systematic investigation of imprinted gene expression and enrichment in the mouse brain explored at single-cell resolution. *BMC genomics*, *23*(1), 1-22.

Funding

This work was made possible by the generous support of the **Wellcome Trust** (220090/Z/20/Z)

Acknowledgements

What a blur. Three and a half years of laughs, trials and long days. I've had a passion for evolutionary biology for a long time and to get to put some of that passion into this thesis has made it genuinely enjoyable (and genuinely tolerable in the harder times).

It would not have been possible without the support of Anthony and Ros. I can still remember the conference which Anthony and I were at where this project was conceived in the full – who actually thought it would all work out, especially with the pandemic. It may look like pure serendipity that the experiments attempted in this thesis bore fruit but there were many hours spent in which Anthony and I were sat in his office, hashing everything out, which has allowed all this to happen, and so for his time and patience, I would like to thank him. As for Ros, I wouldn't be doing this project without her initial input. It was the initial meetings in her office, talking about evolution and parenting that really set me on this course and her enthusiasm for this project and support of me whenever I put myself out there has been greatly appreciated.

Throughout this project I have been bolstered by the members of the Isles and John labs whom I would like to thank. From the early support of Harri, Hugo, Hannah and Simona when I was first getting started in this area of research and in the animal unit, to the present, with Alice, Isadora and the students I have helped supervise along the way (Cerys, Lauren, and Anna) who gave back to me more than they realized.

I have been blessed with friends and family who have seen me through this too. Molly, thank you for doing the heavy lifting in our friendship, you're a great friend. Warwick crew, thanks for always taking an interest and sorry I've been so bad at keeping in contact, and Cassia, thanks for listening to me no matter, you're very important to me. Cardiff friends (Steve, Pete, Natalie, Alice, and Jaz) thanks for the fun times and mutual Ph.D. moaning. Mum, Dad, and Nan, thank you for the encouragement despite your own trials. I know you want what's best for me and for that I thank you.

Thank you to Crumb and Hux, for always being great and a welcome comic relief with your shenanigans and general weirdness.

Finally, Hannah, I'm sorry for all the late nights driving me to the scanner so I could offload it, or the weeks where I'd leave at 7am and come home at 9pm from the animal unit. It's not been easy for you, but I never felt your love and support waiver for a moment, and you made it so that our Cardiff home was always a place I wanted to come back to. I love you.

Summary

Imprinted genes are primarily associated with prenatal biology but have also been shown to substantially impact the post-natal brain and, its output, infant and adult behaviour. One behaviour classically associated with genomic imprinting is maternal care. Maternal care, and parenting more generally, is an essential behaviour in mammals in order to keep offspring alive who are unable to care for themselves and is controlled by a highly conserved neural circuit centred in the medial preoptic area (MPOA) of the hypothalamus. The mechanism by which imprinted genes influence parenting behaviour is mostly unknown and recent findings have cast doubt on the prior research which established a relationship between these genes and parenting in the first place.

Using a systems level approach, this thesis first analysed a set of single cell RNA sequencing (sc-RNA seq) datasets from the mouse body and brain, analysing in which tissues, regions and cell types, a set of imprinted genes would show enrichment. I hypothesized that, given their relationship to parenting, preoptic area neurons might be one of the enriched areas for imprinted genes. The results of this first study showed that imprinted genes do show enrichment in specific areas of the mouse brain, and one of those regions was the parenting associated neurons of the preoptic area, neurons expressing galanin (*Gal)*, tyrosine hydroxylase (*Th)* and calcitonin receptor (*Calcr)*. This produced a list of 21 imprinted genes contributing to the enrichment in these neurons. These genes represented potential candidates for involvement in parenting behaviour and suggested that imprinted genes may converge on this function as a gene set.

To validate this approach, one of the 21 genes was selected for further study, making up the rest of this thesis. *Magel2* was selected as the candidate gene, a paternally expressed gene with no previously characterised parenting deficits. I confirmed that *Magel2* was enriched in *Gal/Calcr* expressing neurons before finding that mothers, fathers and virgin females, null for *Magel2,* displayed parenting deficits in retrieval, nest building and pup-directed attention. Finally, I showed that *Magel2*-null mothers have reduced activity in the Preoptic Area (POA) following exposure to pups and substantial reduction in galanin expression in the POA. Together, I have found a novel imprinted gene associated with parenting, *Magel2*, and a mechanism by which *Magel2*, and perhaps other imprinted genes, could regulate parenting behaviour, by modulating the neuroendocrine ability of galanin expressing neurons.

1 General Introduction

1.1 Genomic Imprinting

1.1.1 What are imprinted genes? – Discovery & Epigenetic mechanisms

Genomic imprinting is an epigenetic process by which the genome is marked by parent-of-origin specific epigenetic marks which cause genes to be expressed in an allelically biased manner rather than the 'default' biallelic expression (Ferguson-Smith, 2011). Genes under this epigenetic regulation are referred to as imprinted genes (IGs), and genes expressed predominantly from the paternal allele of the genes are referred to as paternally expressed genes (PEGs) and vice versa for maternally expressed genes (MEGs) (see Figure 1.1). Imprinting does not exist in all phyla but appears to have evolved independently in mammals and plants (a few hundred genes from across the genome) and some form of allele expression has been observed in insects (namely paternal X chromosome and whole genome deletions) (MacDonald, 2012, Matsuura, 2020).

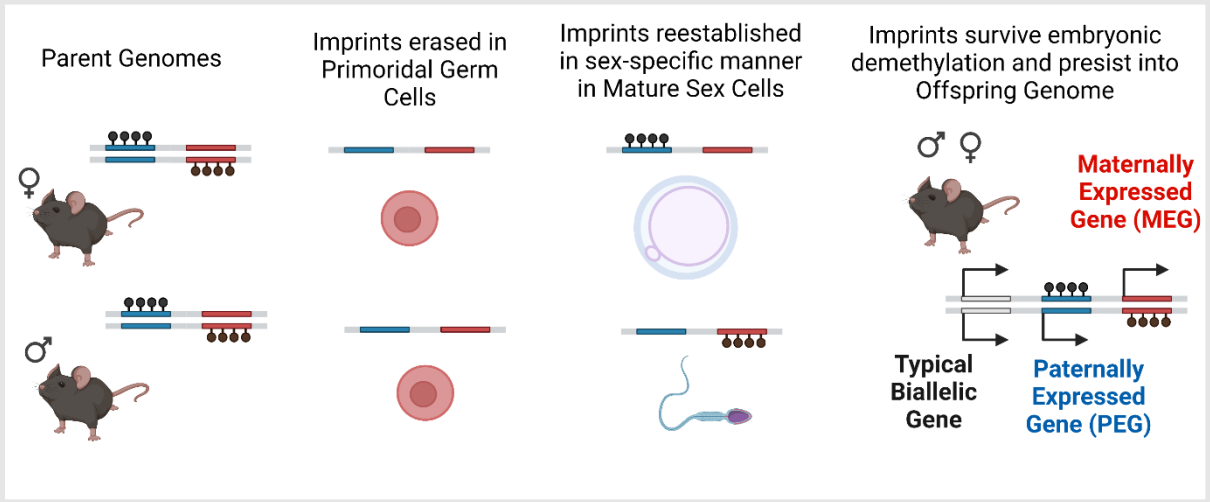


Figure 1.1. Imprinted gene expression and DNA methylation acquisition. Imprinted genes are methylated in a parent-of-origin specific manner. This methylation is erased in the primordial germ cells and reestablished in the sex cells, following the same parent-of-origin specific pattern. Genes are then either paternally expressed genes (PEGs) if the maternal allele is methylated and gene product can only be derived from the paternal allele or maternally expressed genes (MEGs) if vice versa with the paternal allele being methylated. Imprints established in the sex cells survive a second round of demethylation that occurs following successful fertilization to become substantiated in the offspring genome

Epigenetic Establishment and Regulation

To maintain the expression bias of one allele, most imprinted regions show differences in DNA methylation around the parental allele whose expression is suppressed (Reik et al., 1987, Kelsey and Feil, 2013) and DNA methylation fulfils the four criteria for an imprinted mark (Bartolomei and Ferguson-Smith, 2011). DNA methylation places sections of chromosome into a transcriptionally repressed state by modifying CpG dinucleotides (Bird, 2002) and imprinted genes overall contain more intragenic CpG islands compared to biallelic genes (Hutter et al., 2006). This is carried out in the germline by de novo methyl transferases e.g., DNMT3A & DNMT3B (Smallwood and Kelsey, 2012). Typically, two alleles of a gene are equally methylated but imprinted gene regions are known as differentially methylated regions (DMRs).

These DMRs are established in the germline and survive the waves of demethylation during embryogenesis (Messerschmidt, 2012, Seisenberger et al., 2012). The DMRs can then act as Imprinting Control Regions (ICRs) and lead to the establishment of somatic DMRs (which can be tissue specific). Due to this intricate epigenetic process, imprinted genes tend to be organised in contained clusters in the genome with cis-acting ICRs (Reik et al., 2001, Kelsey and Feil, 2013). For many of these clusters, one genetic element carries the primary imprint mark (and is thus directly imprinted and acts as the ICR) while the flanking genes have their marks established indirectly as a consequence of the primary mark. For example, in the *Igf2r (Insulin like growth factor 2 (receptor)*) domain, the primary imprinting mark overlaps with the promotor for *Airn* (Andergassen et al., 2019, Abramowitz and Bartolomei, 2012), which when expressed, promotes the monoallelic expression of the other genes. Deletion of ICRs removes imprinted marks at multiple somatic DMRs (Wutz et al., 1997, Lin et al., 2003, Williamson et al., 2006). Singleton imprinted genes do exist and the establishment and regulation of the monoallelic expression of some of these genes may have a different origin such as retrotransposition into host genes (Cowley and Oakey, 2010).

There are several non-canonical ways imprinted expression can be established and regulated (e.g., via histone modification or the action of small RNA's) but this is beyond the scope of this thesis (but reviewed here – Wan and Bartolomei (2008), Ferguson-Smith (2011), Kelsey and Feil (2013)). However, it is worth discussing that the modern view of imprinting accounts for other varieties in the establishment of an imprinted gene's allelic expression other than absolute (with complete silencing on one allele in every mammalian cell) because

often this expression can be temporal and tissue-specific or even cell type-specific (Laukoter et al., 2020, Bonthuis et al., 2015). More pressingly, the allelic silencing itself can be partial rather than complete, leading to an allelic bias rather than an allelic silencing (DeVeale et al., 2012).

Discovery and Identification

Mammalian imprinted genes were first discovered in the context of parthenogenetic and androgenetic mice (Barton et al., 1984, Surani et al., 1984, McGrath and Solter, 1984). Diploid mouse embryos created with two female or two male pronuclei were not viable post-implantation showing that the maternal and paternal genomes are not equivalent and are both necessary for development. More specific studies using uniparental disomy for specific chromosomes and regions found that it wasn't the whole genome that wasn't equivalent but specific chromosomal regions (Cattanach and Kirk, 1985). It wasn't until 1991 that the first imprinted gene (*Igf2r*) was narrowed down from the broad chromosomal regions (Barlow et al., 1991) which began a process of trying to identify and establish – how many imprinted genes are there?

Various groups have attempted genome wide approaches to identify and validate new imprinted genes (Henckel and Arnaud, 2010). Early approaches utilised microarray and Ribonucleic Acid (RNA-) sequencing using mice of differential strains in reciprocal crossing to find new imprinted genes and finding ~100 genes to be imprinted in the various pre-natal and post-natal tissues (Schulz et al., 2006, Wang et al., 2008). The latest of these attempts however have been heavily criticized (Wang et al., 2011, Gregg et al., 2010) because they found 100's of new genes with parental bias effects but later confirmatory studies (DeVeale et al., 2012, Okae et al., 2012) found that many of these genes did not show imprinted effects in more targeted allelic expression measures, subsequently highlighting the vulnerability of large-scale methods for false discovery. Indeed, an important consideration when investigating imprinted genes at the single-cell level is that monoallelic gene expression is dynamic and stochastic, and that bursts of transcription can occur from one allele at a time, creating the illusion of monoallelic expression if sequenced at that moment (Deng et al., 2014, Symmons et al., 2019). More sophisticated approaches have followed (Babak et al., 2015, Andergassen et al., 2017, Perez et al., 2015, Bonthuis et al., 2015) using pyrosequencing to support new imprinted gene discoveries and found 10's of new genes each, expanding the list of genes from ~100 to closer to ~250. Not all these genes have been

associated with a DMR or epigenetic modification, and so may raise the question whether they should be classed as 'imprinted genes' however, these genes do demonstrate allelic bias in assessments with separate methodologies. The current literature suggests that it is likely that the list of true stable allelically biased genes is nearing completion, which means that to date, the currently identified 260 mouse imprinted genes (Tucci et al., 2019) make up ~0.5% of the mouse genome. What caused this small selection of genes to express predominantly from one allele is explored below.

1.1.2 Why are certain genes imprinted? – Evolutionary Theories

Diploidy protects organisms from deleterious outcomes from mutations in only one allele and hence the paradoxical events that have caused certain genes to become imprinted is necessary to explain (Otto and Gerstein, 2008). Within evolutionary theory, successful organisms are those which copy their genetic material into the future. The relative success with which an individual achieves this is called their 'fitness.' From the gene's point of view, success is leaving the maximum number of copies of itself in the population. For imprinting to evolve, the benefit of imprinting to the parent's fitness must outweigh the vulnerability it creates for the offspring. It is well observed that the two major lineages to produce classical imprinted genes, and therefore establish a beneficial fitness trade-off, are eutherian mammals (relying on an invasive placenta) and flowering plants (whose endosperm has a placenta-like function) (Hore et al., 2007, Scott and Spielman, 2006) in which IGs evolved independently (Feil and Berger, 2007, Pires and Grossniklaus, 2014). Comparative genomics have shown that several classical imprinted genes are either not present in the genomes of non-mammals or are expressed in a biallelic manner (Fresard et al., 2014, Lawton et al., 2005). This is also the case for monotremes (Killian et al., 2001, Killian et al., 2000) and extensive work in marsupials (with less invasive placenta) have shown that only a few DMR's exist in the marsupial genome (Renfree et al., 2009, Rapkins et al., 2006). This strongly suggests that the extended gestation, placentation and offspring-parent dynamics of eutherian mammals are the evolutionary landscape for mammalian imprinting evolution and this relationship between the imprinting and enhanced provisioning of offspring has led to several hypotheses to explain the manifestation of genomic imprinting.

Conflict/Kinship Theory

The most widely accepted theory for the origin of genomic imprinting in mammals is kinship theory, often referred to as parental conflict theory. This was first formally put forward by Moore and Haig (1991) and views the evolution of genomic imprinting as a consequence of a wider evolutionary battle between mothers and fathers (Haig, 2000).When it comes to procreation, mothers and fathers are not in evolutionary agreement about the best way to benefit their offspring. The fitness of the father's genome would be maximally enhanced by encouraging the offspring to grow and develop at as much expense to the mother as she can tolerate, through an invasive placenta, high levels of suckling and feeding and through elevated offspring growth and metabolism. By the nature of reproduction in most mammals, the father has no guarantee that the next child the mother begets will be his and no incentive to protect the mother's reproductive fitness at the expense of his child. The mother, on the other hand, will aim to produce many offspring in her time and providing everything to one child is not a sensible allocation of her fitness (Frost and Moore, 2010, Piedrahita, 2011).

For many imprinted genes, where the gene is growth enhancing and encourage placental development, the maternal allele has been silenced (creating a PEG) and vice versa when the gene is growth restricting (creating a MEG), which is the first level of conflict (Haig, 2015, Patten et al., 2016). Conflict theory captures this element of a parental battle across the offspring genome to benefit each parent's long-term fitness with the ultimate resolution of the conflict coming from complete epigenetic silencing of one of the alleles. A second level of conflict emerged when it was realized that imprinted gene phenotypes interact and contrast. *Igf2* and *Igf2r* are key examples of this idea, combatively regulating resource acquisition (Smith et al., 2006). The PEG (*Igf2*) encodes a polypeptide in fetal circulation that binds to insulting receptors and stimulates an array of metabolism, proliferation and differentiation growth effects (St-Pierre et al., 2012, Sferruzzi-Perri et al., 2011, Chao and D'Amore, 2008, Sibley et al., 2004, DeChiara et al., 1991). The MEG (*Igf2r*) is one of the receptors for *Igf2* and sequesters it for internalization and degradation, effectively stunting the growth effect of *Igf2* (Wutz et al., 2001, Barlow et al., 1991). This battle then ensues with the paternal half of the genome in the offspring produces *Igf2* to enhance growth while silencing *Igf2r* while the maternal half of the genome expresses *Igf2r* to contain the effect of the paternal allele expressing *Igf2* while suppressing the maternal version of *Igf2*. A third level of conflict theoretically exists between whole systems (i.e., regulation of dopamine vs. serotonin) but that remains to be validated (Wilkins et al., 2016). Overall, in this theory, this growth

suppression/expression exchange is believed to be the origin of most if not all imprinted loci and is born from the competing interests of the maternal and paternal genes.

Coadaptation

A contrasting theory, to potentially explain some imprinted gene effects that were difficult to square with Conflict theory is the Coadaptation theory, first proposed by Curley et al. (2004) and formalized by Wolf and Hager (2006). Coadaptation here is the process by which two individuals undergo adaption as a pair, in this case mother and offspring. Hence, it proposes that imprinted expression has evolved to select the most fitness promoting allele at certain loci for both mother and offspring since the fitness of a genotype depends on its interactions with the phenotypes of others and not just on the direct phenotypic effects of the gene. For example, the maternal allele would be selected in situations in which matching the expressed allele in the offspring with the mother's genome would provide beneficial outcomes, hence coordinating certain traits in a pleiotropic manner in mothers and offspring. This means behaviours such as maternal care, lactation can be co-ordinated with care-solicitation and suckling in the offspring. The larger proportion of maternally expressed genes in the genome is viewed as evidence that there is a bias towards expressing maternally inherited genes that needs explanation. However paternal-expression can still be understood to evolve in situations where a mismatch between mothers and offspring would result in higher fitness. This ultimately proposes that imprinting is the act of coordinating expression of genes and in this sense the maternal and paternal genomes are not in conflict but conspiring to find the optimal solution for fetal development (Wolf, 2013).

Though Coadaptation theory has some support, its general explanative power is restricted to a few specific situations for imprinted genes, and while it is unlikely to be the ultimate origin of the imprinting phenomenon, it does appear that certain imprinted genes do work to coordinate mothers and infants together (O'Brien and Wolf, 2017, Wolf et al., 2015). Examples of coadaptation include *Grb10*'s (Growth factor receptor bound protein 10) effect on nutrient demand in pups and milk supply in mothers (Wilkins, 2014, Cowley et al., 2014) or *Peg3*'s effect on ultrasonic vocalizations (USVs) , thermal regulation and general demand of pups for maternal care as well as its influence on lactation and maternal care quality (Curley et al., 2004, McNamara et al., 2018a). Although Coadaptation is a possible explanation for certain situations, Conflict Theory can explain the majority of these co-adaptive actions of imprinted genes across the offspring and placenta (Úbeda and Gardner,

2015), but it tends to be the role of imprinted genes in post-natal adaptions that cannot be easily explained by Conflict Theory and so have seen the generation of theories such as Coadaptation as well as the final idea summarized here, Dynamic Control.

Dynamic Control

The third, and the final, idea considered here has emerged based on how some processes that imprinted genes regulate are sensitive to environmental influences and hence would need to rely on genetic components that can be modulated (Ferguson-Smith, 2011, Charalambous et al., 2012). In "Dynamic Control," genomic imprinting is seen as an epigenetic mechanism to regulate gene dosage, allowing some modulation of genes to environment and supporting the temporal and cell-specific nature of imprinted gene expression. One primary example used to support this idea is *Dlk1* and *Igf2* which lose their PEG status during the process of neurogenesis, suggesting that an increase in *Dlk1/Igf2* expression is necessary for this process (Ferrón et al., 2011, Surmacz et al., 2012, Montalbán-Loro et al., 2021, Ferrón et al., 2015). Furthermore, changes in dosages of *Dlk1* and *Dio3* have been shown to promote or suppress brown fat recruitment and hence gene dosage here is utilised to shift animals from the 'nest phase' of life to independent life (Charalambous et al., 2012). Very recently, changing the dosages and consequently the expression ratios of *Glt2* and *Dlk1* have been shown to alter body weight phenotypes from healthy to lethally underweight, continuing to support the idea that this locus is capable of altering phenotypic outcomes with alterations in dose (Weinberg-Shukron et al., 2022).

Another recent example of dynamic control is the relationship between imprinted genes and placental hormone regulation (Creeth et al., 2019). For example, when the dosage of *Phlda2* was decreased (from 0% to 100% to 200% normal), mothers spent more time performing nest building and less time engaged in pup directed behaviour (PDB), likely linked to the effect *Phlda2* has on restricting spongiotrophoblast size and the hormonal output it provides (Tunster et al., 2016a). Theoretically, *Phlda2* dosage could be modulated to shift behavioural activity of mice which is technically what imprinting achieves. In response to adversity, another gene from this region, *Cdkn1c*, has shown to increase its expression in response to adverse early environments (namely a maternal low protein diet and exposure to poor maternal care) (Van de Pette et al., 2017). *Cdkn1c* regulates dopaminergic neurogenesis (Joseph et al., 2003) and mocking up a loss-of-imprinting model (i.e. double expression) of *Cdkn1c* produces a hyper-dopaminergic animal (McNamara et al., 2018b), strongly

suggesting the *Cdkn1c* can module dopaminergic ability in the animal during early post-natal development (Isles et al., 2019).

It could be suggested that this role for imprinting is limited to stem cell/neurogenesis applications but, on the whole, for the behavioural effects of imprinted genes postnatally, it has been suggested that imprinted genes present an epigenetically regulatable element by which to modulate behaviour in response to environmental influence, theoretically carrying across generations or being acted upon as an evolutionarily malleable target (Curley and Mashoodh, 2010).

1.1.3 Where do imprinted genes act? – Expression summary & the Imprinted Gene Network (IGN)

Regardless of the ultimate origin of imprinted genes, where and how imprinted genes are expressed can provide a fundamental insight into the functional role of these genes.

Imprinted Gene Expression

A fairly robust image of how many imprinted genes existed and where and when in the body they were having their significant effects was built through working on individual genes and clusters and viewing the work collectively. This classical work on individual imprinted genes (e.g., Li et al., 1999, Kozlov et al., 2007, Lefebvre et al., 1998, Garfield et al., 2011) showed a convergent expression profile in the placenta, embryo and the adult brain. Whether the expression profiles of all imprinted genes show some form of enrichment in these tissues has yet to be confirmed, however for a selection of imprinted genes, expression profiles, consisting of a list of tissues indicating whether the genes are expressed or not, have been compiled. The first was by Steinhoff et al. (2009) using microarray data from the mouse and human body, embryo and placenta. Expression of IGs was seen (contributed by different genes for each species) across key energy metabolism organs – adrenal gland, pituitary gland, placenta and pancreas and the brain. It is worth noting that many of the imprinted genes express ubiquitously across tissues while some show tissue-specific expression which drive the formation of hotspots. Babak et al. (2015) and Andergassen et al. (2017) looking at imprinted expression across multiple mouse tissues and timepoints came to a few broad conclusions when it comes to imprinted gene expression. Most imprinted genes were either expressed from one allele or not expressed at all in tissues. More imprinted genes are

expressed in neonatal and extraembryonic tissues compared to adult tissues, but the brain is the strongest locale for imprinted gene expression in the adult body.

In summary, imprinted genes as a gene set are expressed broadly in the animal body. Most genes are imprinted in embryonic and placental tissues with the number of genes showing imprinted expression decreasing into adulthood with endocrine and neurological tissues expressing the largest number of imprinted genes postnatally (Andergassen et al., 2017). However, when carrying out this profiling of multiple imprinted genes, another phenomenon was found concerning their expression, which was the intimate co-expression of these genes across multiple chromosomes, forming some form of imprinted gene network (IGN).

Imprinted Gene Network

Imprinted genes have been classically studied in isolation or in chromosome restricted domains which is logical since imprinted genes have not appeared to converge on a single biochemical process and are members of a variety of gene families (Patten et al., 2016). IGs do however, converge on a set of physiological phenotypes (discussed in full in Section 1.1.4). Hence when Varrault et al. (2006) found heavy co-expression between their imprinted gene of interest, *Zac1*, and a selection of other IGs and suggested these genes operate in an expression network, it fitted with what was already known. They further showed that knocking out *Zac1* in embryonic tissue had consistent impacts on other imprinting genes (either up or down regulating them) and hence indicating *Zac1* had a regulatory role on other imprinted genes. Later, Gabory et al. (2009) showed the same pattern, consistent upregulation of some and downregulation of other imprinted genes when knocking one out the long non-coding RNA (lncRNA) *H19* in embryonic muscle tissue. They also showed this effect held true in adult muscle tissue (Martinet et al., 2016) but importantly, in the original study, they failed to find this effect in the placenta, suggesting this network effect is not true for all tissues at all timepoints. Finally, Al Adhami et al. (2015) found that this network, now expanded to a large collection of IGs and closely associated biallelic genes, is involved in regulating cell-cycle exit into differentiation and the prevention of re-entry likely through growth signalling and extra-cellular matrix (ECM) remodelling mechanisms. The imprinted gene network became highly upregulated at the point of cell-cycle exit in a spike of expression and increasing IGN expression experimentally decreases proliferation, indeed Lui et al. (2008) had shown previously that imprinted genes were dynamically expressed during the later periods (descaling) of somatic growth.

Much remains to be made clear, such as whether this network still operates after somatic growth and in all mouse tissues. Additionally, it is still unclear whether this network is mostly made up of convergent expression or whether it is down to direct regulation of one imprinted gene on another. Several examples of direct regulation do exist such as IPW on *Dlk1* locus (Stelzer et al., 2014) and *Mir-379/410* on *Peg3* and *Plagl1* (Whipple et al., 2020) which suggests direct regulation of other imprinted genes could substantially explain the co-ordinated up and down expression. Either way, the idea of a coordinated, regulated network of imprinted genes has gained traction. If this network has roles beyond somatic growth, then this would go some way to explaining why imprinted gene clusters on disparate parts of the genomes have heavily overlapping phenotypes in mouse models and in imprinted disorders (Abi Habib et al., 2019, Eggermann et al., 2021) since disturbing one domain, in practice, disturbs many.

1.1.4 What do imprinted genes do? – Phenotypic summary

This network of co-expressed genes with convergent locales of high expression has a shared set of phenotypic outcomes which has greatly informed the evolutionary theories described in Section 1.1.2.

The predominant phenotypes seen in imprinted gene mouse models overlaps heavily with the clinical profiles of the various human imprinting disorders, which exist for most major imprinted domains on the human chromosome. They include Transient neonatal diabetes mellitus (TNDM) (6q24), Silver-Russell syndrome (7p13/11p15), Beckwith-Wiedemann syndrome (11p15) Temple syndrome (14q32), Prader-Will (PWS)/Angelman syndrome (15q11.2), Pseudohypoparathyroidism (20q13) and Mulchandani-Bhoj-Conlin syndrome (20). These syndromes each display clinical phenotypes, but there is remarkable overlap between the conditions (Eggermann et al., 2021). Almost all of the conditions mentioned previously involve some form of *in-utero* and/or post-natal growth restriction and some form of metabolic disturbance such as hyper/hypoglyemia or hyper/hypoinsulinism). Behaviourally, infants are likely to have feeding difficulties, often manifesting in obesity, whilst children have reproduction and puberty problems, hypotonia and some level of cognitive impairment (Monk et al., 2019, Mackay and Temple, 2017, Cassidy et al., 2000, Butler, 2009, Eggermann et al., 2021).

Perhaps the most prevalent and noteworthy phenotype arising from mouse models with disturbances of imprinted genes are the effects on embryonic and postnatal growth. Many

models are born with *in-utero* growth restriction (IUGR) (review - Piedrahita (2011), Cleaton et al. (2014), Peters (2014)) and several maintain growth problems postnatally (Millership et al., 2019, Cleaton et al., 2016). In accordance with Conflict Theory, most of these genes are PEGs, and indeed several MEGs have been shown to enhance growth (Peters, 2014, Madon-Simon et al., 2014). The origin of the growth restriction does appear to be associated with improper cell differentiation with issues in skeletogenesis and tissue development/growth (Lui et al., 2008) (i.e., this happens to be the point in development with heavy specialization and differentiation) linking directly to the role the IGN is suggested to play based on its genetic interactions (Abi Habib et al., 2019).

The other predominant problem leading to IUGR appears to be aberrant placental function. Imprinted gene models have been shown to have aberrant placenta which are small or overgrown, do not attach properly, and/or do not exchange nutrients properly (Monk, 2015, Tunster et al., 2013, Sibley et al., 2004, Sibley et al., 2005, Charalambous et al., 2010, Coan et al., 2005, Yevtodiyenko and Schmidt, 2006). It has also been shown that within these models the placenta may have reductions in the cell lineages necessary to hormonally coordinate the mother (Creeth et al., 2018, John, 2017, Tunster et al., 2018). Considering the importance of the placenta for the prolonged gestation in eutherian mothers, the impact on the offspring is severe and several models are embryonically or neonatally lethal (Peters, 2014, Cleaton et al., 2014)

Coupled with the phenotypic effects on fetal growth, imprinted genes have also been shown to impact metabolism and growth more generally in neonates and adults. Once out the womb many models fail to catch up their weight during the neonatal stage and some models are partially lethal at this stage (Peters, 2014, Cleaton et al., 2014). Suckling deficits are common for these neonates which explains much of the problem behaviour (Schaller et al., 2010, Curley et al., 2004, Plagge et al., 2004, Cowley et al., 2014, Dent and Isles, 2014, Curley et al., 2005). Many of these models also have issues with sugar/fat metabolism in general and TNDM is a classic imprinted disorder in humans matching this phenotype (Smith et al., 2006, Curley et al., 2005). Interestingly, if the imprinted gene model is viable at the neonatal stage, several lead to adiposity in early-life following growth retardation in the womb (Weinstein et al., 2010, Peters, 2014), and indeed this is a hallmark feature of imprinted disorders such as PWS (Cassidy et al., 2000).

In adult mice, several models have deficits in reproductive performance (Mercer and Wevrick, 2009, Butler, 2009), several models have issues with motility which is likely skeletomuscular developmental in origin (Mercer et al., 2009, Yu et al., 1998) and several models have issues with thermogenesis and heat coordination (Charalambous et al., 2012, Curley et al., 2005). However, the adult mouse brain is the prominent site for imprinted expression in the adult mouse and consequently there are a diverse of collection of brain and behaviour phenotypes observed for these genes which is mirrored in the cognitive disfunction seen in imprinted disorders.

1.1.5 What do imprinted genes do in the brain? – Brain and Behaviour summary

Imprinted genes are most commonly expressed in the brain in adult tissues, but this is true for the whole genome (~80% genome is expressed in the brain (Negi and Guda, 2017, Lein et al., 2007)) so this alone doesn't guarantee functional consequence. Reassuringly, the high-profile/well-studied imprinted genes do show regional specificity in the brain which is indicative of a more targeted role. Imprinted genes such as *Peg3* (Kuroiwa et al., 1996), *Mest* (Lefebvre et al., 1998), *Magel2* (Bischof et al., 2007) and *Dlk1* (da Rocha et al., 2007) all show high expression in regions of the hypothalamus, which is widely recognized as the hub of imprinted expression In the brain (Pulix and Plagge, 2020, Ivanova and Kelsey, 2011). Gregg et al. (2010) assessed this relationship to some degree by cataloguing the expression status (expressed/not expressed) of 45 imprinted genes across 118 regions, These genes were convergently expressed in 26 hotspots – spanning the midbrain (Ventral Tegmental Area(VTA), Periaqueductal Gray (PAG)), hindbrain (Dorsal Raphe and other Raphe nuclei) a few distinct hypothalamic nuclei (Preoptic Area(POA), Arcuate Nucleus (ARC)), the Bed Nuclei of Stria Terminalis (BnST) and a few amygdala nuclei (Central and Medial) but it should be noted that most imprinted genes showed expression in all regions of the brain and this finding was driven by the few region-specific genes. In contrast a selection of comparable biallelic genes had a similar number of hotspots concentrated in the cortical, hippocampal and olfactory regions (which would be expected for a generic gene set as the cortex is more molecularly diverse). Again, whether imprinted genes, as a gene set, show enriched expression in any areas of the brain has yet to be assessed.

Brain Phenotypes

One thing noted about the gynogenetic, and androgenetic chimeric embryos was that brain size was abnormally large in the former and abnormally small in the latter while the size of the body showed the opposite pattern (Keverne et al., 1996a). *Grb10* was found to replicate this phenotype alone and this was significant as it was a MEG in the body while a PEG in the brain (Garfield et al., 2011, Charalambous et al., 2010, Charalambous et al., 2003). One other observation from the chimeras was that gynogenetic cells mainly localized in the cortex, striatum and hippocampus while androgenetic localized to the hypothalamus. A phenomenon that is still mostly unexplained given the overwhelming association between all imprinted genes (MEGs and PEGs) and the hypothalamus.

In terms of the molecular role that imprinted genes are playing in the brain, it has been shown that in the developing brain, several imprinted genes play roles in neurogenesis, neuronal migration, dendritic outgrowth and apoptosis, which provides a very transparent mechanism through which brain size and cell numbers can be altered while the brain is developing (Perez et al., 2016). Postnatally, several imprinted genes converge on regulating synaptic transmission such as *Kcnk9*'s impact on neuron firing patterns (Brickley et al., 2007) and *Ube3a*'s effect on amplitude of action potentials (Wallace et al., 2012, Smith et al., 2011). It is worth noting that while several imprinted genes have the functions listed above, to suggest that this is a mechanism by which imprinted genes regulate brain activity in general would require these phenotypes to be seen in substantially more genes followed by formal statistical analysis which has not yet been explored.

More established is the role imprinted genes appear to be having on specific neural circuits, predominantly in the hypothalamus. The hypothalamus is a unique area of the brain with a large selection of unique neural types, regulating specific physiological processes, clustered into regions (Saper and Lowell, 2014). One substantial role of the hypothalamus is to interact with the pituitary gland in a neuroendocrine relationship to regulate the hormonal output of this organ, but the pituitary gland is a site of imprinted gene action itself (Davies et al., 2008, Scagliotti et al., 2021). PWS involves clear pituitary gland dysfunction with reduced levels of testosterone and growth hormone (GH) (Miller et al., 2008). Growth hormone and the Somatotropes of the pituitary that release it, have been shown to be sites of activity for a number of imprinted genes: *Nnat* (Huerta-ocampo et al., 2004), *Dlk1* (Ansell et al., 2007, Yevtodiyenko and Schmidt, 2006, Charalambous et al., 2014) and the *Gnas* locus (Picard et al., 2007).

Within the hypothalamus imprinted genes have been shown to have phenotypes in several neuron populations such as *Asb4*'s general high expression and responsiveness to fasting in proopiomelanocortin *(Pomc)* neurons of the hypothalamus (regulation of food intake) (Li et al., 2010), *Kcnk9*'s role in membrane potential maintenance in orexin neurons (Guyon et al., 2009). The roles of *Magel2* and *Peg3* in oxytocin (*Oxt*) neuron number in the paraventricular nucleus (PVN) (Li et al., 1999) and in regulation of oxytocin mature peptides (Schaller et al., 2010). Several imprinted genes (e.g., *Magel2* and *Ube3a*) display very high expression but unknown roles in the Suprachiasmatic Nucleus (SCN) (Kozlov et al., 2007, Jones et al., 2016). Outside of the hypothalamus, imprinted genes have been shown to express highly in and effect the differentiation of dopamine neurons (McNamara et al., 2018b, Jacobs et al., 2009) as well as evidence that key dopamine and serotonin metabolic genes show imprinted expression in the hypothalamus (Bonthuis et al., 2015) and regulate monoaminergic output (Bonthuis et al., 2022, Dent and Isles, 2014, Davies et al., 2008). It has never been assessed whether the exemplar imprinted genes are representations of a genuine enrichment of function for IGs as a whole or whether these are outliers. Regardless, imprinted genes (both MEGs and PEGs) are primarily acting on some of the most fundamental regions of the brain for motivated behaviours – the hypothalamus, monoaminergic nuclei and pituitary.

Behaviour Phenotypes

Changing the activity, size or cell numbers of the brain doesn't necessarily have to produce clear and obvious behavioural differences, but in the case of the imprinted gene models; the behavioural phenotypes are prominent (see Figure 1.2). In mutant pups, several imprinted genes have been showed to regulate suckling behaviour (Schaller et al., 2010, Curley et al., 2004, Plagge et al., 2004, Cowley et al., 2014), activity (Weinstein et al., 2010) and maternal demand through the production of ultrasonic vocalizations (Jiang et al., 2010, McNamara et al., 2018a, Bosque Ortiz et al., 2022). From this summary, it is thought that imprinted gene activity in the pup brain is targeted on resource acquisition. In adult brains, the behavioural impacts are more widespread.

Imprinted genes have been shown to directly alter feeding behaviour (Plagge et al., 2004, Schaller et al., 2010, Chamberlain et al., 2004, Koza et al., 2006), with one mechanism suggested to be the modulation of activity of feeding neurons such as the *Agrp (Agouti related neuropeptide)* and *Pomc* neurons of the Arcuate nucleus (Cassidy and Charalambous, 2018). IGs have been shown to regulate mood, with imprinted genes been shown to reduce

and increase anxiety and depression like symptoms in roughly equal measure (Yokoi et al., 2006, Gotter et al., 2011, Luo et al., 2015, Fountain et al., 2017b). Social behaviour is aberrant with mice showing inappropriate or reduced sexual interest (Mercer and Wevrick, 2009, Swaney et al., 2008) as well as interesting effects on social behaviour such as *Cdkn1c*'s role in social dominance (McNamara et al., 2018b), *Magel2* and *Nesp*'s role in response to social novelty and impulsivity (Fountain et al., 2017b, Plagge et al., 2005, Dent et al., 2016) and *Grb10*'s role in social structure (Garfield et al., 2011, Rienecker et al., 2018). Of recent interest is the role of imprinted genes in the regulation of sleep (Lassi et al., 2012, Ehlen et al., 2015, Tucci, 2016, Lassi et al., 2016) and circadian activity (Kozlov et al., 2007, Curley et al., 2005, Coulson et al., 2018), particularly the high expression and role in the circadian neurons of the suprachiasmatic nucleus (Panda et al., 2002). To date, it has not been reliably established whether the predominant mechanism for these behavioural consequences are caused by neurodevelopment issues affecting the post-natal cell numbers of these vital neuroendocrine cells in the hypothalamus, or whether these genes are acting post-natally in either the electrical or endocrine activity of these cells, however, this collection of genes suggests that the neuroendocrine function of the hypothalamus is a target of imprinted gene action in the brain (although whether this is some form of enrichment has also not been assessed).

There is one more behaviour which: is commonly associated with imprinted genes, is regulated by the neuroendocrinal hypothalamus, was one of the first discovered phenotypes and is the focus of this thesis, parenting behaviour. Several imprinted genes have been shown to directly impact maternal care behaviour (Lefebvre et al., 1998, Li et al., 1999) i.e., the coordination of mothers to care for offspring. This relationship, in theory, fits with the wider phenotypic profile of imprinted genes, since correct parental care (providing protection, warmth and nutrition) is essential for offspring growth and development, yet explaining why imprinted genes influence this behaviour is less straightforward than the embryonic phenotypes since these genes are the effect of grandparents, coordinating their offspring, to care for their grandchildren (Haig, 2014). A full consideration of imprinted genes and parenting will take place in Section 1.3 but first I will present a full overview of the biology of parenting.

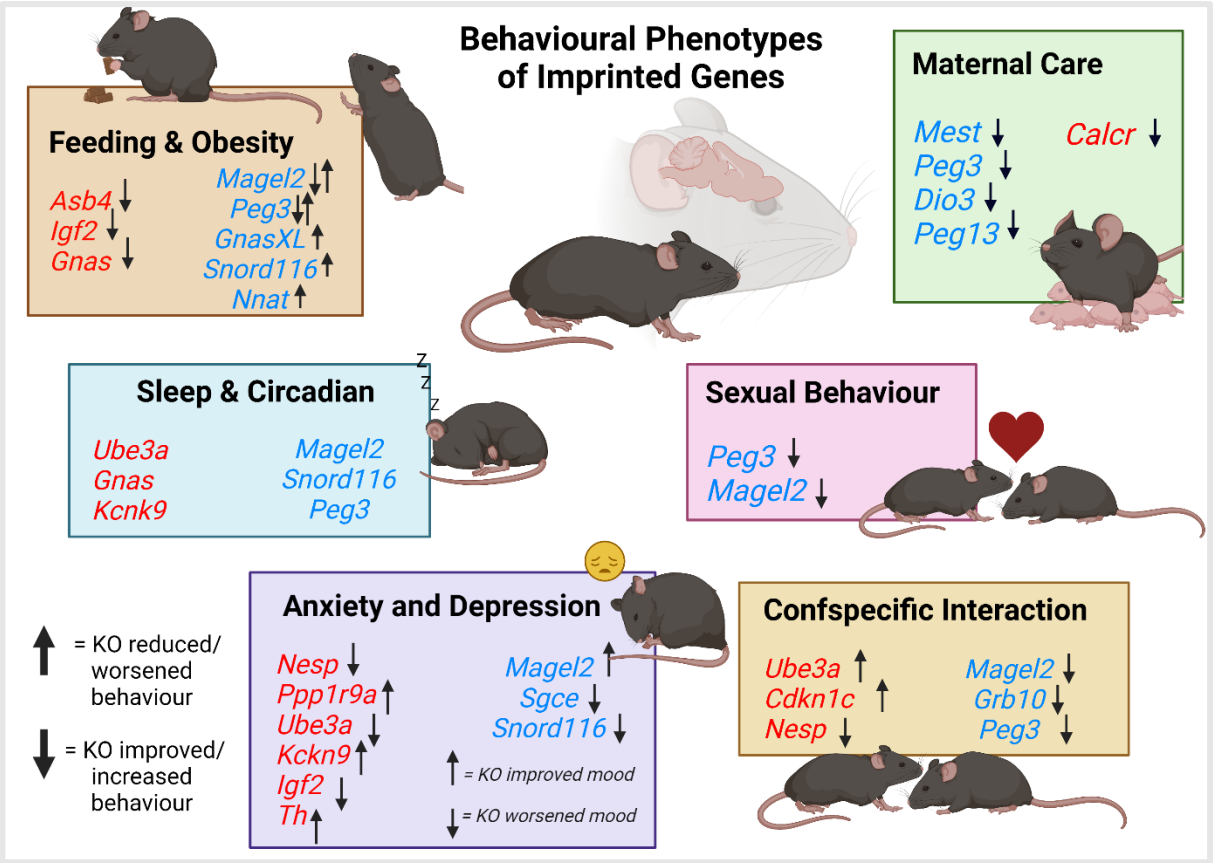


Figure 1.2. Imprinted gene brain and behaviour phenotypic summary. Imprinted genes are listed with behavioural phenotypes shown in mutant mouse models, either from a Knockout (KO) of the gene or an over-expression model. The direction of the arrows indicated whether the model showed a decrease or increase in the relevant behaviour. No arrows are shown for circadian behaviours as the direction of change is not able to be quantified. PEGs are shown in blue and MEGs in red. For references and further details, please refer to the text in Section 1.1.5.

1.2 Parenting Behaviour

1.2.1 What is parenting behaviour?- General definition and perspectives

A broad, species-wide, definition of parenting behaviour describes it as a multifaceted social behaviour that is directed towards the survival and optimal development of offspring, often at the expense of the individual producing the behaviour (Gross, 2005, Numan, 2017). Parenting in this case can be performed by any individual motivated to improve the survival of an offspring, this naturally includes mothers and fathers of the offspring but also alloparental figures too (such as other relatives, friends and community members) (Rogers and Bales, 2019).

From a behavioural perspective, parenting is of great interest since it is composed of multiple, stereotypic, species-specific behavioural components that together work to promote offspring

survival (Kohl, 2018). Parental behaviours carried out by organisms to enhance offspring survival can generally take one of the following forms: gamete provisioning, nest-selection and ovi-positioning (egg-layers), nest/burrow building, egg attendance and brooding (egg-layers), offspring attendance and brooding, offspring food provisioning, care after nutritional independence and in select mammal species, care for mature offspring (Kölliker et al., 2014). For a more species-specific example, in mice, parenting consists of displays of retrieval behaviour, chemo investigation, licking/grooming, crouching, nursing and nest building. Each behaviour can be seen as a separate component, yet it is the collective of behaviours we term parenting (Dulac et al., 2014).

From an evolutionary perspective, parenting is one of the rare behaviours with a clear conserved/innate basis in humans. Offspring survival (and hence gene propagation) is essential to natural selection and if parenting behaviour is necessary to survival of that species' infants, it will be heavily influenced by evolutionary mechanisms. However, parenting is not universal across the animal kingdom with different taxa displaying this behaviour at varying frequencies and some species have thrived and evolved without needing adults to commit themselves to the survival of offspring (Kölliker et al., 2014). Within each taxon, it is evident that this behaviour is deeply conserved. Mammals are often considered the most invested parents and the essential nature of parenting to promote offspring survival suggests that the neural circuits and mechanisms to execute and motivate this behaviour in mammals are also highly conserved (Kohl et al., 2017).

From a neuroscience perspective, parenting presents a rare scenario in which a high convergence of neural mechanisms is expected between humans and other mammals. Furthermore, the work carried out thus far to elucidate the mechanisms of parenting behaviour in mice and other mammals has shown exactly what was predicted evolutionarily, a highly conserved, well-defined neural circuit presenting a clear well-refined set of targets in the brain to further understand how the brain produces a complex multi-faceted behaviour such as parenting but also a target for any form of intervention to address parental attachment problems in human beings.

When considering all behavioural components that comprise parenting behaviour, each one can be understood as a specific adaption to deal with one or more ecological problems the offspring/parental figure may be facing. Some behaviours take on special importance in parenting. Nest building, as an example, is a common behaviour in birds and some mammals

for warmth and protection; however, nest building should only be considered a parental behaviour when the animal builds the nest to enhance offspring survival e.g., overcome the ecological problem of keeping offspring away from predators and in thermal homeostasis.

Parenting behaviour itself is 'innate' since the motivation and desire to interact with infants is rarely rationally decided upon but it is still sensitive to environmental insults and the variety of factors that can alter parenting and offspring outcomes is vast. Why and how parenting behaviour came to be are fundamental questions for biology and tend to be answered from two main perspectives, proximate and ultimate. I will begin with a discussion of the ultimate origin of parenting, namely the situation that has caused its evolution across the animal kingdom and how mammalian parenting evolved specifically.

1.2.2 Why do animals perform parenting behaviour? – Evolutionary origins & Ecological varieties of parenting

Parenting behaviour is beholden to evolutionary pressures, but unlike the evolution of imprinted genes, which have emerged and remained consistent in different lineages, parenting behaviour has evolved and changed between and within lineages. This is due to the pliable nature of parenting. George Williams stated that "Natural selection favours the evolution of behaviour that will maximize lifetime reproductive fitness" (Williams, 1966). The most valuable form of fitness is the number of biological offspring an individual begets (often termed personal fitness) and so a behaviour (or more specific the neural capacity to perform a behaviour) will evolve in situations in which the fitness benefits of carrying out the behaviour exceed the costs of performing. On the surface it would appear as if parenting should be the same in all lineages, because it is in the organisms' selfish genetic interest to care for offspring that cannot survive by themselves, even at the expense of their own time and resources (Gross, 2005).

What we find is that parental investment is not a fixed constant in parental animals but will vary according to previous reproductive success, other mating opportunities, the security of the environment and the viability of the current litter (Gross, 2005). It is also not constant within classes of animals with different mammal's showing varying levels of concern. One does see a more consistent picture when comparing classes of animal as a whole (Figure 1.3). The only universal for all animals when it comes to getting offspring into the next generation is pre-birth nourishment (Kölliker et al., 2014) and this is where some species of fish, amphibian and reptile end their parental investment (Doody et al., 2013). If these classes do

show parental care, it will be in the form of egg attendance, transport and offspring attendance and transport (i.e., tadpole transport) (Kölliker et al., 2014). Another egg-laying class with much higher parental investment is birds, the offspring of birds hatch from eggs but are born in a more vulnerable state (blind, featherless and altricial) compared to reptiles who are born able to regulate temperature and forage from birth (Doody et al., 2013, Case, 1978, Burley and Johnson, 2002). As a consequence of these vulnerable offspring, 82% of species display biparental care as it takes two adults to provision a full nest of birdlings (Cockburn, 2006).

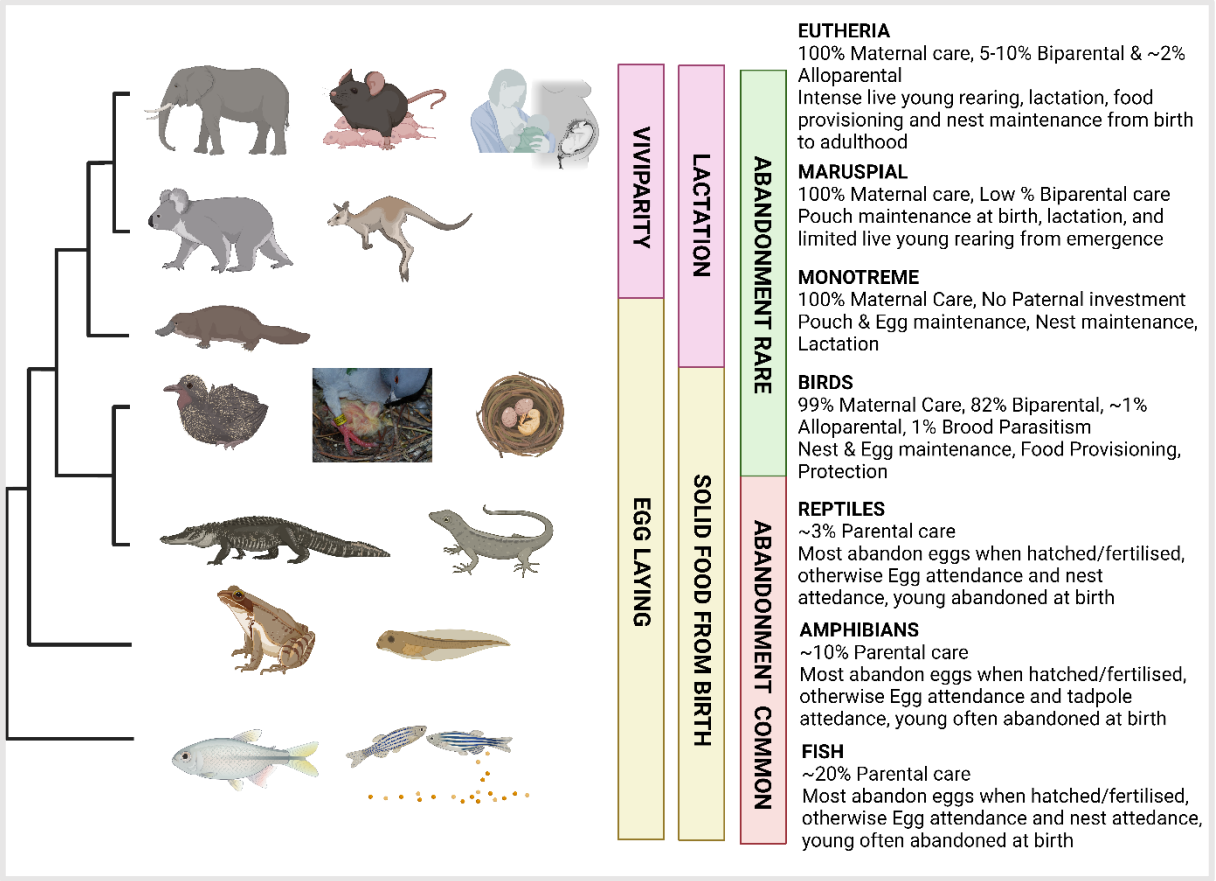


Figure 1.3. Parenting in the animal kingdom and the difference between groups of mammals. Traits between different classes of animals and clades of mammals are contrasted. Viviparity & Placentation vs. Egg Laying, Lactation and Mammary Glands vs Solid feeding from birth and proportion of species that display abandonment at birth. Levels of maternal/biparental care are considered as well as class typical parenting behaviours.

Mammalian parental care is the focus of this thesis as it is the only vertebrate lineage in which imprinted genes have also evolved. Mammals face similar challenges to birds since many species give birth to altricial and highly dependent offspring requiring high parental investment. This appears to be the primary driver of parental care evolution and the benefits of altricial development appear to be the longer period of nervous system development the

offspring gain which means more time to adapt themselves to the environment they are being born into (Faust et al., 2020, Scheiber et al., 2017, Charvet and Striedter, 2011). Mammals uniquely experience long gestation, viviparity, placentation and post-natal feeding from the mammary gland, which all impact/rely on the female and puts mammalian mothers into an intimate bond with their offspring that isn't recapitulated anywhere in the animal kingdom. Consequentially, maternal care is present in 100% of mammalian species (Rogers and Bales, 2019). It is for this reason that mammals are seen as having the most intense parental investment of any species with many mammals maintaining their parent-offspring connections into the formation of herds and social groups.

1.2.3 Do all individuals parent the same? – Mammalian Mothers, Fathers and Others

Taking on the nutritional, protective and developmental concern for another organism is a vast commitment for an animal to make. More importantly, from the gene's view, investing in offspring consumes somatic resources and time so that it impacts the future reproductive success of the animal. What an animal invests in a particular offspring cannot be recuperated and so all individuals will calculate how much investment it is worth making and which individual does the parental investing is a common source of conflict. Across the animal kingdom, some species rely on both parents to provide care and take on the somatic burden (e.g., birds and some mammals), some species have maternal dominant care (e.g., most mammals), others paternal dominant (e.g., some fish and amphibians) and a rare selection of species even involve investment from outside the parental dyad (e.g., humans, elephants). It is worth considering the different nature of these parental figures – mothers, fathers and the wider community – within mammalian parenting as a combination of these figures form the earliest social network for offspring.

Mothering in mammals is the most reliable and predominant form of parenting. The mother is solely responsible for provisioning the offspring pre-birth. Unlike birds and reptiles, the offspring are maintained inside the female's body, and for marsupials, within a maternal pouch as well. During this time, the offspring derive nutrition from the foetally derived placenta or from the egg sac and there is little the father or other individuals can do pre-birth to provision the offspring outside of supporting the mother. Bird and reptile females do produce eggs and hence have a large investment to begin with, but as soon as these eggs are laid, responsibility for their protection and warmth can be shared. Unfortunately, mammalian

mothers are trapped in a sunk cost at the birth of their child. They have already dedicated a large amount of resources to the offspring and cannot abandon them at birth otherwise they lose that investment. Additionally, internal fertilization means that only the mother can be certain of maternity. This combination of factors results in 100% of mammalian species demonstrating maternal care (Rogers and Bales, 2019). This maternal dedication does not hold for other groups of animals. Birds mirror mammals in their maternal dedication with only ~1% not displaying some form of maternal care (Cockburn, 2006) but within reptiles, amphibians and fish, it is much more common to find absent maternal care, resulting in the father stepping in to take on the bulk of the responsibility (Gross, 2005).

Fathers are an interesting party in the parental debate since they display a wider variation of offspring engagement strategies, from aggression, to disinterest, to highly doting. Though mammalian fathers naturally have the same motivation for the survival of their offspring as mothers, the quantity and consistency of their care varies heavily from species to species, depending on the specific environment and opportunities they have (Gubernick, 2013). As already discussed, to parent or not to parent is a calculation based on fitness. For mammals (unlike egg-laying species), the mothers, at birth, are shouldering a large sunk cost hence the fathers have an opportunity to abandon the litter knowing the mother will shoulder the burden alone. The father's personal fitness will be enhanced by abandoning the current litter and impregnating other females. However, this is not always a valid strategy. In some species such as birds and some mammals, the offspring are too demanding to take on alone and it becomes necessary for another individual to intervene, and the father is the next most motivated individual to aid the survival of his own offspring. As might be expected, biparental care in mammals is relatively low with an estimated 5-10% of mammalian species displaying it (Rogers and Bales, 2019). In comparison, biparental care is estimated to exist in 82% of bird species (Numan and Insel, 2003, Cockburn, 2006). Regardless, in species with paternal care, the behaviour displayed by fathers tends to be identical to mothers exempting behaviours requiring female anatomy (nursing and egg laying). Some species of bird and a large proportion of reptiles, amphibians and fish do display male-only parenting systems so it is by no means certain that the mother will be burdened with the offspring (Gross, 2005).

One of the things that makes fathering behaviour interesting in mammals is that many virgin males are hostile to infants and so must perform a behavioural switch to become parental (Rogers and Bales, 2019), and this brings us on to a discussion of virgin mammals and their relationship to offspring. In most species, animals other than the mother and father will not

interact with conspecific offspring, and when interaction occurs, it is usually hostile. This is particularly true in male mammals capable of paternal care, and so fathering offspring then necessitates a transition to offspring care, distinct from the default response of aggression. Infanticide is a behavioural strategy that can be adopted by male mammals and is fitness promoting in harem style situations in which reproduction is monopolized by a minority of males (Lukas and Huchard, 2014). Infanticide brings about two fitness promoting consequences the removal of offspring that are not your own and the promotion of fertility in the females as a result of cessation of lactation. The transition to parenting behaviour from this position of hostility requires mating with a female and the reassurance that offspring surrounding the male will have a chance of being its own (Rymer and Pillay, 2018, Numan and Insel, 2003).

Offspring being too demanding for one parent doesn't necessarily demand commitment from the father. Alloparental behaviour or communal nesting is known in 120 mammal species and 150 avian species and often involves the more distant relatives of offspring to provide some care but can also involve community style rearing of offspring in which genetically unrelated individuals will nurse or 'babysit' offspring (Riedman, 1982, Kenkel et al., 2017). Although this appears a display of altruism, all alloparenting can still be understood as a fitness calculation. 'Inclusive fitness' was termed by W. D. Hamilton (Hamilton, 1963, Hamilton, 1964), and this is an expanded view of fitness suggesting that individuals not only have their own offspring output to add to their fitness total but the offspring output of any individual's that have genetic overlap with yourself. Nieces, nephews, grandchildren and even children of cousins twice removed will all have a certain percentage of genetic overlap with the individual and the survival of these children will also add to one's fitness. Therefore, the consistent manifestations of parenting in relatives other than mothers and fathers can be understood this way and will be selected for in situations where the demand of offspring requires the resources of more than one individual. The interaction of non-relatives with offspring can be understood within the lens of reciprocal altruism, since babysitting and communal nursing are tradeable resources and providing the service now may benefit you in the long term when the favour is recalled.

Within mammals, it is widely recognized that the level of parental investment has dramatically increased from the commitment of the other classes of animals. One interesting consequence of the broadening burden of parenting is the associated impact on the species in general. Mammals tend to be much more social than reptiles, fish and amphibians, living in

herds and packs with offspring present amongst the group. This is generally seen to be a consequence of an increase tolerance to the presence of other animals. The social bond between parent and child seems to be the first social connection a child will form and the intricate bond between mother and child in mammals seems to broaden into consequences for mammal species generally. Mammals have a higher infant receptivity in general with non-related individuals (particularly females) not reacting with hostility at all to infants. In species such as mice and humans, it appears as if the co-adaptation between adults and infants has got to such an extreme extent that infant contact operates as a reward in and of itself (Ferrey et al., 2016, Ferris et al., 2005, Mattson and Morrell, 2005). With this in mind, parenting is not just a strategy for increasing fitness but a tool by which species have been able to reshape their social landscape and increase their tolerance to allow group style living and behaviours.

As already discussed, parenting is a costly behaviour and the larger your genetic overlap with the child, the larger the benefit and the associated cost one will be willing to bare. Hence within mammals, mothers and then fathers still endure most of the cost with minor assistance when alloparents are available. They also display motivation levels that exceed alloparents. Humans are one of the rare mammal species displaying both paternal care and alloparental care as well as several species of rodent. One of the fascinating neurobiological questions for this area of research is, do these separate groups of animals share a proximate mechanism for this behaviour? Or does the mechanism differ for a behaviour which is essentially performed the same by mothers, fathers and others? Recent advances in neuroscience have elaborated the mechanisms for this behaviour rather extensively in mammals such as rodents, sheep and humans and will be summarized below.

1.2.4 What are the proximate mechanisms of mammalian parenting? – The Parental Brain

Although many insights discussed in the previous section apply to mammals generally and other species, the current section focused on parenting in rodents for a number of reasons. Firstly, rodents have been used as the primary experimental model for the majority of parenting studies to date. Secondly, mice are the experimental animal in which imprinted genes have been shown to impact parenting and finally, as mice were used to model parenting behaviour in the experimental chapters of this thesis.

Neural mechanisms of parenting

Parenting is made up of many different sub-behaviours and requires a large number of distinct neural mechanisms (Numan, 2017, Kohl et al., 2018). The core circuitry for parenting i.e., the circuit necessary to prompt parental response upon exposure to pups is deeply conserved amongst eutherian mammals and between the sexes (Kohl et al., 2017, Kohl and Dulac, 2018, Rogers and Bales, 2019). In the following section I have described the common neural circuity for parenting behaviour.

The predominant figure in the mapping of the core parenting behaviour circuitry was Michael Numan. The work identifying this circuitry has been performed predominantly in mice and rats but work with humans and with other mammals has found a heavy degree of convergence. In 1974, using rats, Numan (1974) identified one specific area of the hypothalamus as essential for parenting behaviour – the medial POA (MPOA). Ablation of the MPOA removes parenting behaviour (Lee et al., 1999) and stimulation of this region can produce parenting (Rosenblatt and Ceus, 1998). Recent advances have shown that one specific neural group with the POA act as the 'parenting hub' and those are neurons expressing galanin (*Gal)* (Wu et al., 2014) alongside other markers such as *Th, Calcr* and *Brs3* (Moffitt et al., 2018). Galanin neurons are the most active in the POA during parenting in mothers, fathers and virgin females (Moffitt et al., 2018) and optogenetic activation of these neurons will produce parenting behaviour, even in innately avoidant animals (Kohl et al., 2018). Ablation of galanin neurons on the other hand, has been shown to remove parenting behaviour (Wu et al., 2014).

Later studies have shown that the MPOA neurons activated by pups project to monoaminergic and somatomotor brainstem nuclei (Numan, 1988, Numan et al., 1977, Numan and Smith, 1984, Numan and Stolzenberg, 2009). The MPOA receives input from the major sensory regions of the brain and the primary output of the MPOA links to the dopamine circuity and in turn the motor circuitry to initiate the behaviour. There is also direct inhibitory activity on an innate avoidance circuit regulated by the medial amygdala (MeA) (see Figure 1.4A) (Numan, 2017). Recent work by Kohl et al. (2018) has mapped this circuitry in exquisite detail (See Figure 1.4B) using retroactive viral tracing and found Numan's circuit to be the core circuitry with additional minor circuits co-ordinating specific inputs and outputs. The MPOA has been shown to be intensely hormonally sensitive (Bridges, 2015, Bridges, 1996), expressing hormone receptors - *Prlr, Pgr* and *Esr1* (Consiglio and Bridges (2009), Ribeiro et al. (2012), and an injection of pregnancy hormones into the MPOA has been shown to be enough to produce parenting (Numan et al., 1977,

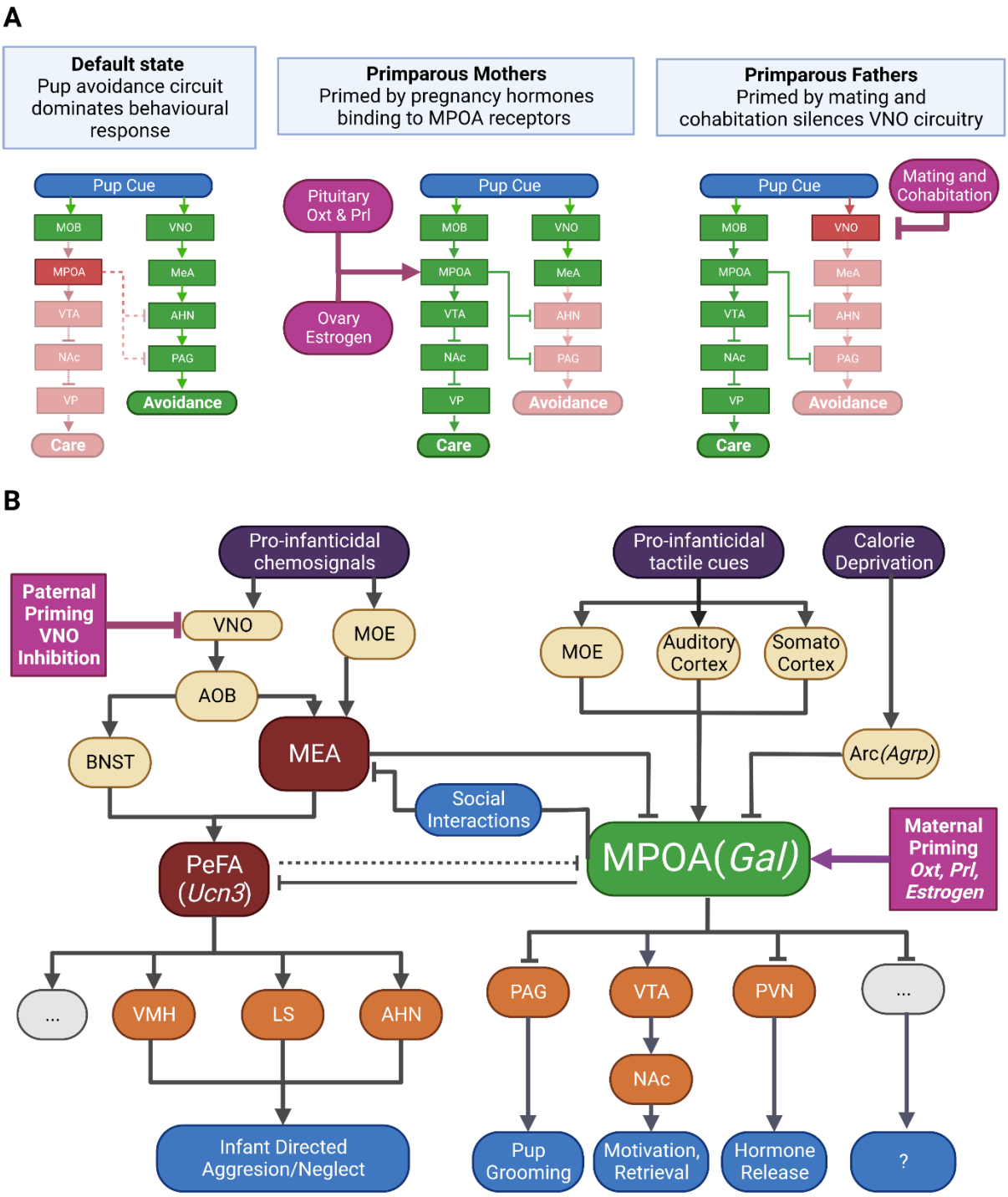


Figure 1.4. Circuits for Care – Parental neural circuitry in mice. (A) Simple parental circuit adapted from Numan (2017) showing the major brain regions in the care and avoidance circuits isolated through the work of Numan and others over the past 50 years. This basic circuit of brain regions was identified from systematic lesion studies and classical pharmacological studies primarily in female rats. The events that occur from primiparous mothers and fathers to activate this circuit and switch the default behaviour of the rodents away from avoidance to care is indicated in pink. (B) Sophisticated parenting circuit adapted from Kohl (2020) showing many more regions involved in parenting and the specific behavioural outputs. This refined circuit was derived from retrograde and anterograde viral tracing from the MPOA, *c-Fos* response and Optogenetic stimulation, primarily in mice. Points of maternal and paternal priming are shown in pink.

Bridges et al., 1990) highlighting the powerful priming role that pregnancy hormones play in the maternal brain.

The MPOA is not the only region of the brain to play an important regulatory role in parenting. Serotonin neurons projecting from the Raphe nuclei have been shown to be fundamentally important (Olazábal et al., 2004, Barofsky et al., 1983). Oxytocin and vasopressin neurons from the PVN have been shown to play a regulatory role in maternal and paternal bonding and have also been shown to project to the MPOA (Caldwell, 2017, Yoshihara et al., 2017, Numan and Corodimas, 1985). Finally, more unexpected regulators of parenting have been observed such as the *Agrp* expressing neurons from the arcuate nucleus, which promote hunger and have an inhibitory effect on parenting (Boillot, 2019). All the neurons mentioned in this section input into the MPOA, and act to regulate the hub region's output of parenting behaviour (Kohl et al., 2018, Kohl and Dulac, 2018).

Murine Mothers/Fathers/Others

Interestingly, in mice and other mammals, the core parenting circuitry holds remarkably similar between different groups capable of parenting behaviour, with the mother, father and alloparent mice all utilising the same core MPOA circuitry to engage in parenting (Rogers and Bales, 2019). Hormones such as prolactin have also been shown to have a consistent impact of parenting, not just in mothers but also in fathers (Grattan, 2015, Smiley et al., 2022, Stagkourakis et al., 2020). However, what does appear to vary between groups is the route to activation of this circuit.

Mothers are the predominant care givers in mice and have a myriad of influences occurring pre-birth to prime the mother for this. The maternal blood stream has varying levels of progesterone, oestrogen and prolactin during pregnancy (as well as placental lactogen from the placenta) and it has been robustly shown that these circulating hormones cross the blood brain barrier and bind to receptors within the MPOA (Kohl et al., 2017, Dulac et al., 2014). These hormones have been shown to induce a priming effect, causing pregnant mice respond to pups without prior training or exposure. Though increased exposure and learning does make mother mice better at parenting, first time mothers still show the effects of hormonal priming and are intrinsically motivated to interact and care for their pups from birth (Stolzenberg and Champagne, 2016, Bridges, 2016, Bridges, 2015).

Father and virgin mice do not have pregnancy hormones to guide their parental responses (Bailey and Isogai, 2022). Virgin male mice are actually hostile to pups and will commit infanticide (Vom Saal and Howard, 1982, Dulac et al., 2014). The only way to reliably produce parenting behaviour in male mice is to allow them to mate with a female and then ensure they cohabit with the mother throughout the duration of her pregnancy (Kuroda et al., 2011, Tachikawa et al., 2013, Elwood, 1985). A few days before birth (coinciding with a large hormonal change in the mothers) the fathers' behaviour switches from aggression to care behaviour and he will perform all the same parenting behaviours as the mother, apart from nursing. Motivation in fathers is intrinsically lower than mothers, but the performance of parenting persists until pups reach maturity at which the father will revert to his infanticidal ways (Dulac et al., 2014). Virgin female mice, on the other hand, are innately interested in pups and though infanticide has been observed, given enough exposure to pups, virgin females have been shown to spontaneously display maternal behaviour (Martín-Sánchez et al., 2015, Stolzenberg and Rissman, 2011). This is an example of the increased tolerance and inherent rewarding property of infants seen in eutherian mammals. Once virgin females have initially displayed spontaneous maternal behaviour, greater exposure to the pups will reinforce and improve this behaviour until it is reliable.

External influences on parenting

In mice, as in all mammals, a selection of environmental and contextual elements can directly affect an animal's ability to produce parenting behaviour. Performance issues of the parental mice can create parenting deficits by association. For example, olfactory discrimination is key to identifying their offspring and, in all animals, appropriate sensory and motor function are critical to manifesting the behaviour successfully (Dulac et al., 2014). Serotonin and dopamine have direct modulatory elements affecting parenting but mood in general, such as anxiety and depression can affect a parent's motivation and ability to produce parenting (Kessler et al., 2011) such as the impact of post-natal depression of maternal bonding (O'Higgins et al., 2013). Finally, for mothers in particularly, hormonal imbalances could alter her 'priming' for pregnancy

The other major external influence on parenting is offspring behaviour. Parenting is a two-way relationship and although parents are primed to care for offspring. Offspring also deploy their own behavioural mechanisms to demand care. In mice, one of the primary behaviours that pups use to signal for care are USV's (Portfors and Perkel, 2014) while in humans,

infants cry, scream and use a variety of auditory and visual cues to demand adult attention. A reduction in the exposure of parents to offspring engagement strategies has been shown to influence parental care both in mice where pups show USV deficits (McNamara et al., 2018a, Bosque Ortiz et al., 2022, Bowers et al., 2013) and in humans, whereby maternal responsiveness has been linked to infant vocalisation (St James-Roberts et al., 1998).

One unique way that offspring have been shown to prime mothers is through the foetally-derived placenta. Prolactin in mammals is one of the major hormones promoting maternal care, it's expression peaks late on in pregnancy and maintains high during lactation (Grattan, 2015). It coordinates a number of changes in the maternal body, which have been described as preparing the body for the first stages of motherhood and one of these changes is priming the POA circuit for parenting (Smiley et al., 2022, Brown et al., 2017, Bridges et al., 1990). Placental lactogens, a prolactin-like hormone, are produced by the spongiotrophoblast cells in the placenta (placental lactogen) which bind to prolactin receptors and influence the same phenotypic outcomes, including modulating maternal care (Bridges et al., 1996, Bridges and Freemark, 1995).

Overall, it has been demonstrated that there are many mechanisms influencing parenting beyond the core POA circuitry, however the majority of these mechanisms still signal to the MPOA either via hormonal-priming or direct neural stimulation. This has highlighted that the POA hub is essential for this parenting behaviours to be performed. Imprinted genes have been shown to impair maternal care when genetically altered in mothers (Li et al., 1999, Lefebvre et al., 1998) and have even been shown to impact the parenting of wildtype (WT) mothers through genetically altered offspring (McNamara et al., 2018a). The following section reviews what we know concerning genomic imprinting and its effect on mammalian parenting behaviour.

1.3 Genomic Imprinting & Parenting Behaviour

As highlighted at the ends of Sections 1.1.5 and 1.2.4, parenting behaviour and imprinted genes appear to be interconnected, as suggested from research which has explored brain and behaviour phenotypes of some well-known imprinted genes. Not only do imprinted genes impact the maternal brain but they can also impact parenting through effects outside of the brain (See Figure 1.5) (Creeth et al., 2019). In this section I have first summarised the

external influences of imprinted genes on maternal behaviour before moving onto the focus of this thesis, the role of imprinted genes in the parental brain.

1.3.1 Which imprinted genes affect parenting externally? – Solicitation of Maternal Care

As previously stated, maternal care is a bi-directional social behaviour which is not only dependent on the mother's ability to engage in care but also in the offspring's ability to solicit and partake in that relationship. As the mammalian parenting ability has evolved, a convergent evolution of offspring behaviour has occurred alongside. Three imprinted genes have been shown to regulate another innate behaviour, the production of USV's in neonatal mice (Jiang et al., 2010, McNamara et al., 2018a, Bosque Ortiz et al., 2022). USV's in mice are produced from offspring when exposed to low temperatures or prolonged separation from their mother, indicating to the mother that they have been displaced from the nest and need returning. Two PEG's, when knocked out, result in a USV's deficit in the offspring and in turn, have been shown to affect the retrieval times and preferences of WT parents (McNamara et al., 2018a, Bosque Ortiz et al., 2022) and this has even been shown to be temperature dependent and related to oxytocin functioning (Da Prato et al., 2022). Interestingly, and in line with conflict theory, the only MEG to affect USV's, *Ube3a*, actually caused an increase in USV emittion when distrupted (Jiang et al., 2010).

Imprinted genes also impact the need for parents to thermoregulate (by crouching over) their offspring and how often/regularly they would need to nurse, both essential parenting behaviour in mice. Imprinted genes have been shown to affect offspring thermoregulation and temperature (Curley et al., 2005, Paulo et al., 2018, Charalambous et al., 2012, Van De Pette et al., 2016), this would be associated with a higher demand of maternal care. Imprinted genes have also been shown to regulate metabolism during pregnancy (Cleaton et al., 2016, Sferruzzi-Perri et al., 2011, Cassidy and Charalambous, 2018) which coordinates mother's ability to thermoregulate her pups with their own temperature. When it comes to suckling behaviour., several PEGs have shown a deficit in suckling behaviour in pups (Schaller et al., 2010, Curley et al., 2004, Plagge et al., 2004, Cowley et al., 2014). These same genes have also been shown to impact the mother's ability to let down milk (Curley et al., 2004, Cowley et al., 2014).

Imprinted genes are well known regulators of endocrine function and deficits in reproductive hormones have been reported in some imprinted loss of functions models such as *Mkrn3*

(Abreu et al., 2013). These hormones are essential for priming the maternal circuitry, but it has yet to be tested whether *Mkrn3* has an impact on parenting or whether other imprinted genes have the same impact. However, one clear example of imprinted genes hormonally regulating parenting behaviour is via placental hormonal signalling. Imprinted genes have a fundamental role in placental development (John, 2017, Monk, 2015, Tunster et al., 2013) and it is widely accepted that the placenta is the primary tissue for imprinted gene action. Several imprinted genes have been shown to regulate spongiotrophoblast development and the placenta's ability to circulate placental lactogen into the mother's blood supply (Tunster et al., 2018, Tunster et al., 2016a, Tunster et al., 2016b, Tunster et al., 2013, Tunster et al., 2011). Within this context, one, placental-specific MEG has been shown to enhance parenting behaviour when disturbed (Creeth et al., 2018) while a PEG has been shown to worsen maternal care when only interrupted in offspring (McNamara et al., 2018a) and has been shown to impact spongiotrophoblast function (Tunster et al., 2018).

The phenotypes described in this section are all important for parenting when considering the bidirectional nature of resource provisioning from the mother. The following section explores how imprinted genes have been shown to directly impact parenting behaviour, by their action in the maternal brain/body.

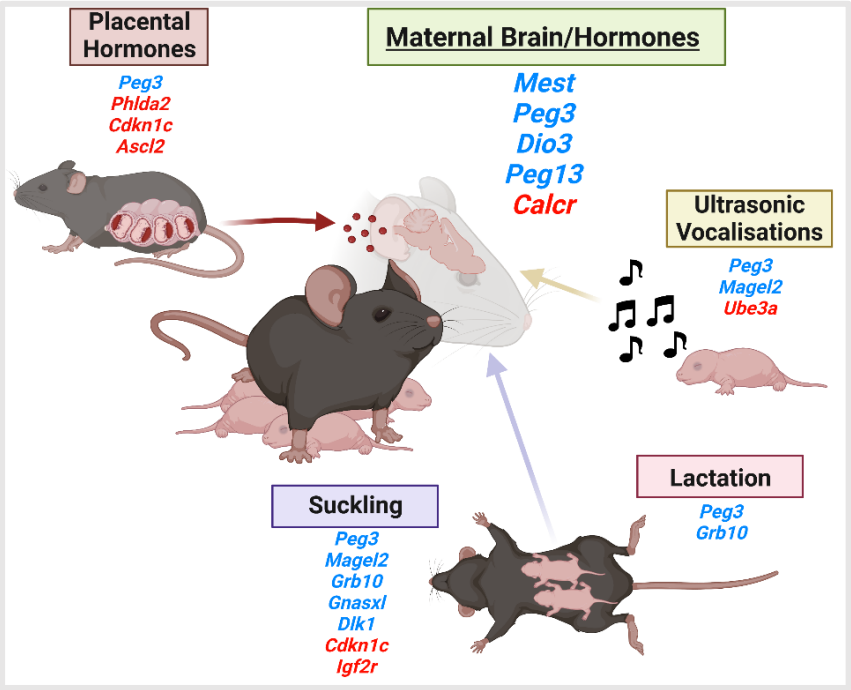


Figure 1.5. Imprinted genes with impacts on maternal behaviour. Imprinted genes are listed with a demonstration of the associated phenotype in a mouse model. References and details of whether the model enhanced or disrupted the behaviour can be found in Sections 1.3.1 and 1.3.2. PEGs are shown in blue, and MEGs are shown in red.

1.3.2 Which imprinted genes affect the maternal brain? – Maternal care phenotype summary

To date, of the ~ 260 imprinted genes that have been identified, only six have been assessed for their impact on parenting behaviour, five of which have shown a parenting phenotype within mothers of global knock out mouse model (Table 1.1) however the role these genes play within the parental brain is still unclear. Of the five genes, C*alcr* is the only maternally expressed gene to be assessed, and has been found to be one of the primary markers of parenting neurons in the POA (Moffitt et al., 2018). It has also been recently shown to elicit parenting deficits by reducing nursing behaviour and risk-taking retrieval behaviour in murine mothers (Yoshihara et al., 2021) as well as reducing parental tolerance and carrying behaviour in marmosets (Shinozuka et al., 2022) but the authors did not indicate they were aware of *Calcr* status as a brain-specific imprinted gene (Hoshiya et al., 2003). To date, *Grb10* is the only imprinted gene which has been assessed for a role in parental care but has not demonstrated a deficit. A caveat to this however is that this study only compared a paternal knock-out (KO) to maternal KO, and a *Grb10* comparison with WT mice remains yet to be carried out. Nonetheless, this model demonstrated an impact on nutrient supply to the offspring and fat deposition (Cowley et al., 2014). Table 1.1 provides an overview of each of the five imprinted genes associated with parenting behaviour but the remaining four will be considered in detail below.

Mest

Mesoderm specific transcript (*Mest* or *Peg1)* was identified as an imprinted gene by Kaneko-Ishino et al. (1995) and was found to be expressed throughout embryo tissues though adulthood expression was shown to be restricted to the nervous system – notably the hypothalamus, amygdala, olfactory bulb and hippocampus. The KO model of *Mest* (*Mest*^{tm1Masu}) was created using an IRES-βgeo-lacZ cassette which was inserted into exons 3-9 and was maintained on a 129/Sv background. The original investigation by Lefebvre et al. (1998) found that mice were viable and fertile, but it was noted that pups and placenta were born smaller and lighter than WT litter mates. KO dams failed to raise pups, only maintaining an 11.3% pup survival in first pregnancy, though this effect was rescued by cross-fostering rescued. KO dams had a deficit in placentophagia with 84.6% of litters having at least one pup abandoned in its extraembryonic tissues and in a 15-minute retrieval and nest building assessment, only 3/8 dams displayed nest building and there was no attempt to retrieve pups.

Olfactory performance was not significantly different between mutants and WTs. These findings were replicated using the original model by Ineson et al. (2012) who found breeding difficulties with their females with offspring dying from maternal neglect, but the phenotype attenuated during second pregnancies.

Table 1.1. Overview of the specific parenting deficits observed in the imprinted genes associated with maternal care. Five imprinted genes (*Mest, Peg3, Dio3, Peg13, Calcr*) have been shown to cause maternal care deficits when disrupted in mothers only. If these genes have shown deficits in the specific parental behaviours in this table, then an 'X' is used, if the deficit has never been tested or observed then a '?' is used and if a deficit has been tested and observed to be equivalent to WT, a 'U' is used for Unaffected. Only studies that found parenting deficits are detailed here, contradictory studies are discussed in the text of Section 1.3.2

Imprinted gene (References)	**Litter Survival**	**Placent ophagia**	**Retrieval**	**Nest**	**Pup Interact ion**	**Lactation & Nursing**	**Attenuated with second pregnancy**	**Neural Mechanism?**
Mest (Lefebvre et al., 1998 & Ineson et al., 2012)	X	X	X	X	?	U	X	?
Peg3 (Li et al., 1999; Curley et al., 2004; Champagne et al., 2009)	X	?	X	X	X	X	X	*Oxt*
Dio3 (Stohn et al., 2018)	X	?	X	X	X	?	X	*Oxt/Avp*
Peg13 (Keshavarz and Tautz, 2021)	X	?	X	?	?	?	?	*5-HT*
Calcr (Yoshihara et al., 2021)	X	U	X	X	X	X	?	POA

Follow up studies using different models have failed to replicate the maternal care findings described above. Beechey (2000), using a model that duplicated the maternal copy of *Mest*, saw the growth retardation phenotype but no deficits in retrieval, nest building or litter survival (PatDp mice were functionally normal suggesting no consequence for over-expression of *Mest*). Most recently, Anunciado-Koza et al. (2022), have revisited *Mest's* effect on behaviour with a different mouse model (*Mest*$^{tm1.2Rkz}$) which was maintained on a C57BL/6J background and produces global inactivation with a floxed *Mest* allele, disrupting exon 3 only. Mutant virgin females and mothers showed no deficits in a 5-minute retrieval test with 2–4-day old pups and no anxiety or depression like behaviour. The failure to replicate a parenting deficit in this modern study could be down to: the difference in mouse strain (Kuroda et al., 2011), the modern sophisticated model that keeps genes such as *Mir-335* intact (which may be the origin of the maternal care deficits) but perhaps more likely is since this newer model doesn't involve the insertion of the LacZ and neomycin cassettes,

which has been shown to directly impact neuronal function in sensory neurons (Peck et al., 2021), the original model might be producing false results when analysed in a brain and behaviour situation, particularly for sensitive behaviours such as parenting. Either way, the relationship between *Mest* and parenting is fundamentally unclear.

Peg3

Paternally expressed Gene 3 (*Peg3*) was identified in 1996 (Kuroiwa et al., 1996) and was the second imprinted gene to have a maternal behaviour deficit characterized (Li et al., 1999) but is by far the most intensively studied. *Peg3* is a zinc finger gene and transcriptional repressor and is expressed in a variety of embryonic meso-endodermal tissues, in the hypothalamus and adult brain. Li et al. (1999) saw only 8% of litters from first time *Peg3* KO mothers reached weaning age (compared to 83% of WT), and since these offspring were functionally WT, they deduced the deficit had to be maternal in origin. Mutant mothers took 11 times longer to retrieve pups and 8 times longer to build nests and failed to display appropriate crouching behaviour in a 15-minute retrieval paradigm. Much like the recovery seen in *Mest* deficient mice, by the third pregnancy, mutant *Peg3* dams had a litter survival rate of 70% showing that maternal behaviour wasn't irrevocably impaired. Dams detected pups at the same time meaning the deficit was unlikely to be olfactory. The PVN and Supraoptic Nucleus (SON) oxytocin neurons were characterized in this model (which are known to stimulate lactation as well as modulate maternal behaviour acquisition) and mutants had ~33% fewer oxytocin neurons in these nuclei than WTs. Oxytocin phenotypes have been shown e.g., *Peg3* regulates expression of *Oxtr* (Frey et al., 2018) and the behavioural deficits on the same genetic mouse model have been replicated by Curley et al. (2004) and again by Champagne et al. (2009) who found different effects of *Peg3* on 129Sv and C57B/6J backgrounds but insults to maternal care in both.

However, just like in the *Mest* model, recent work has called the maternal care deficits into doubt. (Denizot et al., 2016) generated a new mouse model for *Peg3* that produced a truncated transcript that could not fold. The mice showed reduced postnatal growth, but this study found no effect on litter size, reproductive capacity and maternal behaviour. Pregnancy success and pup mortality was comparable to WT mice and within a standard retrieval paradigm in the home cage with 8 pups; mutant mice display comparable retrieval speed, crouching behaviour, nest quality and nest building latency. This was true for virgin females and primiparous mothers. Lactation and suckling behaviour were unaffected in their mutant

mice unlike previous reports and furthermore there was no disruption in circulating oxytocin nor any differences in oxytocin neuron number in the PVN. Just like in *Mest*, several reasons could explain this lack of deficit in the newer model, for one, the authors note their model is the first to disrupt the anti-sense RNA *APeg3* which is transcribed from exon 9 (others disrupt 5 or 6) which they disrupt. *APeg3* has only been shown to downregulate *Peg3* (Frey and Kim, 2014), yet it cannot be ruled out that some combination of *APeg3*, left isolated from *Peg3* causes maternal care problems, although I acknowledge that it is much more likely that the use of LacZ in the older models is to blame.

Dio3

Type 3 Deiodinase (*Dio3*) has paternally biased expression in most of the developing fetal tissues including the brain, placenta and skeletal muscles (Hernandez et al., 2002, Tsai et al., 2002). It is also expressed in the mature brain where it has a role in protecting the brain from thyroid hormones; importantly, expression of *Dio3* has been reported to be biallelic in mature brain cells (Hernandez and Stohn, 2018). However, this gene has been associated with maternal care deficits (alongside increased aggression behaviour) by Stohn et al. (2018). *Dio3* -/- mice were found to perform similarly to WT litter mates on social behaviour/olfaction tests but were distinctly more aggressive than the WTs and 85% of first-time dams failed to display retrieval behaviour compared to only 18% of WT, a deficit that, again, was mostly attenuated for the second litter. To investigate the basis of this phenotype, Stohn et al. (2018) investigated the oxytocin and vasopressin make-up of the *Dio3* -/- males and females. Females had a significant reduction in serum levels of *Avp* and *Oxt* (but not males) compared to WT, conversely males had significant reduction in *Oxt/Avp* messenger RNA (mRNA) in the hypothalamus (by 75%) compared to WT, but again this difference was not seen in the other sex. *Oxtr* and *Avpr1a* were higher in HET females compared to WT but not males and *Avp* was more highly expressed in *Dio3*-/- dams vs. WT dams vs. virgins. Hypothalamic *Oxt*/*Avp* mRNA was also higher in P5 males and females vs. WT. This sexually dimorphic *Oxt*/*Avp* alteration were accompanied by no significant difference in numbers of *Oxt*/*Avp* neurons in the PVN or the SCN in adults, meaning the deficit had not manifested in neuron numbers like the *Peg3* deficit.

Peg13

Peg13 is widely expressed across the body from the paternal allele but demonstrates its highest expression in the brain (Davies et al., 2004, Lorenc et al., 2014). Most recently,

Keshavarz and Tautz (2021) used a mouse model with a semi-deletion (deletion of 3' half) of the gene as the full deletion was semi-lethal and showed a severe skin lesion phenotype. Het- and Homozygous mice were fully viable with no morphological phenotype. Breeding was impaired with fewer successful litters sired from mutant male mice. Litter survival rate prompted the authors to assess maternal behaviour and within a 15-minute retrieval paradigm, mutant mothers (-/- and +/-) were significantly slower to retrieve 8 displaced pups to the nest than their WT comparisons. Many mutant mothers retrieved no pups in the 15-minute time period. The authors did not assess maternal behaviour further than this (so we do not know the impact of *Peg13* on other facets of maternal behaviour), but other behavioural tests show that these mutant mice also display higher anxiety-like behaviour in their open field and elevated plus maze as well as reduced locomotor activity in the open field. Both these deficiencies could easily account for the maternal deficits seen in retrieval behaviour in this study so a more comprehensive assessment of parenting would be needed before concluding *Peg13* affects this behaviour. However, the authors did provide some suggestion of a mechanism by which *Peg13* may be legitimately influencing maternal care through showing a reduction in gene expression key for the serotonin pathway (*Tph2* and serotonin receptors). As already stated in Section 1.2.4, serotonin plays an important role in parenting. More so, a host of imprinted genes saw reduced expression in the *Peg13* mutant mouse, further supporting the idea that these genes work in a network and could be influencing parenting behaviour through a combined effect rather than the impacts of a single gene.

1.3.3 Where does parenting fit with evolutionary theories for imprinted genes? – How to fit maternal care with conflict theory

The reluctance of parents to overcommit resources to offspring has influenced the evolution of imprinted genes (see Section 1.1.2) and for the evolution of parenting styles in animals (see Section 1.2.2). As discussed in Section 1.1.2, the leading hypothesis for the evolution of imprinted genes is conflict theory (Moore and Haig, 1991). Conflict theory is archetypally represented by the embryonic roles of *Igf2* and *Igf2r* on growth (DeChiara et al., 1991, Barlow et al., 1991), which has been suggested as the clearest demonstration of antagonist function of MEGs and PEGs. Imprinted gene influence on infant growth, size and metabolism tend to conform to the pattern that PEGs promote growth and consumption, while MEGs suppress (Iwasa, 1998, Hurst and McVean, 1997). Also conforming to this pattern are the imprinted gene influences on parenting from outside the maternal brain. Work with PEGs

have suggested they enhance USVs (McNamara et al., 2018a, Bosque Ortiz et al., 2022), placenta signalling and growth (Tunster et al., 2018, Ono et al., 2006, Esquiliano et al., 2009) and improve offspring suckling (Schaller et al., 2010, Curley et al., 2004, Plagge et al., 2004, Cowley et al., 2014), while MEGs have been shown to have the opposite effect for USVs (Jiang et al., 2010, Smith et al., 2011) and on placental growth and signalling (Tunster et al., 2016a, Tunster et al., 2016b, Tunster et al., 2011). This has been suggested to be the consequence of PEGs promoting the acquisition of resources by the offspring from the mother by causing the pups to suckle more, call more and prime the mother to be hyper attentive via the placenta, while MEGs attempt to limit this process (Creeth et al., 2019).

In the maternal brain, predominantly PEGs have been shown to create maternal care deficits when manipulated (Li et al., 1999, Stohn et al., 2018, Lefebvre et al., 1998), which is in line with the idea that PEGs enhance provisioning to the offspring. However, since a paternal effect on the maternal brain is equivalent of the father's genome altering his daughter's brain to make her commit intensively to one child at the expense of the others, this goes against his evolutionary interests in principle and as such conflict theory fails to work as a straightforward explanation in this instance. This is the type of postnatal behaviour in which the other theories of imprinting come to the fore (Haig, 2014). Coadaptation theory was also discussed in Section 1.1.2 as a counter theory which looks at the imprinting process as a form of cooperation rather than competition. In light of the maternal care phenotypes, it could be understood that a genetic symbiosis is formed between the maternal care of the mother and the maternal demand behaviour of the offspring, coordinated by the modulated expression of imprinted genes acting in both organisms in different tissues and timepoints. High quality maternal behaviour is not nearly as effective without heavily coordinated offspring behaviour to demand and take advantage of the behaviour offered. The selection pressure was joint for mother and offspring, requiring a coordination of allelic expression between them.

It is important to note that imprinted genes can still be understood to be affecting the maternal brain through conflict theory. One reality could be that the pleiotropic nature of these genes means that the impact on maternal care can be withstood for the advantages for growth that imprinting these genes provides in which case explaining maternal care phenotypes are a red herring. However other realities of murine lifestyle are worth considering. Murine females tend to communally nest, preferentially with female relatives, which requires an element of sharing maternal labour. Fathers would want their offspring to provide optimal maternal care to support his grandchildren while mothers would want a balanced approach to not provide

impact to their cohabiting relatives (Haig, 1999). Another possibility relates to inbreeding, specifically if it can be shown that offspring tend to mate in their youth with paternal side relatives then the interests of the parents become misaligned again, with fathers favouring their female offspring to expend themselves on their early offspring which are more likely to be theirs while mothers favour a balanced approach across the lifespan (Wilkins and Haig, 2003).

A relationship between imprinted genes and parental care is by no means assured and whether the maternal brain is a site of unique conflict or a site of coadaptation or a site of gene dosage and environmental sensitivity is yet to be explained. Gaining clarity on these points will only serve to help us understand further the existence and purpose of genomic imprinting. Notably, these theories centre around females and maternal care only. The impact of finding imprinted gene deficits in paternal care has yet to be considered and could have implications. Similarly, an idea overlooked in the current literature is the conjoined evolutionary trajectory of eutherian parental care and imprinting which suggests that imprinting's influence on parenting may be more fundamental than first assumed.

1.4 Aims

Taken as a whole, the literature surrounding the role that imprinted genes play on parenting behaviour stands at a precarious point. The two standout genes of the late 1990's (Li et al., 1999, Lefebvre et al., 1998) linking imprinted gene models to maternal care have both been brought into question with modern, neater mouse models (Denizot et al., 2016, Anunciado-Koza et al., 2022). Although there are disputes concerning the way the studies were carried out and the differences in the models, it is convincing that models utilizing LacZ are inappropriate to characterize neural/behavioural phenotypes (Peck et al., 2021) and so *Peg3* and *Mest* have been rightly brought into question. Within the past 5 years, more loss-of-function imprinted gene models have demonstrated a parenting deficit (Keshavarz and Tautz, 2021, Stohn et al., 2018) without the limitations of the earlier LacZ models. Interestingly, the first MEG, *Calcr* has been shown to play a crucial role in parenting (Yoshihara et al., 2021), with levels of expression strong enough to warrant consideration as a marker gene of the parenting neurons in the MPOA (Moffitt et al., 2018). This means that a role for imprinted genes in the maternal brain cannot be ruled out, and on the contrary, makes an approach such as what will be performed in this thesis timelier and more important.

Based on the above, in this thesis I aimed to approach the question "What role do imprinted genes play in parenting behaviours?" in a new manner. Rather than first selecting a gene candidate and assessing mouse behaviour as had been done previously, I instead started my investigation by focusing on what is known about the neural mechanisms controlling parenting behaviour in the brain as described in Section 1.2.4. Specifically, I explored whether imprinted genes demonstrated meaningful expression in the parenting circuitry of the brain and used this expression data to identify new imprinted gene candidates for parenting behaviour. This approach broadened out into a wider question – "where do imprinted genes show enrichment in the brain?" which I followed down to cell subpopulation level specificity. Once gene candidates for a role in parenting had been identified, I selected a gene for an investigation of its role in parenting as an exemplar of the approach, and, since the parenting circuit is conserved in males and females, I sought to expand beyond maternal care and look at paternal and alloparental care also. The specific aims for my thesis are detailed below.

AIM 1: Do imprinted genes demonstrate enriched expression in the 'parenting' neurons of the murine brain? (Chapter 3)

My first aim was to bring a systems level approach to identifying the role imprinted genes may be playing in parenting behaviour, particularly the galanin neurons of the POA known to be the hub region for this behaviour. This was done by data-mining existing single-cell RNA sequencing datasets from the adult mouse brain and quantifying imprinted gene expression patterns. Between 2016-2020 many different datasets were produced including datasets that took single cells from multiple organs across the mouse body at the same time, all the way down to in depth sequencing of specific regions of the brain. I opted to take a hierarchical dataset analysis strategy and carried out an over-representation and gene-set enrichment analysis for imprinted genes in the following ways. First, I asked whether imprinted genes would show enriched expression in the brain compared to other organs of the adult mouse, this was assessed in two independent datasets. Second, I asked whether imprinted genes would display enriched expression in specific regions or cell types of the brain using two datasets that sequenced the whole mouse brain. Thirdly, I pursued this regional enrichment and assessed imprinted gene enrichment in datasets of specific regions of the brain, this included the areas of the hypothalamus key to parenting such as the POA (and the galanin neurons) as well as other sites in the brain important for parenting such as the dopamine and serotonin system as well as an analysis of the pituitary gland. The goal was to show that imprinted genes display a 'gene set' like enrichment in these regions and to create a list of

imprinted genes whose expression pattern would predict an involvement in parenting, namely by displaying high expression in the galanin neurons of the POA. This approach aimed, in principle, to show that imprinted genes are displaying gene-set like behaviours in the neural circuitry for parenting. The rest of my thesis served to validate this methodology by selecting one of these imprinted gene parenting candidates that had not previously been associated with parenting behaviours and carry out a parenting assessment to confirm that this bioinformatic approach could be informative for behavioural genetics.

AIM 2: Does my imprinted gene candidate express in the POA 'parenting' neurons when assessed in-situ using RNAscope®? (Chapter 4)

After selecting a gene candidate, my next aim was to confirm that the significant up-regulation of this gene in parenting-associated neurons in the brain found in sc-RNAseq data was valid by carrying out my own in situ RNA quantification using RNAscope® technology. Namely, if I found enrichment of imprinted genes in the galanin neurons of the POA, I wanted to demonstrate clear expression of my gene candidate in those neurons. *Gal/Th* and *Gal/Calcr* expressing neurons have been shown to be highly active during parenting and would act as the primary targets. I planned to conduct three-plex RNAscope® in WT mouse brain sections through the POA, with both *Gal/Th/gene candidate* and *Gal/Calcr/gene candidate* probe combinations. This approach allowed me to visually demonstrate that the gene candidate co-expresses in cells with the parenting markers and allowed me to quantify the number of gene candidate RNA molecules in a particular cell. In turn this allowed me to assess whether there was an in-situ enrichment to match that seen in the single-cell data. I also quantified the expression of my gene candidate in oxytocin (*Oxt* probe) and vasopressin (*Avp* probe) neurons of the PVN to act as an additional target implicated in parenting behaviour, shown to differ in the *Peg3* model already.

AIM 3: Does a mouse model for my imprinted gene candidate display differences in a parenting assay carried out in mothers, fathers and virgin females? (Chapter 5)

The third aim was to carry out a parenting assessment on a mouse model of my gene candidate identified in Aim 1 and validated in Aim 2. The parenting assessment consisted of the Retrieval/Nest Building assessment as well as a Three Chambers pup-preference test. Mutant mothers, fathers and virgin females were contrasted in their performance against WT mothers, fathers and virgin females. WT animals were paired with mutant animals to produce

litters and so a WT control group was also included, paired with other WT's to produce litters. The ultimate aim was to show that a behavioural deficit could be shown based on a prediction drawn from single-cell RNA sequencing analysis as well as potentially find the first parenting deficits in fathers and alloparents in an imprinted gene model.

AIM 4: Does immediate early gene expression differ in the maternal brain of my imprinted gene candidate mouse model following exposure to pups? (Chapter 6)

The fourth and final aim was to return to RNAscope® and tie any behavioural deficits back to the neural circuitry examined in Aim 1 and 2 by examining neuronal activation (*c-Fos* levels) of parenting neurons upon exposure to pups in my mutant model vs. WT. This was performed in mothers as the most reliable parenting group. Probes were used for *Gal/Calcr/Fos* allowing a quantification of *c-Fos* RNA molecules in *Gal/Calcr* neurons (the most strongly activated neuron population during parenting (Moffitt et al., 2018)) following exposure to pups in my groups. Additional comparisons of number of *Gal* and *Gal/Calcr* positive neurons were also calculated for mutants and WTs allowing preliminary analysis of structural differences

2 General Methods

2.1 Bioinformatic Approach

2.1.1 Datasets

All sequencing data were acquired through publicly available resources and Table 2.1 details the basic parameters of all datasets used in this thesis alongside the repositories that the original data can be accessed from. Once processed, each dataset was run through the same basic workflow (See Section 2.1.3 and Fig. 2.1), with the minor adjustments laid out for each dataset detailed in Appendix A1.

Due to the high variability in sequencing technology, mouse strain, sex and age, and processing pipeline, I did not perform analyses on merged datasets although this has been attempted very recently (Steuernagel et al., 2022). Rather I chose to perform my analyses independently for each dataset and look for convergent patterns of imprinted gene enrichment between datasets of similar tissues/brain regions. As with any single-cell experiment, the identification of upregulation or over-representation of genes in a cell-type depended heavily on which other cells were included in the analysis to make up the 'background.' Analysing separate datasets (with overlapping cell-types alongside distinct ones) and looking for convergent patterns of enrichment was one way of counteracting this limitation.

The raw data analysed in this thesis (author information available in Table 2.1) are available from the following Gene Expression Omnibus (GEO) repositories, Mouse Cell Atlas (MCA) – GSE108097, *Tabula Muris (TM)* – GSE109774, Aging Mouse Brain – GSE129788, Hypothalamus (Chen) – GSE87544, Hypothalamus (Romanov) – GSE74672, Arcuate Nucleus – GSE93374, Suprachiasmatic Nucleus – GSE132608, Dopamine Neurons – GSE108020, Ventral Mid Brain – GSE76381, Dorsal Raphe Nucleus – GSE134163, Pituitary Gland (Ho) - GSE146619, Pituitary Gland (Cheung) - GSE120410, Pancreas - GSE84133, Muscle Tissue - GSE143437, Mammary Gland - GSE106273, and the following Sequence Read Archive (SRA) repository, Mouse Brain Atlas (MBA) – SRP135960.

Table 2.1. Dataset specific sequencing and processing information for Chapter 3 datasets. Datasets are organised by level of analysis and compared for single-cell sequencing protocol, Animal and Tissue processing (whether Males (M) and Females (F) were included), Cell Quality Filters used, No. of Cells in final dataset and Data Normalisation procedure followed (log transformations of Unique Molecular Identifiers (UMIs), Transcripts per million (TPM), RNA counts per cell (RPC).

Dataset	Level	Protocol	Animal/Tissue	Cell Filter	Cell No.	Normalisation
Mouse Cell Atlas (Han et al., 2018)	Multi-Organ	Microwell-seq	C57BL/6J , M+F, 6-10 weeks, Embryonic day (E)14.5 and neonatal	1500 highest quality cells per tissue	61,637	100,000 transcripts log transformed
Tabula Muris (Schaum et al., 2018)	Multi-Organ	Smart-seq2	C57BL/6J 7 mice (4F/3M), 10-15 weeks, virgin	reads>50,000 genes>500	44,879	ln(CPM+1)
Mouse Brain Atlas (Zeisel et al., 2018)	Whole Brain	10X chromium	CD1, M+F, Post-natal day (P)12 – 30, week 6 and week 8,	600<UMI 1.2 UMI:gene	160,796	5,000 reads per cell log transformed
Whole Brain (Ximerakis et al., 2019)	Whole Brain	10X chromium	C57BL/6J, 8 mice 2-3 months, whole brain minus hindbrain	200<UMI<30,000 250<gene<6,000	16,028	10,000 reads per cell log transformed
Whole Hypothalamus (Romanov et al., 2017)	Whole Hypothalamus	STRT-seq (Fluidigm C1)	C57BL/6J, M+F, 14-28 days	Molecules > 1500 (excluding rRNA and mitochondrial RNA)	2,882	10,000 reads per cell log transformed
Whole Hypothalamus (Chen et al., 2017)	Whole Hypothalamus	Drop-seq	B6D2F1 mice (C57B6 female × DBA2 male) - 7 Female, 8-10 weeks	Genes>2000	3319	log(TPM+1)
Arcuate Nucleus (ARC) (Campbell et al., 2017)	Specific Hypothalamic Nucleus	Drop-seq	C57BL/6J – 53mice - 4-12 weeks, virgin, M+F	genes>800	20,921	10,000 reads per cell, log transformed
Suprachiasmatic Nucleus (SCN) (Wen et al., 2020)	Specific Hypothalamic Nucleus	10x chromium	C57BL/6J at timepoint ZT8	None stated (but prefiltered cells provided)	1,251	5000 transcripts per cell, log transformed
Dorsal Raphe Nucleus (DRN) (Huang et al., 2019)	Monoaminergic Nuclei	inDrop	C57BL/6J -8 mice (4M, 4F), 8-10 weeks	18,000>UMI>500 6000>gene>200 mito<0.1	39,411	10,000 UMI per cell log transformed
E11.5 - E18.5 Ventral Midbrain (La Manno et al., 2016)	Monoaminergic Nuclei	STRT-seq (Fluidigm C1)	CD1, E11.5 - E18.5 - 271 embryos	2000>Molecules >26,000	1,907	10,000 reads per cell log transformed
Whole Brain Dopamine (Hook et al., 2018)	Monoaminergic Nuclei	Smart-seq2	C57BL/6J at E15.5 and P7	2000<genes <10,000 1000>RNA >40,000	396	log2(RPC+1)
Pituitary Gland (Cheung et al., 2018)	Pituitary Gland	10X chromium	C57BL/6 6 mice, M, 7-week-old	Genes >= 200	13,620	10,000 UMI per cell log transformed
Anterior Pituitary Gland – 10x (Ho et al., 2020)	Pituitary Gland	10X chromium	CD1, 2M+2F, 7-8 weeks old	Mito < 0.1 200 < Genes < 3000	2,780	10,000 UMI per cell log transformed
Anterior Pituitary Gland - Dropseq (Ho et al., 2020)	Pituitary Gland	Drop-seq	CD1 1M+1F, 8-week-old	Mito < 0.1 UMI > 300 100 < Genes < 4,000	4,663	10,000 UMI per cell log transformed
Pancreas (Baron et al., 2016)	Mouse Organ	inDrop	C57BL/6 and ICR mice, M+F, 5 mice	Reads > 1,000 Genes > 750	1,886	Log(TPM+1)
Skeletal Muscle (De Micheli et al., 2020)	Mouse Organ	10X Chromium	C57BL/6J, M+F, 3-7 months of age, tibialis anterior muscles	UMI > 1000 Genes > 200	34,438	By total UMI per cell, log transform
Mammary Gland (Bach et al., 2017)	Mouse Organ	10X Chromium	C57BL/6J, F, ~8 week at different lactation stages (E14.5, P6, P11 and nulliparous)	UMI > 1000 genes > 500 mito < 0.05	23,184	Size Factors Scran, log2 transform

2.1.2 Imprinted gene list

The gene list for the analyses was based on the list of murine imprinted genes recently published in Tucci et al. (2019). Although the original list of imprinted genes was 260 genes long, only 163 genes had transcripts detected in the most comprehensive of the scRNA-seq datasets, meaning ~100 genes were undetected in any single cell libraries. I further refined this list to 119 imprinted genes (Appendix Table A3.1) which excluded the X-linked genes and consisted of mostly the canonical protein-coding and long noncoding RNA imprinted genes, but the criteria for inclusion was those genes with at least two independent demonstrations of their Parent of Origin Effect (POE) status in the literature. The only exceptions to multiple independent demonstrations of a POE were four genes (*Bmf, B3gnt2, Ptk2, Gm16299*) identified by (Perez et al., 2015) where a POE was assessed across 16 brain regions and 7 adult tissues within one study. For brain-specific analyses, the MEG/PEG status of a gene was matched to the observed status in the specific tissue. Small non-coding RNAs (snoRNAs) such as micro-RNAs (miRs) and small nucleolar RNAs, which represent ~10% of identified imprinted genes, were excluded from the analysis as their sequences were not detected/subsumed by larger transcripts in all but one of the datasets. Another caveat with short-read RNA-seq libraries is that much of the expression data for a given transcription unit cannot discriminate differentially imprinted isoforms nor do some of the technologies (e.g., Smart-Seq2) possess stranded libraries to distinguish antisense transcripts. For complex imprinting loci, such as the *Gnas* locus, most reads as result map to only a few overlapping transcripts e.g. *Gnas* and *Nespas* ignoring several overlapping and antisense genes.

2.1.3 Basic pipeline

Figure 2.1 details the basic workflow for the bioinformatic approach. Data were downloaded in the available form provided by the original authors (either raw or processed) and, where necessary, were processed (filtered, batch-corrected and normalized) to match the author's original procedure. Cell quality filters were specific to each dataset and summarised in Table 2.1. A consistent gene filter, to remove all genes expressed in fewer than 20 cells, was applied to remove genes unlikely to play a functional role due to being sparsely expressed. Datasets of the whole brain/hypothalamus were analysed both at the global cell level (neuronal and non-neuronal cells) and neuron specific level (only neurons) with genes filtered for the ≥ 20 cell expression at each level before subsequent analysis. Cell identities were supplied using the outcome of cell clustering carried out by the original authors, so that each

cell included in the analysis had a cell-type or tissue-type identity. This was acquired as metadata supplied with the dataset or as a separate file primarily from the same depository as the data but occasionally acquired from personal correspondence with the authors. Cells were used from mice of both sexes when provided and all mice were aged 15 weeks or younger across all datasets. Although the focus of the analyses was adult mouse tissue, embryonic data were included in some comparisons when no alternatives were present. However, embryonic and post-natal cells were never pooled to contribute to the same cell populations.

Positive differential expression between identity groups were carried out using one-sided Wilcoxon rank-sum tests (assuming the average expression of cells within the current identity group is 'greater' than the average of cells from all other groups). The test was performed independently for each gene and for each identity group vs. all other groups. The large number of *p* values were corrected for multiple comparisons using a horizontal Benjamini-Hochberg correction, creating *q* values. Fold-change (FC) values, percentage expression within the identity group and percentage expressed within the rest were also calculated. I considered genes to be significantly positively differentially expressed (significantly upregulated) in a group compared to background expression if it had a $q \leq 0.05$. and a Log2 Fold Change (Log2FC) > 0.

The same custom list of imprinted genes with reliable parent-of-origin effects (Appendix Table A3.1) was

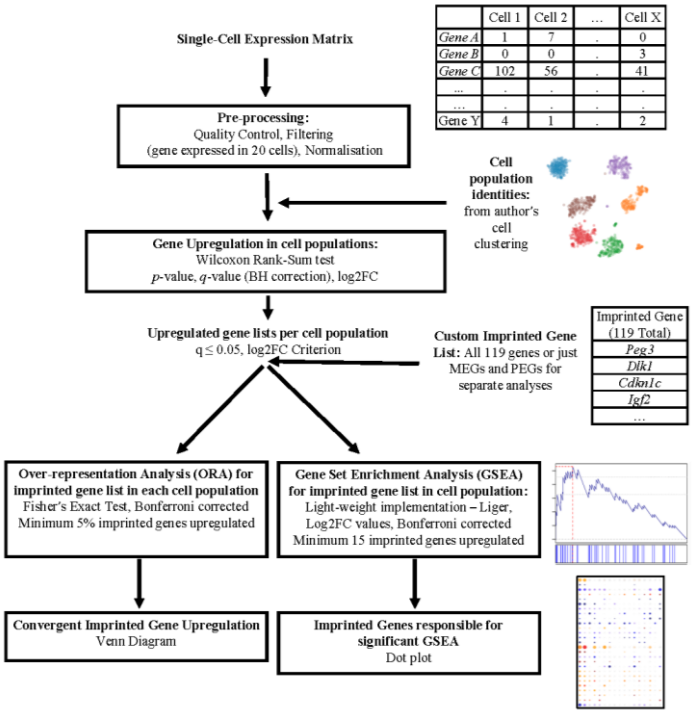


Figure 2.1. Basic Bioinformatic workflow schematic. Single Cell Expression Matrices were acquired through publicly available depositories. Data were processed according to the author's original specifications and all genes were required to be expressed in 20 or more cells. Cell population identities were acquired from the author's original clustering. Positive differential gene expression was calculated via Wilcoxon Rank-Sum Test. Upregulated genes were considered as those with $q \leq 0.05$ and a Log2FC ≥ 1 for analysis levels 1 and 2, while this criterion was relaxed to Log2FC > 0 for level 3. The imprinted gene list was used to filter upregulated genes and two different enrichment analyses were carried out, over-representation analysis via Fisher's Exact Test and Gene Set Enrichment Analysis via Liger algorithm (Subramanian et al. (2005), https://github.com/JEFworks/liger). Venn diagrams and dot plots were utilised for visualisation.

used for all analyses, and all genes were included as long as the gene passed the 20-cell filter. The first statistical analysis for enrichment was an Over-Representation Analysis (ORA) using a one-sided Fisher's Exact Test ('fisher.test' function in R core package 'stats v3.6.2'). The aim was to assess whether the number of imprinted genes considered to be upregulated as a proportion of the total number of imprinted genes in the dataset (passing the 20-cell filter) was statistically higher than would be expected by chance when compared to the total number of upregulated genes as a proportion of the overall number of genes in the dataset (passing the 20-cell filter). To limit finding over-represented identity groups with only a few upregulated imprinted genes, an identity group was required to have ≥ 5% of the total number of imprinted genes upregulated for ORA to be conducted. Subsequent p-values for all eligible identity groups were corrected using a Bonferroni correction. This provided a measure of whether imprinted genes are expressed above expectation (as opposed to the expression pattern of any random gene selection) in particular identity groups.

To further examine the presence of imprinted genes within tissues/cell types, and to provide a different perspective to over-representation, I conducted a Gene-Set Enrichment Analysis (GSEA) for imprinted genes amongst the upregulated genes of an identity group using a publicly available, light-weight implementation of the GSEA algorithm (Subramanian et al., 2005) in R (https://github.com/JEFworks/liger). This was done in a manner similar to Moffitt et al. (2018) since I was using this computational method to identify enrichment of the gene sets *within the upregulated genes* of the different identity groups. Here, the GSEA was conducted for each individual identity group using Log2FC values to rank the upregulated genes. The GSEA acts as a more conservative measure than the ORA since it tests whether imprinted genes are enriched in the stronger markers of a group (the genes with the highest fold change for a group vs. the rest) and hence whether the imprinted genes are enriched in those genes with a high specificity to that tissue/cell type. To prevent significant results being generated from just 2 or 3 genes, identity group to be analysed were selected as having a minimum of 15 upregulated imprinted genes (i.e. the custom gene set) to measure enrichment for (a value suggested by the GSEA user guide (https://www.gseamsigdb.org/gsea/doc/GSEAUserGuide Frame.html)) and to prevent significant results in which imprinted genes cluster at the tail, identity groups were selected as having an average fold change of the upregulated imprinted genes greater than the average fold change of the rest of the upregulated genes for that group. Again, multiple *p* values generated from GSEA were corrected using a Bonferroni

correction. If no cell populations met these criteria, GSEA was not run and not included for that analysis.

2.1.4 Statistics and Data Presentation

Venn diagrams of the upregulated imprinted genes making up over-represented identity groups across datasets (within a level) were reported. Full lists of upregulated imprinted genes can be found in the 'Upregulated_IGs.csv' file for each analysis in the open science framework repository created for the outputs of these analyses (https://osf.io/jx7kr/).

To further elucidate the genes responsible for significant GSEA's, dot plots of the imprinted genes upregulated in that identity group were plotted across all identity groups with absolute expression and Log2FC mapped to size and colour of the dots, respectively. Graphical representations of significant GSEA's (post-correction) are included in the main text or as Appendix Figures, all other graphs, including additional dot plots not discussed in this thesis, can be found in the repository (https://osf.io/jx7kr/).

Dot plots and statistical analyses were conducted using R 3.6.2 (Team, 2013) in RStudio (Team, 2015). Custom R scripts unique for each dataset analysed can be found at the following github repository (https://github.com/MJHiggs/IG-Single-Cell-Enrichment). All Figures were created with the aid of BioRender.com

2.2 Behavioural Approach

2.2.1 Mice

Animal studies and breeding were approved by the Universities of Cardiff Ethical Committee and performed under a United Kingdom Home Office project license (30/3375 then PP1850831, Anthony R. Isles). All mice were housed under standard conditions throughout the study on a 12 h light–dark cycle with lights coming on at 08.00 h with a temperature range of 21°C ± 2 with free access to tap water, and standard chow. Mice were either Wildtype (WT) on a C57BL/6J background acquired from Charles River Laboratories, or *Magel2*-null mice – which were either *Magel2*-FLOX-EM1.1:Het and *Magel2*-FLOX-EM1.1:WT derived from WT female x heterozygous *Magel2*$^{(m+/p-)}$ male (Paternal Deletion) crosses on a C57BL/6J background obtained from the Mary Lyon Centre at MRC Harwell and the following award is acknowledged: MC_UP_2201/1.

2.2.2 Cohorts and Breeding System

Parenting behaviour was assessed in mother, father and virgin female mice. These groups were acquired by pairing mice in-house to produce litters for experimental testing. The breeding and experimental set-up for each group is described below.

Mothers and Fathers

To generate mothers and father cohorts, mice were paired together aged 9 – 12 weeks, females were weighed periodically to confirm pregnancies and litters were born when mice were aged 12-17 weeks. For WT animals, male and female non-siblings were paired together to generate litters. For *Magel2*-null animals, Heterozygous (HET) males and females were paired with non-sibling WT animals of the opposite sex to produce litters. The day the litter was born was considered Postnatal day 0 (P0). On day P1 and P2, the home cage (with the mother, father and pups) was carried to the test room and placed in the testing apparatus with the camera suspended overhead for a habituation period. Both habituation periods lasted 20 minutes, on P1 the cage lid was left on and on P2 it was removed. Mothers and fathers were only included in the assessment if they raised a minimum of 3 pups in their litter to P2 (the minimum for the test). Mothers and fathers were assessed with the same three pups from their litter, the largest 3 based on P2 weight. On the day of the test, P3 (father) / P4 (mother), all test animals underwent a pre-test 20-minute habituation period with the cage lid and enrichment removed from the cage. Following this, the test was set up and began. The non-test animal and the remaining pups were removed from the test room before testing began.

Virgin Females

Virgin female mice were housed in same genotype pairs. When a litter was born (from separate WT x WT matings), and 3 unique pups from these litters designed for testing with the virgin females, the pair of virgin females were habituated on day P1 and P2 (identically to the mothers and fathers). The virgin females were then tested twice, once on P3 and once on P4 (each time proceeded by the 20-minute habituation). The pair of virgin females in a cage were tested one after another, the order of testing was reversed on the second day. Females who had been exposed to pups were kept isolated and not reintroduced to females that hadn't been exposed until their test was also carried out.

2.2.3 Behavioural Testing

All behavioural tests were carried out during the light phase of the light cycle (between 08:00 – 20:00) in a dimly lit room (< 30 lux). All mice were handled using the tunnel technique to avoid undue stress before and after the tests. Following the birth of a viable litter and the habituation period, fathers performed the Retrieval/Nest Building assessment when test pups were P3, mothers on P4 and both carried out the Three Chambers test on P5. Virgin females performed the Retrieval/Nest Building assessment when test pups were P3 and a second assessment when P4 and also carried out the Three Chambers assessment on P5.

2.2.4 Retrieval & Nest Building

Set up

Animals were tested within their home cage which was cleaned to the same standard 2 days prior to testing (the day after the litter was born) by replacing soiled bedding and sawdust. For the test, the home cage had all enrichment removed and was placed, with metal cage lid removed, inside a plastic storage container. A High Definition (HD) webcam was attached to a table mounted tripod and set up, so it was positioned directly above the centre of the home cage (See Figure 2.2A).

Directly prior to the test, the test animal (mother/father/virgin female) was removed from the home cage and placed in a holding cage. The other animal in the pair was then placed in a new clean cage (identical to the home cage) with the enrichment from the home cage, a fresh nest disc and all the pups bar the three test pups. The three test pups remained in the home cage and were positioned against one of the short ends of the cage (the end opposite the home nest – see Figure 2.2B) with two in the corners and one directly in between these two. The home nest was shredded completely and placed all the way along the opposite side of the cage from where the pups were placed (the side in which the home nest was previously located). The recording was started, and test animal was returned to the home cage and placed directly onto the shredded nest. Animals had one-hour total time in which to complete the behaviour test from their first olfactory investigation of the pups. The goal of the test was to retrieve the three scattered pups to the nest material and to re-construct the home nest using the scattered material.

Video Analysis

All videos were recorded on a standard HD webcam and videos were analysed using Behavioral Observation Research Interactive Software (BORIS) which allowed for event logging video coding (Friard & Gamba, 2016). The events/metrics recorded are discussed below. All animals were recorded, and behaviour coded for the 60 minutes or until they completed the test. Time began when the test animal investigated the pups for the first time.

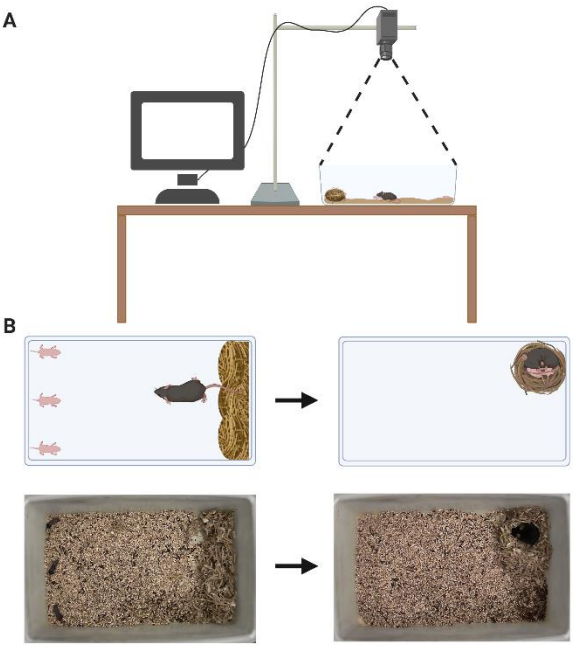


Figure 2.2. Experimental setup for the Retrieval /Nest Building Task. (A) The task was performed in the home cage which was positioned under a standard HD webcam suspended directly above the cage using a clamp. Data were stored and the camera controlled from a PC attached. (B) Retrieval/Nest Building Exemplar Test. *Top* –Image recreation of test start and finish. *Bottom* – Acquired images of exemplar mouse, at the start and at the end of the test.

Retrieval Behaviour Scoring

The first metric was the time taken for the animals to successfully retrieve the pups to the nest. A pup was considered retrieved to the nest when it was carried by the test animal from the pup end of the cage to the nest end of the cage and was placed inside the alcove of the nest or was completely covered from above with nest material. All animals were scored based on the time it took them to retrieve the first pup to the nest under these conditions and the time it took to retrieve all three pups to the nest. How many pups successfully retrieved by the end of the trail was also recorded (0, 1, 2, or 3).

Nest Building Scoring

All animals were also scored on the quality of their parental nest at the end of the 60-minute trial.. All nest scoring was carried out using a 1-5 rating scale used by Neely et al. (2019) (adapted from Deacon, 2006). A score of 1 meant the shredded paper remained scattered throughout the cage or remained untouched; a score of 2 meant some of the material was constructed into a nest, but over 50% of the material was not used for nest construction (i.e., remained scattered or the majority of the original material remained untouched); a score of 3 meant a noticeable nest was constructed, but several pieces were still scattered; a score of 4 meant almost all the material was used for the nest, but a few pieces of material remained

scattered or were near the nest; a score of 5 meant all material was used to make an identifiable nest (See Figure 2.3).

Pup-Directed Behaviour Scoring

During the trial, the proportion of the time spent engaged in pup-directed behaviour was also measured. This was measured up until the mouse achieved full retrieval (or the trial ended) for one metric and the proportion of behaviour for the entire duration of the trial. Pup-directed behaviour was considered any of the following: sniffing or licking or grooming pups, retrieval attempts/carrying pups, nest building, or time spent in the nest whilst pups had been retrieved into it. Non-pup directed behaviours included nest building before pups were retrieved, exploration behaviour, self-grooming or any other behaviour not directed at the pups.

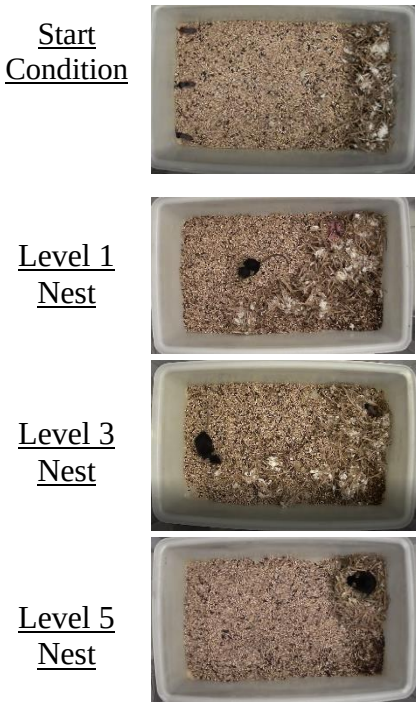


Figure 2.3. **Exemplar Nest building scoring from the Retrieval/Nest Building assessment.** Animals start with nest material fully deconstructed. Level 1 represents no visible nest structure; Level 3 represents a recognisable nest shape yet not fully formed and Level 5 represents a fully formed nest.

Optimisation (Appendix A4)

Since this behavioural experiment had not been carried out in the lab before, I carried out a pilot optimization experiment using WT mice to validate the set up. The aim was to run a small group of mice on the Retrieval/Nest Building assessment as well as on a more traditional retrieval set up without the nest building component. To get a rough measure of sensitivity, I tested several different groups of mice, namely, first-time and second-time mothers and fathers as well as a small group of virgin females. The assessment was only performed for 30 minutes compared to the 60 minutes in the main assessment due to the constraints of the older animal (PPL) license. The data from this optimization (Appendix A4) confirmed that all groups of mice were capable of completing the retrieval and nest building components of the task within a 30-minute period, although several mice didn't start retrieving until near the end of this time limit, prompting me to extend the test time to 60 minutes. The data also revealed an increased sensitivity to detect differences between groups of mice capable of parenting

behaviour from the Retrieval/Nest Building assessment as opposed to the standard retrieval. This convinced me to use this assessment as the main assessment of parenting in this thesis.

2.2.5 Three Chambers

On P5, all animals carried out a Three Chambers assessment with the same three pups as used in the retrieval test. The Three Chambers apparatus consisted of a white Perspex arena (40 x 30 x 30 cm, h x w x d) divided into three equal chambers connected in a row. Two guillotine doors (5x5 cm, operated by a pulley system) were used to connect each of the exterior chambers to the middle chamber. A HD webcam was attached to a table mounted tripod and set up, so it was positioned directly above the centre of apparatus (See Figure 2.4A)

For the pup preference assessment, mice were initially habituated to the middle chambers for 5 minutes with the guillotine doors closed. After 5 minutes, 3 pups (and fresh bedding) were placed at the outer edge of one of the exterior chambers with a protective cage placed over the top of the pups and weighted down. At the outer edge of the other exterior chamber, a novel object (a large Lego brick and an equivalent amount of fresh bedding) were placed and covered with an identical protective cage (See Figure 2.4B). The two Guillotine doors were then opened simultaneously. From this point, the mouse had 10 minutes in which to freely explore all chambers. Apparatus was wiped down thoroughly between trials and the chamber with pups in was alternated between trials.

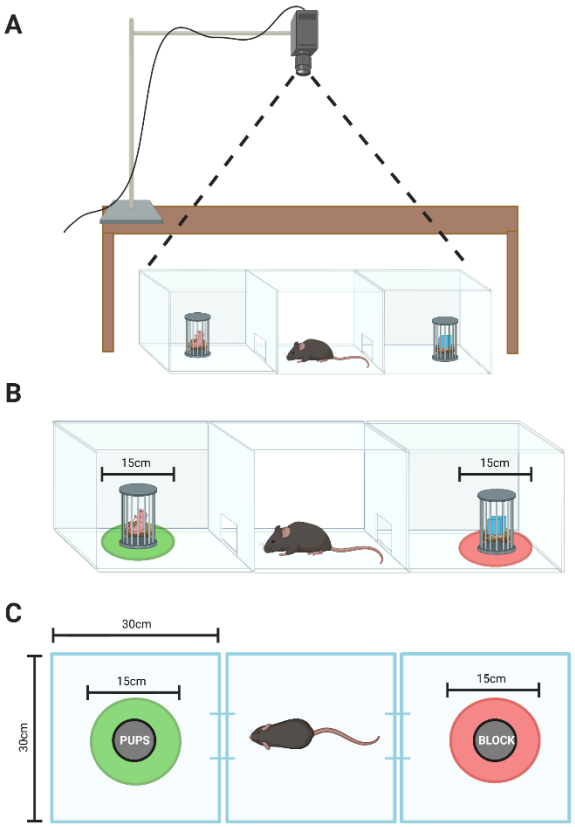


Figure 2.4. Experimental Setup for the Three Chambers Task. (A) The task was performed in a white Perspex area which was positioned under a standard HD webcam suspended directly above the cage using a clamp. Data were stored and the camera controlled from a PC attached. (B) Three Chambers Exemplar set up. Mice had 10 minutes to freely explore the three zones, one containing three pups under a protective cage and the other containing a novel object under an identical protective cage. (C) Three Chambers set up as viewed from above. 15 cm zones were used around the cage with the pups and novel object in, to score amount of time animals spent in these zones.

Video Analysis

The Three Chambers assessment was scored using Ethovision. Each chamber of the three chambers was set as a unique area as well as unique areas set up for the immediate circular zone around the two protective cages (15 cm radius). I assessed the number of seconds that the test mice spent in the pup chamber compared to the object chamber, but primarily, I recorded the time the animal spent within a 15 cm diameter of the pup cage as well as the time spent within a 15 cm diameter of the novel object cage (Figure 2.4C). Animals were automatically tracked using their nose and tail as markers and the mouse's position was tracked based on where the centre of their body was at any particular time. Hidden zones were utilized for any place the mouse couldn't be observed from above. Each mouse had 600 seconds of positional information, each starting from the centre of the middle zone.

2.2.6 Statistics and Analysis

All statistical analysis were carried out, and graphs plotted, using R 3.6.2 (Team, 2013) in RStudio (Team, 2015). All statistical tests were done independently for mothers, fathers and virgin females. All Figures were created with the aid of BioRender.com

For retrieval & nest building metrics, for mothers and fathers, statistical tests consisted of one-way Analysis of Variances (ANOVAs) with post hoc Tukey's test for continuous variables. For non-continuous variables such as nest quality score and number of pup's retrieved Kruskal Wallis tests with post hoc Dunn's tests were used. The latter were also used where variables severely violated parametric assumptions. Assumptions were assessed using Levene's test of equality of variances and Shapiro-Wilk test of normality. For virgin females, two-way ANOVAs (accounting for whether the assessment was the first or second assessment) with post hoc t-tests for continuous variables and a nonparametric analysis of longitudinal data ('nparLD') with post hoc Dunn's test for noncontinuous variable and variables violating parametric assumptions. Box's M test for equality of multiple variance-covariance matrices was additionally used.

For the Three Chambers assessment, each cohort was assessed on the difference between time spent in the pup zone vs. time spent in the novel object zone using two-way one-sample t-tests against zero, simply asking whether any cohort of animals showed significant preference for one zone over the other and hence having a time difference greater or less than zero.

2.2.7 Parenting confounders

Olfaction

The number of seconds to first sniff the pups in the Retrieval/Nest Building assessment was recorded for each mouse and was assessed with one-way ANOVA with post hoc Tukey test.

Motility

Motility analysis was also carried out using Ethovision for Retrieval/Nest Building and Three Chamber assessments. Hidden zones were utilized for the nesting material in the Retrieval/Nest Building and any areas in the Three Chambers in which the mouse could not be reliably tracked. Animals were automatically tracked using nose and tail and mouse position was tracked based on where the centre of the body was at any particular time. Velocity was calculated during the Retrieval/Nest Building and Three Chambers tests as well as the number of chamber crosses that each group performed during the Three Chambers test. Each of the zones in the Three Chambers were marked as separate areas in Ethovision and the number of times a mouse moved from one zone to another were scored. Average velocity per mouse was automatically calculated based on mouse position across time in the two trials. The whole home cage environment was used as an area to score velocity in for Retrieval/Nest Building. The velocity and number of crosses in each group were assessed with one-way ANOVA with post hoc Tukey test.

Litter size

Litter size at P2 was recorded for each pairing and was assessed with Kruskal Wallis test with post hoc Dunn's test. Nests were checked every morning and evening P0 - P2 and litter loss or dead pups were recorded when found.

Effects of Mutant pups

Magel2 deletions have been shown to influence pup behaviours such as suckling (Schaller et al., 2010) and isolation induced USV production (Bosque Ortiz et al., 2022), the latter has also been shown to influence maternal preference for retrieval at P8 (but not when the pups were younger at P6). *Magel2* is a paternally expressed gene, hence the groups at risk of this confounding variable are the *Magel2*$^{(m+/p-)}$ fathers and the WT females they were paired with. By testing retrieval behaviour at P3 and P4, I avoided the P8 window in which USV production altered maternal behaviour but, to address this further, I attempted to minimise the

number of paternally inherited *Magel2*$^{(m+/p-)}$ pups used for testing by selecting the three heaviest pups (at P2) for every subject (which were distinguished on subsequent days by colouring the back of the pups with marker pen).

I also sought genotyping confirmation that the number of *Magel2*-null pups were minimized in my test litters. Hence, following behavioural assessment in *Magel2*-null mice (after Three Chamber test at P5), tail clips were taken from the all the pups used in the behavioural assessments. Genotyping was carried out by MRC Harwell to confirm the number of paternally inherited mutants were present in the testing pup population. The animals with mutant pups in their test litters were then assessed separately to determine whether they would show a retrieval preference for WT pups over mutants.

2.2.8 Mouse Behaviour for tissue harvest

Mother mice scheduled for tissue harvest were either culled from their home cage or underwent a brief behavioural exposure. Mice undertaking the behavioral exposure were habituated to the Retrieval/Nest Building set up as described above until the pups reached P4. Prior to tissue collection, pups and adult mice not to be dissected were removed from the home cage and placed in a temporary holding cage with new nest material. The mouse for tissue harvest was left in the testing room for 1 hour prior to the test to standardize *c-Fos* activity in the POA. 3 pups were then introduced to the home cage on the opposite side of the cage from the nest. Time was started upon first sniff of the pups. After 30 minutes post-first-exposure, the pups were removed, and the animal transported for perfusion and tissue harvest.

2.3 Molecular Approach

2.3.1 Perfusion and Brain Dissection

Animals designated for tissue harvest were administered an appropriate dose of Euthatal™ intraperitoneally until the mouse was unresponsive to a sharp toe pinch. The mouse was positioned on a grid suspended above a collection dish (for excess run off) within a fume cupboard. The first incision was made below the ribcage of the mouse and two lateral incisions made either side of the rib cage so that this section of mouse tissue could be clamped open to expose the heart. The right atrium was cut, and the left ventricle was punctured carefully with a blunted needle attached to the perfusion set up (so as not to puncture through the back of the ventricle). The animal was then perfused by hand, initially

with 1x-RNA-free phosphate buffered saline (PBS) (~20 ml) from the first 100 ml syringe before switching the flow from the tubes and delivering an equal amount of 10% Neutral Buffered Formalin (NBF; Sigma Aldrich, Gillingham/UK) from the second syringe.

Once fixation was completed, the animal was immediately dissected. The mouse was first decapitated with scissors, cutting from behind the ears at the base of the skull. The skin surrounding the skull was removed with a series of cuts between the skin and skull. An incision down the centre of the skull allowed the careful peeling of the two sides of the skull with forceps to expose the whole brain. Excess bone was removed, and the skull inverted. The brain was then carefully cut away from the tissue holding it to the base of the skull, so that it detached upside down on a pre-chilled dissection plate. Whole brains were placed in fresh 10% Neutral Buffered Formalin (NBF) for a further 24 hours before tissue preparation.

2.3.2 FFPE Tissue Preparation

The POA was the target for all molecular work in this thesis and so after 24 hours, the dissected, fixed, brain tissue was sectioned into a 3 mm block using a brain slicing matrix with 1 mm slice channels (Zivic Instruments, Pittsburgh/USA). This 3 mm section was taken with the POA situated in the centre of the block. The rest of the brain was appropriately disposed of, and the 3 mm block underwent standard pre-paraffin treatment. First the tissue was washed with 1x-RNA-free-Phosphate Buffered Saline (PBS) and then underwent a standard ethanol series (30%, 50%, 70%) for 30 minutes each before storage in 70%.

Tissue blocks were delivered to Cardiff University's School of Biosciences Histology facility to undergo the rest of the process, namely a standard xylene series, followed by paraffin before embedding the tissue block in fresh paraffin, orienting the tissue block so that the caudal surface of the 3 mm brain block was at the front. Blocks were then ready for sectioning.

2.3.3 Slide Preparation

Brain blocks were sectioned at 10 µm coronally. Due to shrinkage across the tissue preparation process, the 3 mm brain blocks lose ~33% of their area and so a complete brain block would deliver ~200 10 µm sections. For every 10 sections taken, section 8 was put aside for H&E staining carried out by the School of Biosciences Histology facility. Section 9 and 10 were mounted on one slide. The early sections in the run of 10 were discarded. This allowed me to have a section representative of every 100 µm of the block. Slides were

collected at this point to continue the protocol. A visual POA sectioning guide can be found in Appendix A5

2.3.4 RNAscope® protocol

Three-plex RNA Scope was performed using RNAscope® Multiplex Fluorescent Reagent Kit v2 (ACD Bio-techne, Abingdon/UK) on these Fixed-Formalin Paraffin Embedded (FFPE) brain sections. The manufacturer's protocol was followed exactly following the 'standard' pre-treatment guidance. Briefly, slides were baked for 1 hour in a dry oven. Both sections on the slides underwent deparaffinization in the fume hood (2 times 5-minute Xylene washes and 2 times 2-minute 100% Ethanol (ETOH) washes) and then a H2O2 incubation (10 minutes at room temperature) before being added to boiling RNAscope® Target Retrieval for 15 mins, followed by a 30 mins incubation at 40°C in the HybEZ Oven with Protease Plus. Each of the previous three steps was followed by two washes in dH20 for 1 minute each. The RNAscope probes (all previously existing in the catalogue and ready to order) for the genes used in each assay were combined into a probe mix after being heated at 40°C for 10 minutes. Probe mix (stored at 4°C between experiments) was added, and slides incubated for 2 hours at 40°C in the HybEZ Oven. One section on the slide received the probe mix while the other section on the slide (containing the adjacent cells) received an equal amount of probe dilutant to act as a no-probe control. Slides were stored overnight at this point in 5x saline-sodium citrate (SSC), and the protocol continued on the following day.

Wash buffer (made up by combining 60 ml 50X ACDBio wash buffer and 2.94l of dH2O) was used twice for 2 minutes between every one of the following steps. Amplification steps (AMP1, AMP2, AMP3) were performed (incubated at 40°C for 30, 30 and 15 minutes respectively) before each channel's signal was developed. All RNAscope® experiments in this thesis were run in three-plex and hence the signal of each channel needed developing separately, channel 1 followed by 2 and then 3. Each channel was developed in the same manner, under red light to minimise fluorescent bleaching of the fluorophores once applied. First a horseradish peroxidase (HRP) (HRP-C1/C2/C3) was administered and incubated for 15 minutes at 40°C to open the specific channel. Following 2 times 2-minute wash buffer step, the fluorophore (TSA Vivid™ 520, 570 or 650, Tocris/Bio-Techne Ltd. Abingdon/UK) designed for that channel was then administered and incubated for 30 minutes at 40°C. All fluorophores were applied at a concentration of 1:1500. TSA Vivid™ 520 fluoresced in green (Fluorescein) and was recommended for high expressors due to the high autofluorescence

produced in this channel. TSA Vivid™ 570 (Orange, Cy3) and TSA Vivid™ 650 (Red, Cy5.5) were easily distinguishable from autofluorescence and recommended for lower expressors. Finally, HRP-Blocker was then administered and incubated for 15 minutes at 40°C to close the specific channel. Slides were then counterstained with the DAPI (4′,6-diamidino-2-phenylindole) provided in the RNAscope® kit (30 seconds) and mounted with Invitrogen™ Prolong™ Gold Antifade (Fisher Scientific, Loughborough/UK) mounting medium. Slides were left to dry at room temperature overnight before being stored, covered, at 4°C until the slides were scanned.

2.3.5 Slide Scanning

Whole brain slides were imaged at 20x magnification within one week of mounting using the Carl Zeiss Ltd. (Cambridge/UK) AxioScan Z1 with HXP 120 V lamp attached. Zen Blue slide scanner program was used to run a scanning profile. To set up the scan, first, a Region of Interest (ROI) was drawn around both whole brain sections separately. From here on the two sections would be treated as two unique objects with identical settings applied. Autofocusing used DAPI to find optimal focus and occurred at a course level (10x magnification at 6 points per section) and then at a fine level (20x magnification at 24 points per section in an onion skin format) to create focal points to take the images. Prior to images being taken, light intensity settings and light duration settings were specified for each fluorescent channel (DAPI, Fluorescein, Cy3, Cy5.5). Light intensity was kept constant throughout this thesis for all channels (DAPI/fluorescein = 50% intensity and Cy3/Cy5.5 = 100% intensity). Where possible maximal intensity was used except for fluorescein in which this was reduced to 50% to weaken the autofluorescence effect produced in this channel. Light duration was variable for different probe/fluorophore combinations but was kept constant per experiment. Light intensities were calculated by first utilizing the ‘set exposure’ function on the first batch of slides to find the optimal light duration for each probe/fluorophore to generate a 30% shift in the pixel intensity distributions in the probe tissue in the POA. This differed slightly between slides and so the average was found (and rounded up to the nearest 100) and this value was used as the light duration for every slide containing that probe/fluorophore. Table 2.2 details the specific settings for each gene/probe used in this thesis. Keeping these values constant allowed comparison between slides. With the lamp settings specified for each channel, images were taken for every portion of the ROI, one image per channel (1000’s of images per section) and stitched together using DAPI as a reference to make a complete image.

Table 2.2. Settings used for each gene/probe combination when acquiring images used Zen AxioScan Z1. Gene/probe combinations are listed with the catalogue code the probe was registered under in the ACD Biotechne catalogue and the specific experiment the probe was used in within this thesis. For each probe combination, the intensity percentage and light duration (milliseconds (ms)) are recorded.

Experiment	Gene / Probe	Probe Catalogue code	HXP 120 V Intensity	HXP 120 V Duration
TSA Vivid™ 520 (Green, Fluorescein)				
Calcr/Gal/Magel2	*Calcr*	ACD 494071,Mm-Calcr	50%	1500 ms
Gal/Calcr/Fos & Gal/Th/Magel2	*Gal*	ACD 400961, Mm-*Gal*	50%	400 ms
Oxt/Avp/Magel2	*Oxt*	ACD 493171 Mm-*Oxt*	50%	100 ms
TSA Vivid™ 570 (Orange, Cy3)				
Calcr/Gal/Magel2	*Gal*	ACD 400961-C2, Mm-*Gal*-C2	100%	15 ms
Gal/Th/Magel2	*Th*	ACD 317621-C2, Mm-*Th*-C2	100%	25 ms
Gal/Calcr/Fos	*Calcr*	ACD 494071-C2,Mm-Calcr-C2	100%	10 ms
Oxt/Avp/Magel2	*Avp*	ACD 401391-C2, Mm-*Avp*-C2	100%	2.5 ms
TSA Vivid™ 650 (Red, Cy5.5)				
Calcr/Gal/Magel2, Gal/Th/Magel2, Oxt/Avp/Magel2	*Magel2*	ACD 502971-C3, Mm-*Magel2*-C3	100%	150 ms
Gal/Calcr/Fos	*Fos*	ACD 316921-C3, Mm-*Fos*-C3	100%	200 ms

2.3.6 Slide Analysis Pipelines

Images were analysed with Zen Blue 3.6 Image Analysis software (see Figure 2.5). Images were first pre-processed in order to counteract some of the predictable error generated during fluorescent microscopy. This was done using two methods. First a 50-pixel radius rolling ball background subtraction was applied for all channels (DAPI, Fluorescein, Cy3, Cy5.5). This background subtraction method was designed to overcome a certain amount of background fluorescence and to address situations in which background fluorescence was uneven by removing a background intensity value from every pixel, calculated uniquely using pixels within the radius specified. This effectively removed a large amount of the background fluorescence in my images. The second preprocessing step was a Gauss smoothing applied to every channel (2 pixel (x,y) for DAPI, 1 pixel (x,y) for the other channels). This was utilised to address a particular form of error. Due to the nature of fluorescent microscopy, single fluorophores are not detected as unique points but instead produce ‘airy pattern’ signals due to the dispersal of photons. This means true RNA signal cannot be quantified as single pixel values but instead as ‘blurred’ multi-pixel disks. In an area without signal, if the detector plate is aberrantly excited and releases electrons without photon exposure, this error will appear as a single fluorescent pixel. This error would be identified as signal. By deploying the Gaussian smoothing, every pixel is transformed into a weighted average of the

surrounding pixels and the intensity of single pixels are drastically reduced while the airy disc signal is maintained and smoothed into a more consistent signal pattern to be detected.

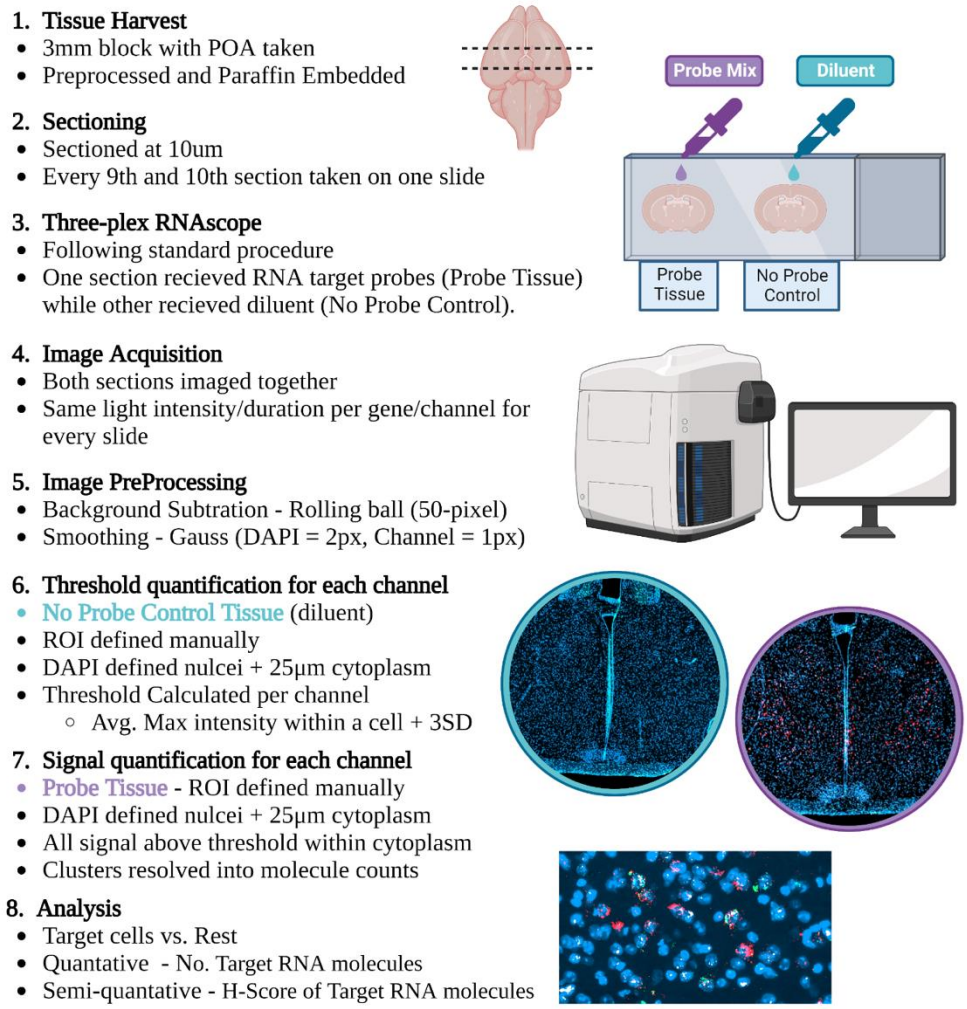


Figure 2.5. Summary of RNAscope® image analysis workflow. A 3 mm block of mouse brain was harvested, fixed and paraffin embedded. Every 9th and 10th section through the block were taken on the same slide. Both tissue sections underwent the full RNAscope® protocol however during the addition of the probe mix, one section was the experimental tissues and received a probe mix (either *Gal/Th /Magel2*, *Gal/Calcr/Magel2*, *Oxt/Avp/Magel2*, *Gal/Calcr/Fos* probe mix) and the other received diluent during this step, to act as the no-probe control. Images were acquired on a Zen AxioScan Z1 at 20x magnification and fluorescent light intensity and duration were kept the same between slides of the same probe mix. Images were pre-processed. The ROI was defined on the no-probe control tissue, nuclei resolved, cytoplasm defined and then the maximum intensity of a pixel for each channel within each cell was recorded. This value indicated the minimum threshold needed for this cell to be classed as positive for that gene. Each gene then had a threshold value derived using the average maximal intensity plus three standard deviations. This threshold value was applied to the probe tissue to identify signal. Clusters were resolved and Target RNA (*Magel2/Fos)* expression was analysed quantitively and semi-quantitively.

Processed images were then ready to have thresholds calculated. All remaining analysis required an ROI to be defined first. For both sections on the slides, the whole hypothalamus, POA, SCN, SON or PVN were manually defined as the ROI. The threshold for the gene signals (Fluorescein, Cy3, Cy5.5) were identified from the no-probe-control section (the adjacent section on the slide with probe diluent instead of probes) which was analysed first. Using Zen Blue's Image Analysis, nuclei within the ROI of the no-probe control section were identified using a 75-intensity threshold for DAPI signal. This identified every pixel with DAPI intensity higher than 75 as 'nucleus' and every other pixel as background which effectively identified nuclear objects. Nuclei had to exhibit circularity of > 0.5 and an area of 50-500 µm to be included. To simulate a cytoplasm around the valid nuclei, a 25 µm border was placed around each nucleus as an estimated cytoplasm. For every cell (nucleus plus cytoplasm) in the no-probe control, the intensity of pixels within each cell were quantified and the maximum intensity value and average intensity values for each of the channels (Fluorescein, Cy3, Cy5.5) were calculated. The maximum intensity value for each cell in the no-probe control is equivalent to the intensity of the strongest pixel within the cell border and hence can be considered the value by which that cell would be considered positive when thresholded. I hence used this value to calculate a threshold value for each channel, by taking the average maximum intensity value for cells in the no-probe control plus three standard deviations (SD). This effectively sets a threshold for each channel (Fluorescein, Cy3, Cy5.5) that would produce a false signal in <0.01% of cells in a tissue without probes imaged under the same light conditions. It is worth noting that standard thresholding procedures would result in a much higher false positive rate meaning my approach was more conservative compared to the standard procedure but more accurate when quantifying single signals.

Next, for the probe-sections, nuclei were again localized within the ROI using DAPI signal intensity (with the same settings as the no-probe control and used across all slides) and a 25 µm cytoplasm. Objects were identified for each channel and fluorescent pixels within the cell border which exceeded the threshold value for that channel were quantified as signal. Neighbouring pixels higher than threshold were grouped into a single object and a watershed of 5 was used for all channels to resolve situations in which multiple distinct signals overlapped and would otherwise be classed as a single object.

2.3.7 Quantifying Signal

RNAscope signal displayed in one of three patterns. Dots were the primary pattern, seen as individual distinct RNA molecules, and were tagged and distinguishable. Clusters were often seen and were created in situations where individual tagged RNA molecules were in too close proximity to be distinguished individually. Finally, some genes were expressed in such high quantities that all 'dots' and 'clusters' were not distinguishable but instead one large fluorescent cluster covering the whole cytoplasm of the cell was observed. This type of expression was termed 'super expressor'.

Most genes assessed in this thesis (*Gal, Calcr, Th, Magel2, Fos*) displayed RNA expression as dots and clusters. Pixels whose intensity exceeded threshold were quantified into objects (consisting of sets of neighboring pixels). Small objects (< 10 µm) were denoted as dots and were considered single molecules of the gene. For objects large than 10 µm (i.e., potential clusters), these could be resolved into molecule counts using the guidance from ACD Biotechne (SOP 45-006). Specifically, the average integral intensity of individual molecules (minus average background) was calculated, and clusters were resolved by dividing the integral intensity of the cluster (minus average background) by this average. The finished data took the form of individual molecule counts for each of the channels for each DAPI identified cell.

For the 'super expressors' which were *Oxt* and *Avp* in this thesis, there were no single molecule signal and no way to resolve the large clusters into individual molecules. The only data extractable for these genes was whether a cell was positive or negative. Given the all-encompassing expression of these genes, novel issues presented themselves. Specifically, neighbouring cells tended to be picked up as false positives from minor overlap at the peripheries of a positive cell. To address this, a 'percentage coverage' criterion was devised for a cell to be classed as a true positive. In practice, a certain percentage of the pixels comprising the cell had to register as positive for the cell to be considered a true positive. If true gene expression for these genes involves cytoplasm filling expression, then cells with only 10% of the pixels as positive are likely positive thanks to an overlap with a neighbouring true positive cell. For *Oxt* and *Avp*, manual visual assessment confirmed a 20% gene coverage was sufficient to dispel false positives. Hence, if a cell were made up of 1000 pixels, 200 had to be positive for a cell to be called a positive cell.

2.3.8 Analysis and Presentation

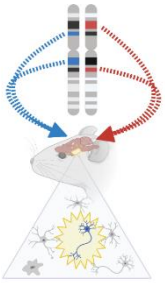
RNAscope® image data were analysed in two ways. Firstly, my quantitative analysis utilised molecule counts for each gene and simple metrics, such as the number of positive cells for a gene probe, could be quantified in the POA. More advanced assessments were also calculated, for example, the number of molecules of target genes (such as *Magel2* and *Fos*) were quantified in cells of interest (e.g., *Gal/Calcr* cells) vs. background expression. Statistical analyses for the experiments in Chapter 4 which focused on *Magel2* RNA counts were carried out using either Wilcoxon Ranked Sums Tests with Bonferroni correction or Kruskal Wallis Tests with post hoc Bonferroni corrected Dunn tests for multiple comparisons. For c-*Fos* RNAscope® image data, I compared the proportions of positive cells for a particular gene (*Gal/Calcr* – 2+ molecules, *Fos* – 5+ molecules) between areas of the brain and cell types. Variability in sections/POA position was accounted for by normalizing *Fos/Gal/Calcr* positive cell counts per animals to counts per 1000 POA cells. WT and *Magel2*-nulls, pup-exposed and non-pup-exposed, males and females, were compared using three-way ANOVA. Additionally, the average number of gene molecules was also compared between *Magel2*-null mice and WTs and analysed with three-way ANOVA also.

Secondly, as suggested by ACD Bio-techne (SOP 45-006), a semi-quantitative metric was also created by deriving a H-Score for the groups compared. The percentage of cells: with 0 gene molecules was multiplied by zero, with 1-3 molecules was multiplied by one, with 4-9 molecules was multiplied by two, with 10-15 molecules was multiplied by three and with 16+ molecules was multiplied by four. The resulting H-Score depicts a measure of cell distribution, and how many cells would be considered high expressors.

Exemplar images of RNAscope sections were acquired using Zen Blue using images without any processing steps applied. Channel displays were set to optimal settings by eye to minimise background fluorescence while maintaining signal integrity. The ‘Image Export’ function in Zen Blue allowed the extraction of high-quality images of small or large ROIs, saved as PNGs.

All graphical representations and statistical analyses were conducted using R 3.6.2 (Team, 2013) in RStudio (Team, 2015). All Figures were created with the aid of BioRender.com

3 Imprinted Gene Single-Cell RNAseq Enrichment



3.1 Overview

Analysis of gene expression, the process by which the information encoded in a gene is transcribed into gene product, has been a possible window into gene function for decades. With the advent of high-throughput RNA sequencing, it has become possible to quantify mRNA for hundreds of thousands of genes simultaneously, breaking free from the restricted single transcript methods (Northern Blots and Real-Time Quantitative Polymerase Chain Reaction (RT-qPCR)). Single-Cell Transcriptomics has now established the next generation of sequencing, allowing mRNA quantification from single isolated cells (the fundamental level for mRNA quantification), rather than bulk, cellularly heterogenous, tissues. Now one can quantify gene expression of hundreds of thousands of genes, across thousands of cells simultaneously, allowing both in depth and cell-specific quantification of gene expression. This chapter sought to utilise this technology to explore the expression of imprinted genes.

The current consensus for the sites of imprinted gene expression has predominately been informed by a convergence in the characterization of a selection of well-studied imprinted genes, typically those first identified in the 1990's and early 2000's (see Section 1.1.3). Primary amongst these convergent expression/functional hotspots are the placenta (Tunster 2013, Peters 2014), the tissues of the embryo and neonate (Plasschaert and Bartolomei, 2014, Hudson et al., 2010), and the developing and adult brain (Perez, Rubenstein & Dulac, 2016). Within the brain specifically, hypothalamic expression has been most notable for imprinted genes with models displaying behavioural implications affecting innately motivated behaviours such as feeding and suckling (Pulix and Plagge, 2020, Ivanova and Kelsey, 2011). In addition, other adult tissues have been implemented such as adipose tissue (Millership, Van De Pette & Withers, 2019), muscle tissue (Martinet et al., 2016) and the mammary gland (Hanin & Ferguson-Smith, 2020). These tissues may at first appear random, but they functionally converge on the resource acquisition and growth of offspring from foetus to adulthood. However, the modern list of IGs now expands far beyond the initial well-studied

examples, and whether the expression profile of this modern imprinted gene set will mirror the early findings, and the theories built from them, has yet to be tested.

Hence, this chapter sought to investigate imprinted gene expression using an enrichment approach, looking at tissues, sub tissues and cell types in which imprinted genes, as a gene set, are over-represented amongst the upregulated genes, or in which they demonstrate a gene set enrichment. This approach sought to validate the associations that exist in the literature, that imprinted genes are predominantly expressed in placenta, embryonic and brain tissues, by carrying out unbiased enrichment analyses for imprinted genes in existing single-cell RNA sequencing (scRNA-seq) datasets. Furthermore, by using scRNA-seq data, I wanted to not only isolate enriched tissues, but also cell types and functional implications based on expression. To achieve this, my approach was hierarchical, by first analysing the large-scale datasets sequencing cells from many tissues in the mouse body and then delving deeper into the enriched tissues to identify enriched cell types and functional implications. This chapter had a particular interest in the organs relevant to proper parental care (e.g., brain, pituitary, ovary) and aimed to use imprinted gene expression data to identify novel imprinted gene candidates which could have a role in parenting based on expression. I was particularly interested in imprinted gene expression in the galanin neurons of the preoptic area, neurons shown to act as a hub when it comes to parenting behaviour (Wu et al., 2014, Kohl et al., 2018, Moffitt et al., 2018).

3.2 Methods

16 unique datasets were analysed across the four levels of analysis, 13 included in the main body of the chapter and making up analysis levels 1-3 (see Figure 3.1) and another 3 datasets were analysed in level 4 (Appendix A2). At each level of analysis, I aimed to be unbiased by using all the datasets that fitted the scope of that level, but the availability of public scRNA-seq datasets was limited, which prevented me from exploring all avenues (for example, a direct comparison of enrichment between hypothalamic nuclei).

Imprinted gene enrichment was assessed using an over-representation analysis (ORA) and gene set enrichment analysis (GSEA) on upregulated gene lists for every population of cells in the detailed analyses (see Section 2.1.3). The same list of 119 imprinted genes was used for all analyses (Appendix Table A3.1), carefully curated to include only genes reliably demonstrating allelic biases in expression (details found in Section 2.1.2).

Level 3 and Level 4 (Appendix A2) analyses were carried out as detailed in Section 2.1.3 with a log2FC > 0 criteria for upregulation. For Level 1 and Level 2 analyses, the criteria for upregulated genes included demonstrating a Log2FC value of 1 or larger (i.e., 2-fold-change or larger). The datasets at these levels represented cells from a variety of organs, regions, and cell-types, and in line with this cellular diversity, the aim of these analyses was to look for distinctive upregulation, akin to marker genes. Once the analysis was restricted to cell subpopulations within a specific region of the brain (Level 3) or within other organs (Level 4), the additional criteria for upregulation was relaxed to demonstration of a positive Log2FC (i.e., the gene has a higher expression in this cell type than background). This was mainly because I was not expecting imprinted genes to be 'markers' of cell subpopulations at this level, but my aim was to identify enriched expression profiles for IGs. This additionally ensured consistent criteria for enrichment within levels, allowing meaningful comparison.

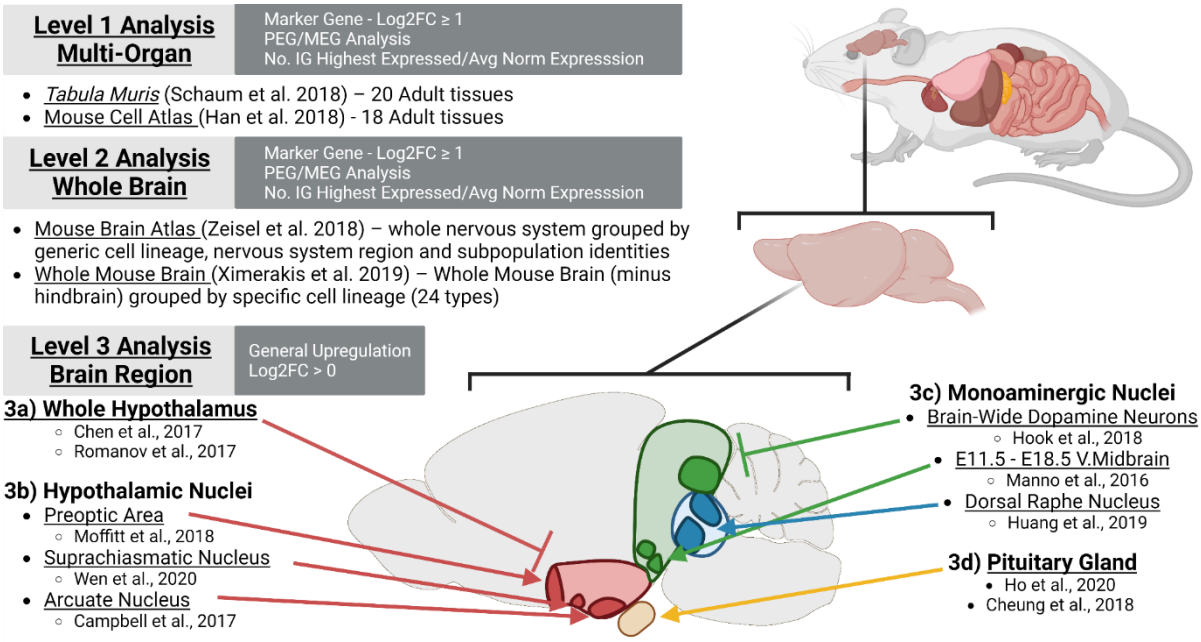


Figure 3.1. The hierarchical set of datasets analysed in Chapter 3. The datasets are sorted into Level 1 (Multi-Organ), Level 2 (Whole Brain) and Level 3 (Brain Nuclei) analyses. The original publication and specific tissue/s analysed are provided for each analysis. White text in dark grey box indicates specifics to the analysis at that level – whether the analysis used the 'marker gene' Log2FC criteria or the relaxed Log2FC > 0 criterion, whether paternally and maternally expressed gene (PEG/MEG) analysis was carried out and whether the number of IGs with highest expression in a cell population and the average normalised expression were reported for IGs.

In addition, within Level 1 and Level 2 analyses, I also carried out parent-of-origin specific analyses. The imprinted gene list was divided into MEGs and PEGs, and the analyses detailed above were run separately for these two gene groups. For imprinted genes with known parent-of-origin variability based on tissue type (*Igf2* and *Grb10*), the parent-of-origin

characterisation of these genes was changed accordingly. The number of imprinted genes with their highest expression in a tissue/cell-type were also reported for analyses in Level 1 and Level 2 in the tables, since these analyses included a variety of cell-types and tissues which may demonstrate meaningful clustering of the highest normalised expression values. The mean normalised expression for all imprinted genes across the series of identity groups in the datasets in Level 1 and Level 2 was also calculated alongside the mean normalised expression for the rest of the genes (Appendix Table A3.2)

3.3 Results

3.3.1 Imprinted gene expression is enriched in the brain, pancreas, bladder and muscle tissue in a multi-organ analysis (Level 1 Analysis)

The Mouse Cell Atlas (MCA) (Han et al., 2018) and the *Tabula Muris* (TM) (Schaum et al., 2018) are single cell compendiums containing ~20 overlapping, but not identical, adult mouse organs. Key overlapping organs include the bladder, brain, kidney, lung, limb muscle, and pancreas while organs included in only one dataset include the ovary, testes, uterus, stomach within the MCA, and the heart, fat, skin, trachea and diaphragm within the TM. These compendiums create a snapshot of gene expression across adult tissues to assess imprinted gene enrichment.

An over-representation analysis was performed on both datasets. All data were processed according to the original published procedure, a list of upregulated genes was produced for each tissue/identity group (vs. all other tissue/identity groups) and a one-sided Fisher's Exact test was performed using a custom list of imprinted genes (Appendix Table A3.1) to identify tissues in which imprinted genes were over-represented amongst the upregulated genes for that tissue. Each dataset in this chapter was analysed independently which allowed me to look for convergent patterns of enrichment between datasets of similar tissues/cell-types. Across only adult tissues, imprinted genes were convergently over-represented in the pancreas, bladder and the brain in both datasets (Figure 3.2A). In addition, in the MCA adult tissue dataset, there was a significant over-representation in the uterus (Table 3.1), and in the *Tabula Muris* analysis (Table 3.2), there was a significant over-representation in the muscle-based tissues - diaphragm, trachea, and limb muscles. In addition to the ORA, to identify situations in which imprinted genes were in fact enriched amongst the stronger markers of a

tissue/cell-type, I performed a Gene-Set Enrichment Analysis (GSEA) on tissues meeting minimum criteria (see Section 2.1.3), which assessed whether imprinted genes were enriched within the top ranked upregulated genes for that tissue (ranked by Log2 Fold Change). No tissue at this level showed a significant GSEA for imprinted genes. Mean normalised expression of imprinted genes across identity groups (Appendix Table A3.2) was the highest for Brain in the MCA and highest for Pancreas in the TM (Brain (Non-Myeloid) was the fourth highest).

Table 3.1. Imprinted gene over-representation in MCA adult tissues (Han et al., 2018). *Identity* – Tissue identities for the cells used in analysis; *Up Reg* – number of upregulated genes with $q \leq 0.05$ and Log2FC ≥ 1 (total number of genes in the dataset in brackets); *IG* – number of imprinted genes upregulated with $q \leq 0.05$ and Log2FC ≥ 1 (total number of IGs in the dataset in brackets); *ORA p* – *p* value from over representation analysis on groups with minimum 5% of total IGs; *ORA q* – Bonferroni corrected *p* value from ORA; *Mean FC IG* – mean fold change for upregulated imprinted genes; *Mean FC Rest* – mean fold change for all other upregulated genes; *No. IGs with highest expression* – Number of IGs with highest mean expression for cells from that identity group.

Tissue Identity	**Up Reg** (20,534)	**IG** (95)	**ORA *p***	**ORA *q***	**Mean FC IG**	**Mean FC Rest**	**No. IGs with highest expression**
Pancreas	2737	42	1.57E-13	**1.89E-12**	8.74	10.32	22
Brain	3401	34	4.43E-06	**5.31E-05**	8.76	125.00	19
Bladder	3183	29	0.000168	**0.002012**	4.45	8.51	8
Uterus	2567	22	0.002827	**0.033919**	4.66	8.46	7
Lung	1203	8	0.192705	1	3.82	151.41	4
Ovary	2219	13	0.223666	1	7.46	11.27	5
Kidney	1714	10	0.268425	1	13.76	182.89	5
Liver	1739	8	0.560145	1	4.55	80.51	3
Stomach	1821	7	0.748590	1	4.24	88.60	3
Thymus	1805	6	0.851579	1	2.78	6.76	2
Small Intestine	1719	5	0.908008	1	7.99	218.64	2
Testis	5212	14	0.995891	1	27.04	5058.36	10
Bone Marrow	1095	2	-	-	5.31	4.43	1
Mammary Gland Virgin	902	4	-	-	3.70	4.03	0
Muscle	1127	4	-	-	8.64	15.05	3
Peripheral Blood	1146	3	-	-	3.78	3.57	0
Prostate	369	0	-	-	0.00	478.10	0
Spleen	1501	1	-	-	4.90	4.77	1

Given the interest in the divergent functions of maternally expressed genes (MEGs) and paternally expressed genes (PEGs), I additionally ran the large-scale enrichment analyses (Levels 1 and 2) using separate lists of PEGs and MEGs. At Level 1, MEGs and PEGs (Appendix Tables A3.3A, A3.3B, A3.4A & A3.4B) revealed a similar pattern of enrichment in both datasets (Figure 3.2B & 3.2C). PEGs were over-represented in the brain in both datasets (MCA - $q = 4.56 \times 10^{-6}$, *TM* - $q = 0.0005$) while MEGs were not. PEGs were also

over-represented in the diaphragm (q = 0.0007), limb muscle (q = 0.0001) and pancreas (MCA - q =1.93x10^{-5}, *TM* - q = 0.0002), with a significant GSEA in the MCA pancreas (p = 0.02, Appendix Figure A3.A). While MEGs were over-represented in the bladder (MCA - q = 0.002, *TM* - q = 0.020), the pancreas (MCA - q = 1.53x10^{-7}) and in the three muscular tissues of the *Tabula Muris* (diaphragm - q = 2.13x10^{-8}, limb muscle - q = 2.43x10^{-7}, trachea - q = 0.004).

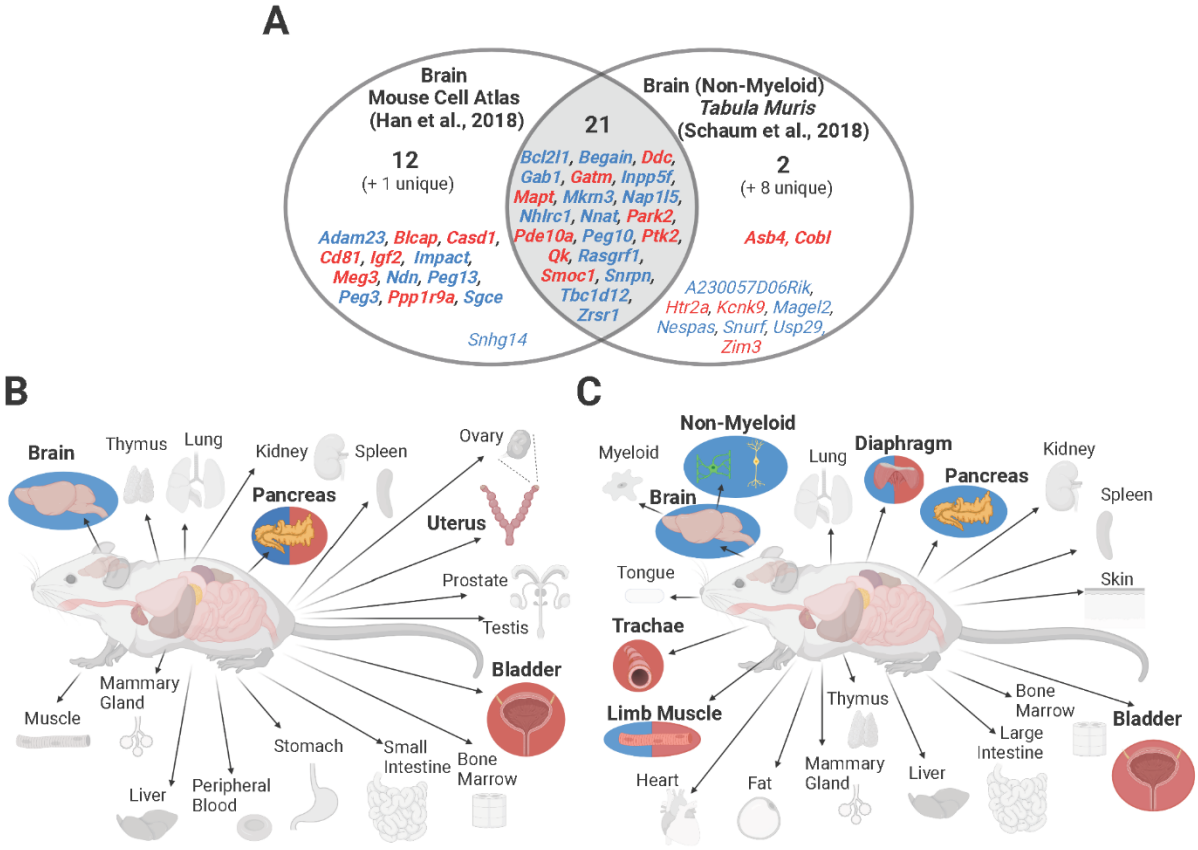


Figure 3.2. Level 1 multi-organ comparison summary graphics. (*A*) Venn diagram of upregulated imprinted genes in the brain in Mouse Cell Atlas and in the brain (non-myeloid) in the *Tabula Muris*. Imprinted genes are listed which show significant upregulation ($q \leq 0.05$ and Log2FC ≥ 1) in the tissues. Although these tissues are not identical, these were the two brain associated over-representations in the enrichment analysis. Parental-bias is indicated by colour (MEG - red, PEG - blue). From the 119 imprinted genes in the gene list, only 92 were common to both analyses (i.e., successfully sequenced and passed gene quality control filters). 34 imprinted genes were upregulated in the brain in the MCA and 31 genes in the *TM*. Genes in common from the two analyses are presented in bold and totalled in each section of the Venn Diagram, while genes found upregulated in one analysis but not available in the other analysis are included in small font and the number indicated in brackets. (*B*) Tissues with over-representation in MCA. Coloured tissues with bold labels were over-represented tissues using all imprinted genes, tissues with a blue circle behind were over-represented in the PEG-only analysis, a red circle represent the same for the MEG-only analysis, and a red/blue split circle were over-represented for both the PEG- and MEG-analyses. (*C*) Tissues with over-representation in *Tabula Muris*. See Figure 3.2B description for details.

Given the enrichment of imprinted genes in the the brain, and the necessity of the brain for parenting behaviour, the remainder of this chapter further delves into the specificities of this brain enrichment, to find the regions and specific cells enriched for imprinted genes. However, the other tissues– bladder, pancreas, muscle tissues - not directly relevant to parenting have been examined in depth too, looking for specific cell populations making up this enrichment and the results of this Level 4 analysis can be found in Appendix A2.

Table 3.2. Imprinted gene over-representation in *Tabula Muris* adult tissues (Schaum et al., 2018). *GSEA p – p* value from Gene Set Enrichment Analysis for identity groups with 15+ IGs and Mean FC IG > Mean FC Rest; *GSEA q* – Bonferroni corrected *p* values from GSEA. All other column descriptions can be found in the legend of Table 3.1.

Tissue Identity	**Up Reg** (20,839)	**IG** (107)	**ORA *p***	**ORA *q***	**Mean FC IG**	**Mean FC Rest**	**GSEA *p***	**GSEA *q***	**No. IGs with max expr…**
Diaphragm	416	19	3.66E-13	**4.75E-12**	6.49	4.83	0.19	0.38	4
Limb Muscle	761	24	6.32E-13	**8.22E-12**	9.02	5.09	0.06	0.11	8
Pancreas	4104	43	8.31E-07	**1.08E-05**	12.52	12.60	-	-	29
Trachea	1979	25	1.78E-05	**0.0002**	3.81	4.57	-	-	5
Brain (Non-Myeloid)	3081	31	0.0001	**0.0016**	12.16	14.17	-	-	14
Bladder	3338	31	0.0005	**0.0068**	3.30	5.30	-	-	16
Fat	1263	12	0.0286	0.3713	3.46	3.68	-	-	1
Heart	1108	10	0.0585	0.7601	2.87	5.14	-	-	0
Mammary Gland	1826	12	0.2264	1	3.52	5.24	-	-	3
Liver	1808	7	0.8307	1	6.19	54.93	-	-	3
Aorta	3515	14	0.8832	1	7.47	16.08	-	-	2
Tongue	4295	15	0.9696	1	4.15	7.16	-	-	8
Large Intestine	4758	11	0.9998	1	5.95	12.22	-	-	5
Brain (Myeloid)	1024	5	-	-	3.39	6.80	-	-	2
Kidney	584	3	-	-	24.94	22.90	-	-	1
Lung	914	2	-	-	2.73	5.41	-	-	0
Marrow	1957	5	-	-	7.65	5.25	-	-	4
Skin	1612	4	-	-	4.15	8.36	-	-	1
Spleen	625	1	-	-	4.63	4.28	-	-	0
Thymus	678	4	-	-	3.46	7.45	-	-	1

3.3.2 Imprinted gene expression is enriched in neurons and neuroendocrine cells of the brain (Level 2 Analysis)

I next analysed cells from the whole mouse brain (Level 2), firstly using the Ximerakis et al. (2019) dataset, in which cells were grouped from the whole mouse brain (minus the hindbrain) into major cell classes according to cell lineage. Imprinted genes were over-represented in neuroendocrine cells and mature neurons (Table 3.3).

Table 3.3. Imprinted gene over-representation in neural lineage types (Ximerakis et al., 2019).
All column descriptions can be found in the legend of Tables 3.1 & 3.2.

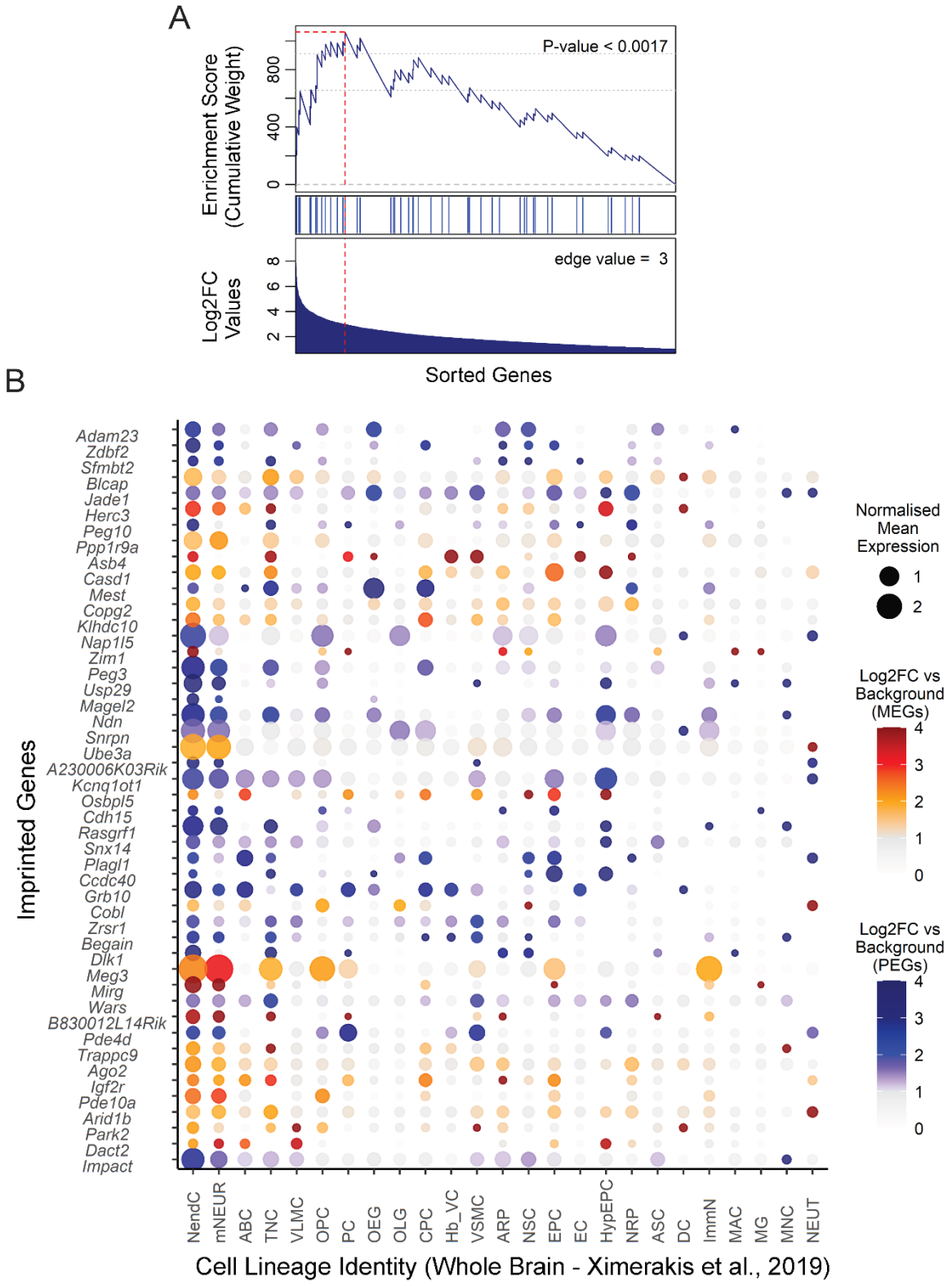
Cell Population Identity (Abbr.)	**Up Reg** (14,498)	**IG** (85)	**ORA *p***	**ORA *q***	**Mean FC IG**	**Mean FC Rest**	**GSEA *p***	**GSEA *q***	**No. IGs with max expr…**
Neuroendocrine cells (NendC)	3868	47	2.12E-08	**3.82E-07**	11.88	5.42	0.0017	**0.0051**	26
Mature Neurons (all types) (mNEUR)	2968	32	0.0002	**0.0035**	8.80	9.28	-	-	2
Arachnoid barrier cells (ABC)	2287	20	0.0396	0.7120	16.84	22.63	-	-	7
Tanycytes (TNC)	1279	12	0.0692	1	6.64	12.01	-	-	8
Vascular and leptomeningeal cells (VLMC)	1714	15	0.0724	1	15.06	13.03	0.0468	0.1404	4
Oligodendrocyte precursor cells (OPC)	1524	13	0.1067	1	3.03	7.17	-	-	1
Pericytes (PC)	1801	14	0.1649	1	8.20	8.22	-	-	2
Olfactory ensheathing glia (OEG)	1086	9	0.1848	1	7.95	26.03	-	-	1
Oligodendrocytes (OLG)	1183	9	0.2561	1	3.73	12.91	-	-	5
Choroid plexus epithelial cells (CPC)	2602	17	0.3524	1	7.43	19.34	-	-	5
Hemoglobin-expressing vascular cells (Hb_VC)	1798	11	0.4889	1	5.25	6.33	-	-	3
Vascular smooth muscle cells (VSMC)	3006	17	0.6093	1	8.94	6.71	0.1376	0.4128	5
Astrocyte-restricted precursors (ARP)	1445	8	0.6214	1	4.50	5.09	-	-	1
Neural stem cells (NSC)	1009	5	0.7138	1	4.00	4.09	-	-	0
Ependymocytes (EPC)	3233	17	0.7346	1	15.04	53.27	-	-	4
Endothelial cells (EC)	1455	7	0.7619	1	5.80	8.54	-	-	0
Hypendymal cells (HypEPC)	1525	6	0.8946	1	17.24	20.80	-	-	5
Neuronal-restricted precursor (NRP)	2339	10	0.8979	1	3.07	10.20	-	-	1
Astrocytes (ASC)	1384	4	-	-	2.22	6.04	-	-	0
Dendritic cells (DC)	1209	1	-	-	3.50	16.02	-	-	1
Immature Neurons (ImmN)	652	4	-	-	3.37	5.79	-	-	0
Macrophages (MAC)	1222	2	-	-	3.47	21.56	-	-	0
Microglia (MG)	1342	3	-	-	19.28	19.22	-	-	3
Monocytes (MNC)	947	2	-	-	16.49	19.13	-	-	1
Neutrophils (NEUT)	519	2	-	-	9.18	62.13	-	-	0

Table 3.4. Imprinted gene over-representation in nervous system cell types (Zeisel et al., 2018). All column descriptions can be found in the legend of Table 3.1.

Cell Population Identity	Up Reg (19,547)	IG (109)	ORA *p*	ORA *q*	Mean FC IG	Mean FC Rest	No. IGs with highest expression
Neurons	5710	44	0.0081	**0.0487**	11.73	24.97	45
Vascular	2473	22	0.0171	0.1029	17.91	26.64	16
Oligos	1587	11	0.2701	1	4.64	11.48	12
Peripheral Glia	2820	16	0.5117	1	5.42	12.64	12
Ependymal	3683	20	0.5912	1	24.52	66.97	15
Immune	1564	7	0.7787	1	13.42	93.05	5
Astrocytes	1539	4	-	-	2.88	10.73	3

Neuroendocrine cells were defined as a heterogeneous cluster, containing peptidergic neurons and neurosecretory cells expressing neuronal marker genes (e.g., *Syt1* and *Snap25*) alongside neuropeptide genes (e.g., *Oxt*, *Avp*, *Gal*, *Agrp* and *Sst*) but distinguished by Ximerakis et al. (2019) by the unique expression of *Baiap3* which plays an important role in the regulation of exocytosis in neuroendocrine cells (Zhang et al., 2017). GSEA additionally showed that the imprinted genes were enriched in the genes with the highest fold change values for neuroendocrine cells only (Fig. 3.3). 26 imprinted genes had their highest expression in the neuroendocrine cells and the mean normalised expression of imprinted genes was almost twice as high for neuroendocrine cells as the next highest identity group (Appendix Table A3.2) The MEG/PEG analysis (Appendix Table A3.5A and A3.5B) for this dataset found that PEGs were over-represented in mature neurons ($q = 0.027$) and neuroendocrine cells ($q = 8.97 \times 10^{-6}$). MEGs were also over-represented in neuroendocrine cells ($q = 0.047$) and uniquely over-represented in Arachnoid barrier cells ($q = 0.014$). Only PEGs replicated the significant GSEA in neuroendocrine cells ($p = 4 \times 10^{-4}$, Appendix Figure A3.B).

Figure 3.3. (Following Page) GSEA and dot plots for imprinted genes upregulated in neuroendocrine cells in the Ximerakis et al. (2019) whole mouse brain dataset. (*A*) GSEA for imprinted genes upregulated in the neuroendocrine cells. In the analysis, genes are sorted by strength by which they mark this neuronal cluster (sorted by Log2FC values) indicated by the bar (middle). Fold change values are displayed along the bottom of the graph. The genes are arrayed left (strongest marker) to right and blue lines mark where imprinted genes fall on this array. The top vertical axis indicates an accumulating weight, progressing from left to right and increasing or decreasing depending on whether the next gene is an imprinted gene or not. The *p*-value represents the probability of observing the maximum value of the score (red dashed line) if the imprinted genes are distributed randomly along the horizontal axis. The *q*-value for this analysis was significant at 0.0036. (*B*) Dot plot of imprinted genes upregulated in the 'Neuroendocrine cells' plotted across all identified cell types (Abbreviations found in Table 3.3). Imprinted genes were plotted in chromosomal order. Size of points represented absolute mean expression; colour represented the size of the Log2FC value for the cell identity group (e.g., neuroendocrine cells) vs. all other cells. Unique colour scales are used for MEGs (red/orange) and PEGs (blue). Where a gene was not expressed in a cell type, this appears as a blank space in the plot



The second dataset at this level was the Zeisel et al. (2018) Mouse Brain Atlas (MBA) which allowed a much deeper investigation of nervous system enrichment with sequencing of the entire murine nervous system and identifying cells by both brain region and cell type. Concordant with the previous findings, primary analysis separating cells by lineage revealed

over-representation of imprinted genes in neurons only (Table 3.4). The overlap between the upregulated imprinted genes for the over-represented neural-lineage cells from the Level 2 datasets are displayed in Figure 3.4. Additionally, PEGs alone demonstrated no significant over-representations in cell lineage types while MEGs demonstrated over-representation in vascular cells only (q = 0.0004) (Appendix Tables A3.6A and A3.6B).

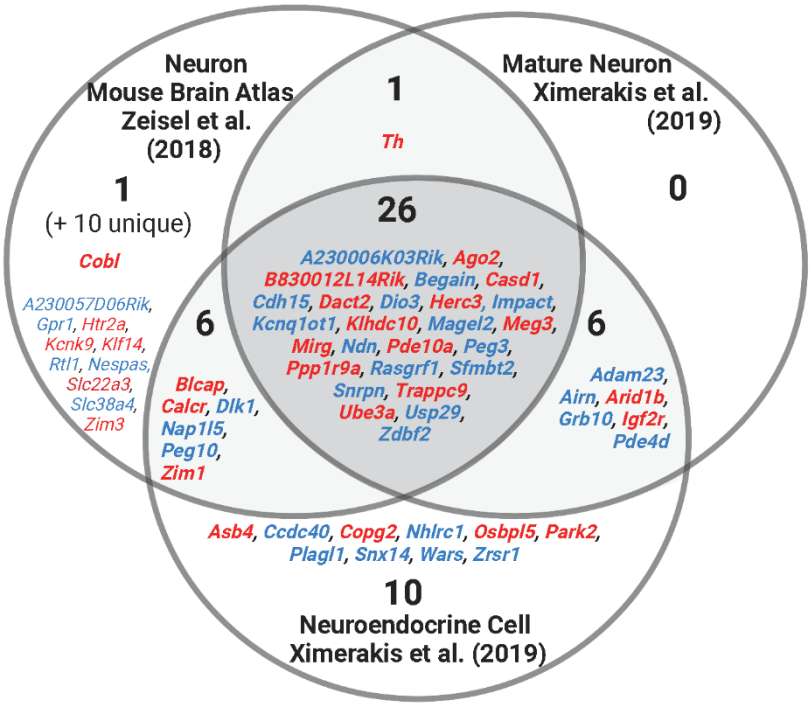


Figure 3.4. Venn diagram of upregulated imprinted genes in the mature neuronal cells in the whole brain datasets of Zeisel et al. (2018) and Ximerakis et al. (2019). Imprinted genes are listed which show significant upregulation ($q \leq 0.05$ and Log2FC ≥ 1) in the cells. Although these cell types are not identical, these were all mature neural lineage cells with over-representations in the enrichment analysis. Parental-bias is indicated by colour (MEG - red, PEG - blue. From the 119 imprinted genes in the gene list, only 88 were common to both analyses (i.e., successfully sequenced and passed gene quality control filters). 45 imprinted genes were upregulated in neurons in the MBA, and in Ximerakis et al. (2019), 33 imprinted genes were upregulated in neurons and 48 genes in neuroendocrine cells. Genes in common from the two analyses are presented in bold and totalled in each section of the Venn Diagram, while genes found upregulated in one analysis but not available in the other analysis are included in small font and the number indicated in brackets.

3.3.3 The hypothalamus, ventral midbrain, pons and medulla are enriched for imprinted gene expression (Level 2 Analysis)

After confirming neuron-specific enrichment of imprinted genes in the MBA dataset, further MBA analysis was performed on cells classified as neurons and then grouped by brain/nervous system regions. Significant over-representation was seen in neurons of the

hypothalamus, ventral midbrain, medulla, and pons (Table 3.5). The pons and medulla had the largest number, 45 and 44 respectively, of imprinted genes upregulated (Figure 3.5A).

Regional analysis for MEGs and PEGs separately (Appendix Tables A3.7A and A3.7B), revealed that PEGs were over-represented in hypothalamus ($q = 6.53x10^{-7}$), ventral midbrain ($q = 0.018$), the pons ($q = 4.65x10^{-5}$) and the medulla ($q = 4.10x10^{-6}$); while MEGs were only over-represented in the medulla ($q = 0.002$) but had a significant GSEA for the pons ($q = 0.027$, Appendix Figure A3.C); see Figure 3.5B.

Table 3.5. Imprinted gene over-representation in cells clustered by nervous system region identity (Zeisel et al., 2018). All column descriptions can be found in the legend of Tables 3.1 & 3.2.

Brain Region Identity	**Up Reg** (18,335)	**IG** (106)	**ORA *p***	**ORA *q***	**Mean FC IG**	**Mean FC Rest**	**GSEA *p***	**GSEA *q***	**No. IGs with max expr…**
Medulla	3147	45	8.38E-10	**1.26E-08**	4.79	4.01	0.1	0.2	15
Hypothalamus	1040	22	9.81E-08	**1.47E-06**	4.92	5.84	-	-	8
Pons	3581	44	1.62E-07	**2.43E-06**	4.20	3.91	0.12	0.23	22
Vent. Midbrain	1228	18	0.0002	**0.0034**	4.90	4.99	-	-	3
Vent. Striatum	689	8	0.0463	0.6941	3.92	4.92	-	-	0
Posterior Cortex	1090	9	0.1788	1	2.64	3.20	-	-	2
Enteric Nervous System	3885	26	0.2311	1	8.98	121.04	-	-	11
Sympathetic Nervous System	2804	18	0.3535	1	11.37	57.96	-	-	9
Anterior Cortex	979	6	0.5016	1	2.72	3.30	-	-	1
Dors. Midbrain	1045	6	0.5663	1	2.20	4.85	-	-	3
Thalamus	1441	8	0.6000	1	2.90	6.36	-	-	0
Hippocampus - CA1	1082	6	0.6008	1	3.01	4.02	-	-	2
Somatosensory Cortex	2121	11	0.6943	1	4.09	3.70	-	-	8
Dors. Striatum	1196	6	0.6974	1	4.03	5.43	-	-	2
Dorsal Root Ganglion	3607	16	0.9088	1	11.56	75.89	-	-	9
Middle Cortex	623	5	-	-	3.29	3.24	-	-	0
Spinal Cord	972	5	-	-	4.57	12.36	-	-	1
Amygdala	452	4	-	-	4.65	4.11	-	-	2
Dentate Gyrus	796	4	-	-	3.79	4.16	-	-	2
Hippocampus	631	4	-	-	4.86	3.82	-	-	2
Olfactory Bulb	445	4	-	-	4.02	8.27	-	-	2
Antero-Middle Cortex	646	3	-	-	4.59	4.31	-	-	1
Cerebellum	240	0	-	-	0.00	32.30	-	-	0

Neurons were then re-clustered into unique subpopulations identified by marker genes (Zeisel et al., 2018) to uncover the specific neural populations underlying the enrichment seen in the hypothalamus, pons and medulla, and midbrain (Figure 3.6; Appendix Table A3.8). Each

neural population was identified by its distinct gene expression and suspected location within the brain (see http://mousebrain.org/ for an online resource with detailed information on each cluster).

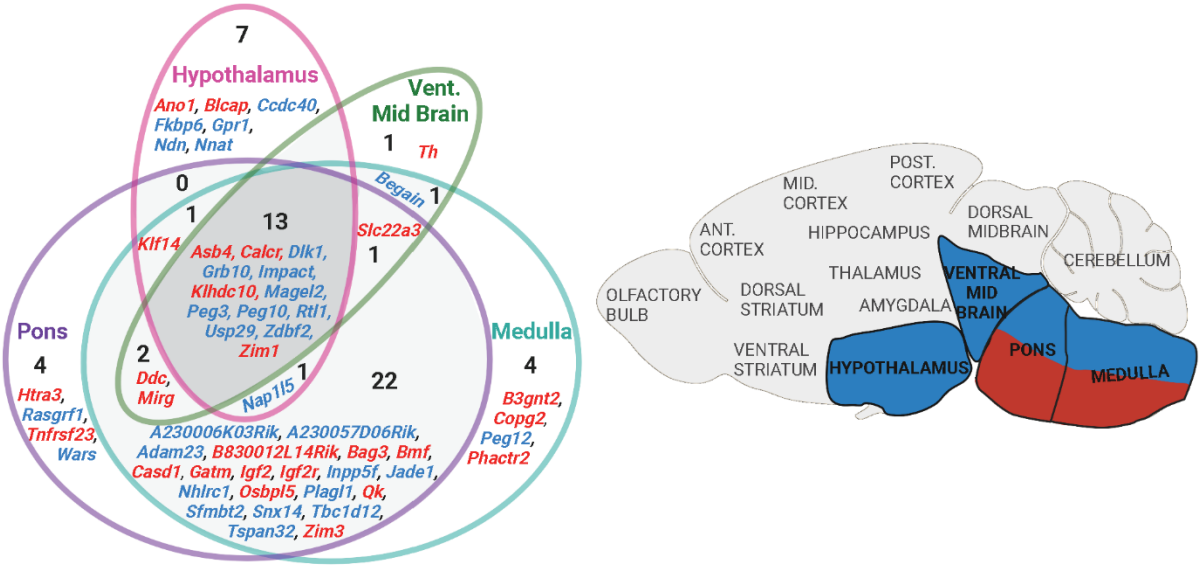


Figure 3.5. **Level 2 Brain Region Analysis summary figures.** (*A*) Venn diagram of upregulated imprinted genes in the neurons of enriched nervous system regions from the Mouse Brain Atlas (Zeisel et al., 2018). Imprinted genes are listed which show significant upregulation ($q \leq 0.05$ and Log2FC ≥ 1) in the regions specified. The number of imprinted genes in each region of the Venn diagram are specified. Parental-bias of imprinted genes is indicated by colour (MEG - red, PEG - blue). (*B*) Brain regions enriched for imprinted gene expression via ORA or GSEA in the MBA (Zeisel et al., 2018). Regions over-represented for all imprinted genes are bolded. Regions over-represented for PEG expression alone are coloured blue while regions enriched for MEG expression alone are coloured red.

The hypothalamus was represented by a selection of inhibitory and peptidergic neurons. Inhibitory neurons with over-representation of imprinted genes included: a Subthalamic Nucleus population (notable genes *Lhx8, Gabrq*), two Preoptic Area/ BNST populations (*Nts, Dlk1* / *Gal, Irs4*), an Arcuate nucleus population (*Agrp, Otp*), and two Suprachiasmatic nucleus populations (*Avp, Nms, Six6, Vip*). For peptidergic neurons, over-representation was seen in a ventromedial population (*Gpr101, Tac1, Baiap3*), a ventromedial/paraventricular population (*Otp, Trh, Ucn3*), a lateral hypothalamic population (*Trh, Otp, Ngb*), an oxytocin magnocellular population of the paraventricular and supraoptic nuclei (*Oxt, Otp*), and an orexin producing population of the dorsomedial/lateral hypothalamus (*Hcrt, Pdyn, Trhr*).

The midbrain, medulla and pons were represented by a number of cell groups, with over-representation seen in the medulla-based adrenergic (HBAR) and noradrenergic (HBNOR) groups and the dopaminergic neurons of the midbrain in the Periaqueductal Gray (PAG) (MBDOP1) and the Ventral Tegmental Area (VTA)/Substantia Nigra (SNc) (MBDOP2).

There were also several inhibitory (MEINH, HBIN) and excitatory neuron (MEGLU, HBGLU) types spread across the nuclei from the three regions (Figure 3.6). The serotonergic populations of the raphe nuclei of these regions (HBSER) were particularly prominent since the pons and medulla-based serotonin neuron populations (HBSER2, HBSER4 and HBSER5) were the only neuron subpopulations out of the 214 total to have a significant GSEA for imprinted genes after correction (Appendix Figure A3.D).

Additional regions of over-representation included neurons in the pallidum and striatum and PVN neurons from the thalamus. It was interesting to note that within this comparison of 214 neuron populations, no neurons from areas such as the cortex, cerebellum or peripheral nervous system were enriched, and neither were they over-represented in the previous regional analysis.

Using this regional enrichment as a guide, the final analysis I carried out in the brain focused on the hypothalamus and monoaminergic nuclei of the brain, both areas have neurons critical in regulating parenting behaviour.

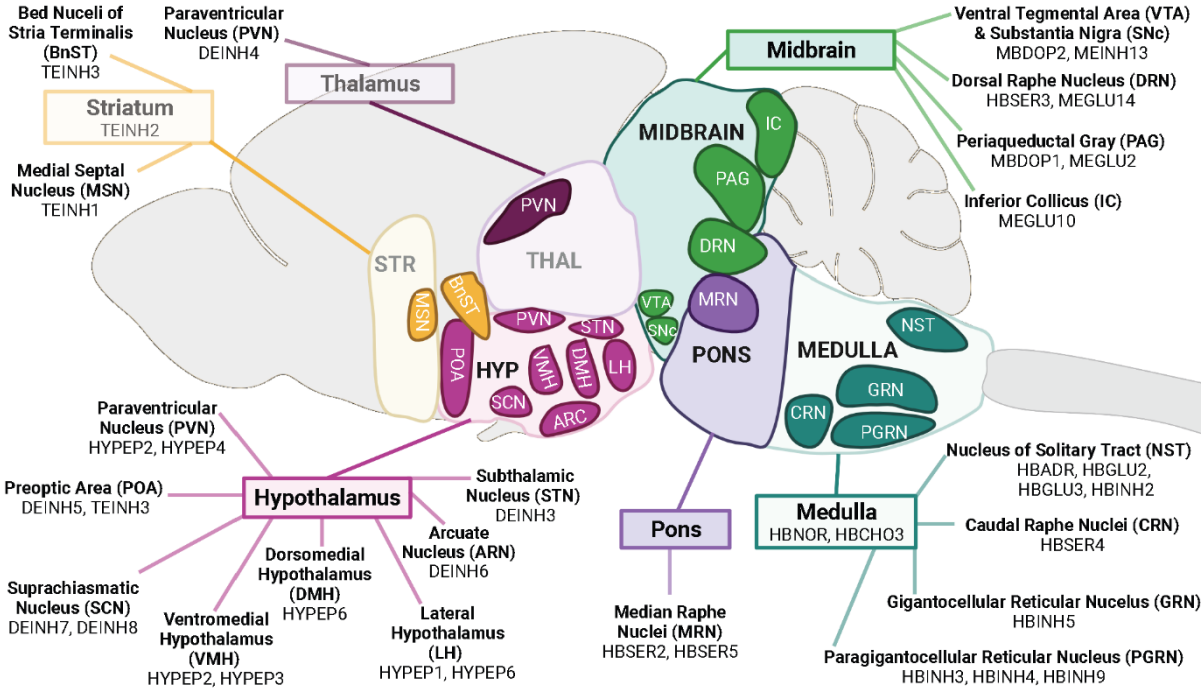


Figure 3.6. Anatomical labelling of all the neural subpopulations with a significant over-representation of imprinted genes ($q \leq 0.05$ and Log2FC ≥ 1) in the Mouse Brain Atlas (Zeisel et al., 2018). The predicted brain nuclei localisation of the 32 neuronal subpopulations (out of 214 populations identified across the nervous system) specified in the MBA and enriched for imprinted genes. Brain regions that were not found to be enriched for imprinted genes are greyed out. The full Enrichment Analysis is available in Appendix Table A3.8

3.3.4 Imprinted gene expression is over-represented in specific hypothalamic neuron subtypes including POA galanin neurons (Level 3A & 3B Analysis)

I first investigated whether any specific neuron populations of the hypothalamus would show enrichment for imprinted gene expression when compared to other hypothalamic neurons. Two datasets with single cell sequencing data for the adult hypothalamus existed (Chen et al., 2017, Romanov et al., 2017). Both clustered their data into neuronal subpopulations allowing me to look for convergent imprinted enrichment across major hypothalamic neuronal subtypes (Level 3A). Analyses revealed a clear neuronal bias in expression of imprinted genes (Appendix Tables A3.9A and A3.10A). Within the Romanov et al. (2017) data, there was a significant over-representation of imprinted genes in neurons (q = 0.02) and a similar observation was seen in the Chen et al. (2017) data (q = 0.001), and both also demonstrated a significant GSEA in neurons (Figure 3.7A-D, Romanov et al. (2017) – p = 0.011, Chen et al. (2017) - p = 0.022).

Within the Chen et al. (2017) dataset, 4/33 hypothalamic neuronal subtypes had a significant over-representation of imprinted genes (Appendix Table A3.9B). The four subtypes were all GABAergic neurons, specifically: a dopaminergic neuron type (*Slc6a3*) with high expression of *Th* and *Prlr* suspected to be the Tuberoinfundibular Dopamine Neurons (TIDA) neurons of the arcuate nucleus (q = 0.0001); SCN neurons (*Vipr2)* with very high *Avp* and *Nms* expression (q = 0.0071); galanin neurons (*Slc18a2/Gal)* present in a several hypothalamic regions (q = 0.0079); and *Agrp* feeding promoting neurons of the Arcuate Nucleus (q = 0.034). Within the Romanov et al. (2017) dataset, 3/62 subtypes had significant over-representation of imprinted gene expression (Appendix Table A3.10B): A*grp*/*Npy* neurons (q = 0.013), the Arcuate Nucleus feeding neurons also reported in Chen et al. (2017); a *Ghrh/Th* neuronal type (q = 0.032), again likely corresponding to neurons from the arcuate nucleus and the top hit ($q = 1.63 \times 10^{-6}$) was a poorly segregated population (*Calcr/Lhx1)*, likely due to a deeper inner cluster heterogeneity. This cluster was interesting since the imprinted genes *Calcr* and *Asb4* were amongst its most significant marker genes, and it was notably the only cluster with high expression of all three of *Th, Slc6a3* and *Prlr.* Romanov et al. (2017) did not identify any of their populations as the TIDA neurons, but the above pattern of gene expression suggests that this cluster may contain these neurons. Furthermore, the suspected

TIDA neurons from the Chen et al. (2017) dataset shared 21/40 upregulated genes of this unresolved cluster (see Appendix Table A3.11 for full comparison).

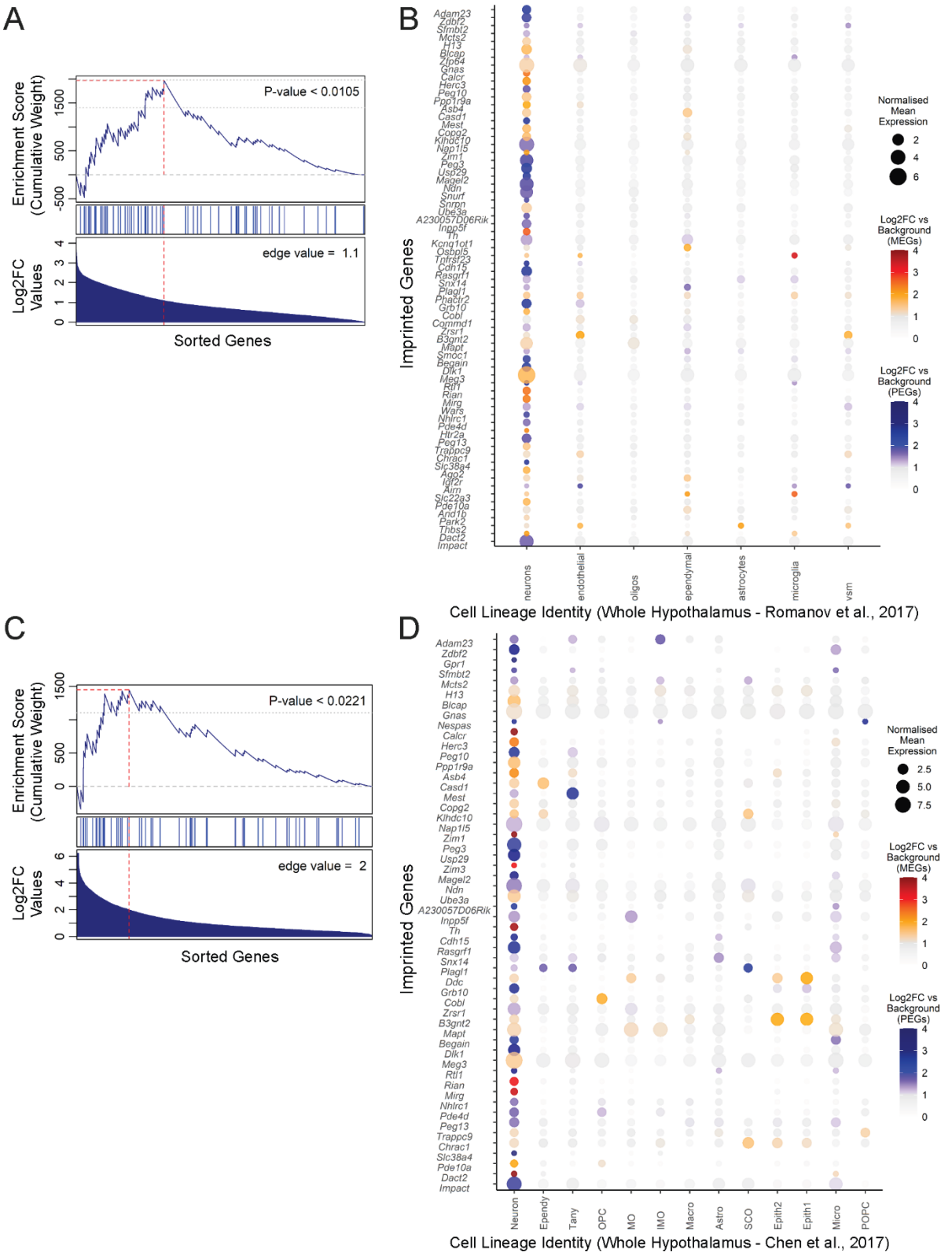


Figure 3.7. (Previous Page) GSEA and Dot plots for imprinted genes upregulated in neurons across the whole hypothalamus. (*A*) GSEA for imprinted genes upregulated in the 'Neuron' cell type in the whole hypothalamic dataset of Chen et al. (2017). See legend of Figure 3.3A for a description of how to interpret the plot. (*B*) Dot plot of imprinted genes upregulated in the 'Neuron' cell type plotted across all identified cell types in the Chen et al. (2017) whole hypothalamic dataset. See legend of Figure 3.3B for a description of how to interpret the plot. Abbr: OPC = Oligodendrocyte Precursor Cell, MG = Myelinating Oligodendrocyte, IMG = Immature Oligodendrocyte, Astro = Astrocyte, Epith = Epithelial, Macro = Macrophage, Tany = Tanycyte, Ependy = Ependymocyte, Micro = Microglia, POPC = Proliferating Oligodendrocyte Progenitor Cell. (*C*) GSEA for imprinted genes upregulated in 'neurons' in the whole hypothalamic dataset of Romanov et al. (2017). See legend of Figure 3.7A for a description of how to interpret the plot. (*D*) Dot plot of imprinted genes upregulated in 'neurons' plotted across all identified cell types in the Romanov et al. (2017) whole hypothalamic dataset. See legend of Figure 3.7B for a description of how to interpret the plot.

Galanin neurons from the POA have consistently been shown to be the key neuron type in this area for coordinating parenting behaviour (Wu et al., 2014, Kohl et al., 2018). Having found enrichment in galanin neurons in the whole brain and whole hypothalamic level analyses, I wanted to know whether this enrichment would occur for galanin neurons of the POA, and so analysed the only existing scRNA-seq dataset for the POA (Moffitt et al., 2018). Additionally, the other consistently enriched neuron populations at multiple analyses came from the arcuate nucleus (*Agrp, Th & Prlr, Th & Ghrh*), and suprachiasmatic nucleus (*Avp, Vip*) and so I additionally tested imprinted gene enrichment within these hypothalamic regions at a high resolution using datasets sequencing neurons purely from these hypothalamic regions (Level 3B).

Preoptic Area (POA) (Campbell et al., 2017)

Imprinted gene over-representation was found in 14/66 of the neuron clusters identified by Moffitt et al. (2018) in the preoptic area (Appendix Table A3.12A). Moffitt et al. (2018) stated that 9 of the 66 clusters likely originate from outside the POA including 3/14 of the imprinted gene enriched neuron clusters (one of which was the *Avp*/*Nms* neurons (q = 0.0003) likely to be from the SCN). To focus my analysis on the POA alone, I removed the non-POA neuron groups and ran the analysis on the remaining neural groups restricted to POA (Appendix Table A3.12B). Imprinted genes were found to be over-represented in 15 neuron types including two galanin enriched clusters (i22, i8), the major *Th* enriched cluster (i17), e15:*Ucn3*/*Brs3* (a cluster activated by parenting), a *Crh* cluster (i35) and a *Trh* cluster (e4), and both major glutamatergic *Ghrh* clusters (e13, e19). Since a considerable number of neuron types were over-represented, the above analysis was run again with the Log2FC ≥ 1 criteria to find out if imprinted genes were enriched in the more distinct markers (Appendix

Table A3.12C) Only two neuron types were over-represented: i35:*Crh*/*Tac2* (q = 0.0126) and i16:*Gal*/*Th* (q = 0.0026). i16:*Gal*/*Th* was identified by Moffitt et al. (2018) as the prominent parenting behavioural neuron group (expressing high *c-Fos* following parenting behaviour in mothers, fathers and virgin females).

Arcuate nucleus (ARC) (Campbell et al., 2017)

The arcuate nucleus was sequenced by Campbell et al. (2017). Imprinted gene over-representation was found in 8/24 arcuate neuron types (Appendix Table A3.13). These included the *Agrp*/*Sst* neuron type (with high expression of *Npy*, q = 0.003) and two *Pomc* neuron types (*Pomc/Anxa2*, q = 0.004; *Pomc/Glipr1*, q = 0.03). *Pomc* expressing neurons are known to work as feeding suppressants (Rau and Hentges, 2017). Additional significant over-representation was found in the *Ghrh* neuron type (q = 0.009), which was also enriched in *Gal* and *Th*. Finally, a highly significant over-representation of imprinted genes was found in the *Th*/*Slc6a3* neuron type ($q = 1.72 \times 10^{-8}$) identified by the authors as one of the most likely candidates for the TIDA dopaminergic neuron population. Marker genes for this identity group overlapped with the TIDA candidates from the previous two datasets (e.g., *Slc6a3*, *Th*, *Lhx1*, *Calcr*). *Agrp* neurons, *Ghrh* neurons and these TIDA candidate neurons were identified in both whole hypothalamic datasets and at the ARC level.

Suprachiasmatic Nucleus (SCN) (Wen et al., 2020)

Analysis of the 10x chromium data of SCN neurons (Appendix Table A3.14) revealed a significant over-representation ($q = 1.51 \times 10^{-8}$) and GSEA (p = 0.004, Appendix Figure A3.E) in the *Avp*/*Nms* neuronal cluster (out of 5 neuronal clusters). This cluster shows the strongest expression for *Oxt*, *Avp*, *Avpr1a* and *Prlr* and is one of the three neural group that Wen et al. (2020) found had robust circadian gene expression, and the only subtype with notable phase differences in circadian gene expression in the dorsal SCN. This cluster likely corresponds to the GABA8 cluster found enriched in the Chen et al. (2017) dataset.

Figure 3.8 presents the overlapping upregulated imprinted genes from the convergently upregulated neuron subtypes in the hypothalamic analysis of Level 2, 3A and 3B.

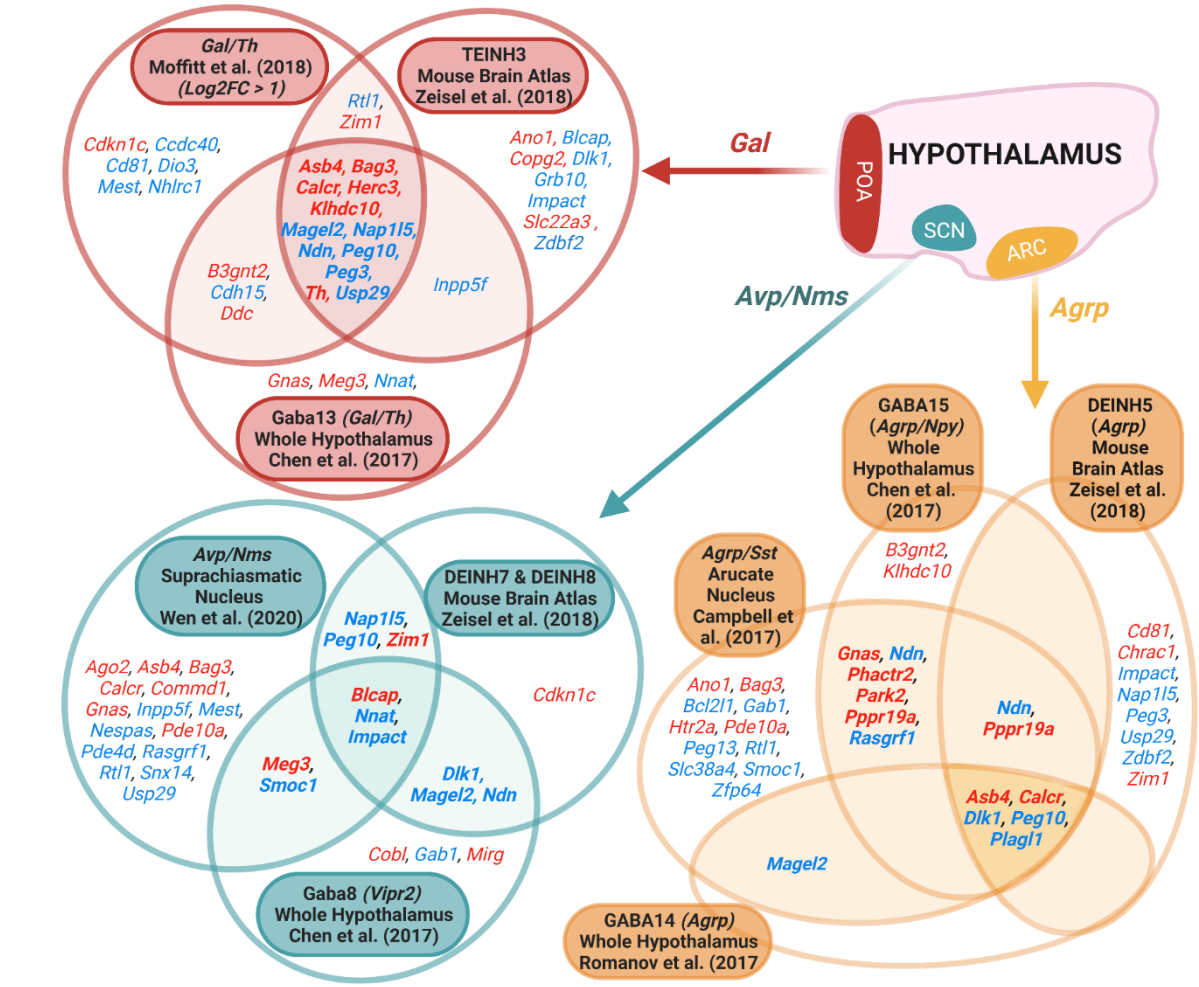


Figure 3.8. Venn diagrams of upregulated imprinted genes in the neuronal subpopulations from level 3B that were also identified in level 2 and 3A. Imprinted gene overlap was contrasted for *Gal* neuronal population of the POA (also expressing *Calcr*, *Brs3*, *Th)* (Zeisel et al., 2018, Chen et al., 2017, Moffitt et al., 2018), *Agrp* neuronal populations of the Arcuate Nucleus (Zeisel et al., 2018, Campbell et al., 2017, Romanov et al., 2017, Chen et al., 2017) and *Avp/Nms* neurons from the Suprachiasmatic Nucleus (Zeisel et al., 2018, Wen et al., 2020, Chen et al., 2017) Imprinted genes are listed which show significant upregulation ($q \leq 0.05$ and Log2FC > 0) in the subpopulation. Parental-bias is indicated by colour (MEG - red, PEG - blue).

3.3.5 Imprinted gene expression is over-represented in monoaminergic nuclei of the mid- and hindbrain (Level 3C Analysis)

In the MBA, Whole Hypothalamus and Arcuate Nucleus analyses, dopaminergic clusters were consistently enriched. Since parenting is an innate motivated behaviour, midbrain dopaminergic functioning is essential (Numan and Stolzenberg, 2009). To explore this further, analysis of Hook et al. (2018) data allowed comparison for dopamine neurons across the brain (specifically from the olfactory bulb, arcuate nucleus and midbrain) at two developmental timepoints (E15.5 and Post-natal day (P) 7). The arcuate nucleus P7 dopamine

neurons emerged as the clearest over-represented subgroups (Appendix Table A3.15). This included the *Th*/*Slc6a3/Prlr* neurons ($q = 1.15x10^{-8}$) and the *Th/Ghrh/Gal* cluster ($q = 4.79x10^{-5}$) the latter of which were referred to as 'neuroendocrine' cells by Hook et al. (2018), and the former a mixture of arcuate nucleus populations with *Prlr* was one of the marker genes, suggesting this includes the TIDA neurons. Additionally, P7 midbrain dopamine neurons were the other group with significant over-representation (specifically from the PAG and VTA) as well as the neuroblasts at this time point.

Although no specific adult mouse midbrain datasets exist, ventral midbrain sequencing at E11.5 - E18.5 by La Manno et al. (2016) allowed me to identify imprinted enrichment within the midbrain at a timepoint when the major neuronal populations are differentiating but still identifiable (Appendix Table A3.16). As anticipated, I found significant over-representation in both mature (DA1; high *Th* and *Slc6a3*, $q = 0.0103$), and developing (DA0, $q = 0.0129$) dopaminergic neurons, as well as the serotonergic neurons ($q = 3.09x10^{-7}$), likely from the midbrain raphe nuclei.

Raphe nuclei from the midbrain/hindbrain are key serotonergic regions of the brain. Serotonin has also been shown to be essential for parenting behaviour (Lerch-Haner et al., 2008). Analysis of all cell types in the Dorsal Raphe Nucleus (DRN) sequenced by Huang et al. (2019) revealed a clear enrichment of imprinted genes in the neuronal populations of the DRN as compared to the non-neuronal cell populations of the DRN (Appendix Table A3.17A). When compared to all other cell populations, significant ORA was seen for Dopaminergic ($q = 0.009$), Serotonergic ($q = 0.012$) and Peptidergic neurons ($q = 0.0008$), however, a significant GSEA was found for all five neuronal populations (Appendix Figure A3.F). When compared against each other (i.e., serotonergic upregulation vs. the other neurons), only the serotonergic neurons of the DRN ($q = 0.0019$) were found to have a significant over- representation of imprinted genes (Appendix Table A3.17B). GSEA's were non-significant but the mean fold change for imprinted genes was markedly higher in both serotoninergic (52% higher) and dopaminergic neurons (68% higher). When contrasting neuronal subpopulations of the DRN, two of the five serotonin subpopulations had significant over-representation of imprinted genes: *Hcrtr1/Asb4* ($q = 0.0014$) and *Prkcq/Trh* ($q = 0.007$) (Appendix Table A3.17C). These clusters were identified by Huang et al. (2019) as the only clusters localised in the dorsal/lateral DRN and the serotonin clusters enriched in *Trh*. Huang et al. (2019) hypothesised that these were the serotonin neurons that project to hypothalamic nuclei, and motor nuclei in the brainstem (as opposed to cortical/striatal projection).

3.3.6 Imprinted gene expression is over-represented in lactotrophs and somatotrophs of the pituitary gland (Level 3D Analysis)

Following on from the enrichment seen above for imprinted gene expression in the dopaminergic arcuate nucleus neurons coordinating pituitary gland output, I sought to identify whether any cells in the pituitary would display matching over-representation for imprinted gene expression (Level 3D). The pituitary was not sequenced as part of the multi-organ or whole brain datasets analysed above and so two independent datasets were analysed that specifically sequencing the mouse pituitary at single cell resolution. Ho et al. (2020) recently sequenced the anterior pituitary gland of male and female C57BL/6 mice using two sequencing technologies, both 10X genomic and Drop-Seq. This identified a variety of cell types from the endocrine and non-endocrine pituitary. I analysed data from both technologies and found that imprinted gene expression was convergently over-represented in the Lactotrophs (prolactin secreting) and Somatotroph (growth hormone secreting) cells (Appendix Tables A3.18A & A3.18B). In a second independent dataset sequencing cells from male mouse pituitary glands (Cheung et al., 2018), I found significant over-representation in the Somatotropes and Thyrotrope (secreting thyroid stimulating hormone) (Appendix Table A3.19). Figure 3.9 demonstrates the overlap in imprinted genes significantly expressed in Somatotropes and Lactotropes across the datasets since these were the only cell-types to be over-represented in more than one dataset. It is notable that the two cell types represented here directly match the two regulatory neurons found over-represented in the arcuate nucleus of the hypothalamus.

As discussed above, Level 4 Analysis, exploring the specific cell type enrichments for the bladder, pancreas, muscle tissues and mammary gland can be found in Appendix A2.

3.4 Discussion

This work utilised an unbiased systems biology approach using publicly available scRNA-seq data and showed that imprinted genes, when treated as a gene set, do show enriched expression in specific cell populations in the adult mouse body and brain. Imprinted genes were over-represented in several mouse tissues: pancreas, bladder, muscle and, most importantly for this thesis, the adult mouse brain. Imprinted gene expression was further over-represented in neurons at every level tested with a marked enrichment in neuroendocrine

cell lineages. Within-brain analyses revealed that the hypothalamus and the monoaminergic system of the mid- and hindbrain were foci for imprinted gene enrichment. While not all imprinted genes follow these patterns of expression, these findings highlight collective gene expression which is non-random in nature. As such, these analyses identify 'expression hotspots', which in turn suggest 'functional hotspots'.

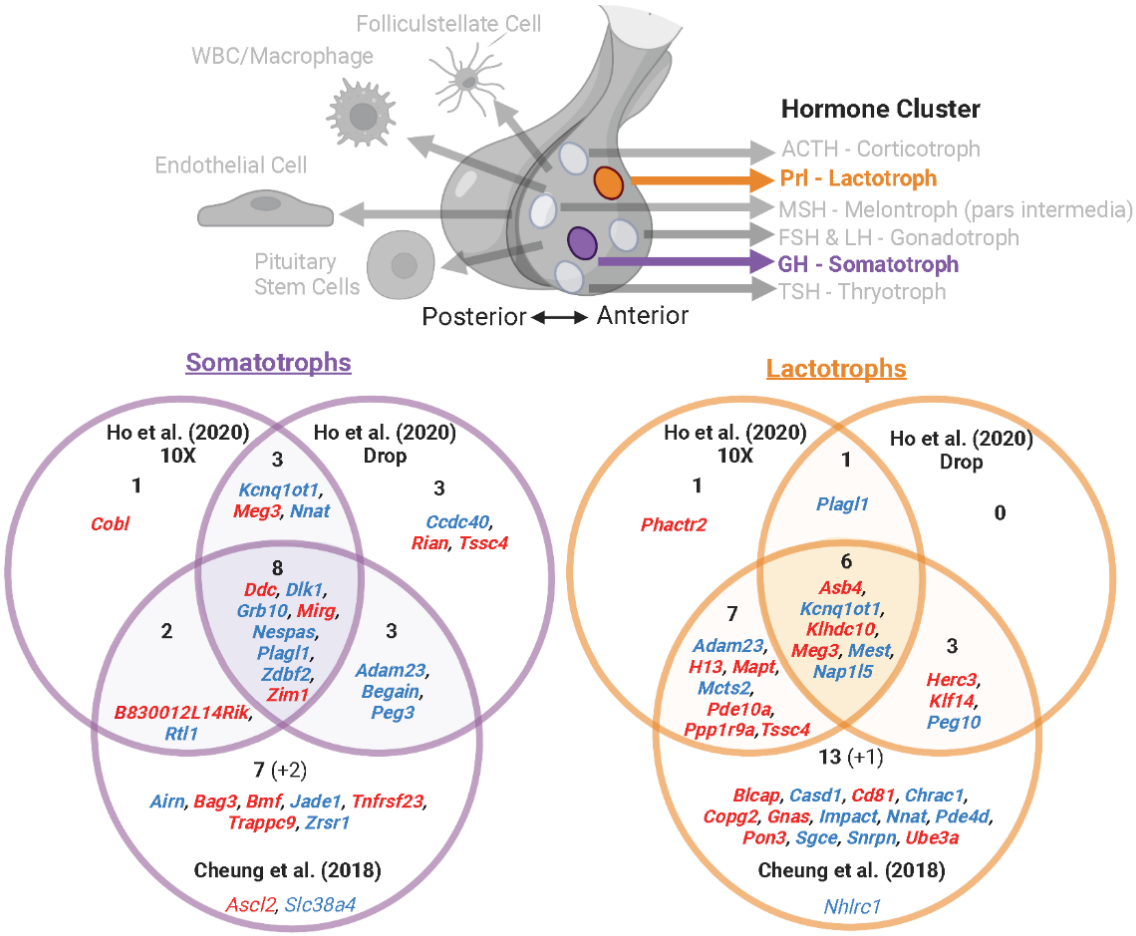


Figure 3.9. Level 3D analysis of Pituitary Gland. *(A)* Pituitary cell types showing over-representation for imprinted gene expression in multiple pituitary datasets. Over-represented cell types are bold and not in greyscale. The hormone/s released from the endocrine cell types are also indicated. *(B)* Venn diagram of upregulated imprinted genes in the Somatotrophs and Lactotrophs in Cheung et al. (2018) and Ho et al. (2020). Imprinted genes are listed which show significant upregulation ($q \leq 0.05$ and Log2FC > 0) in the cell types. Parental bias is indicated by colour (MEG –red, PEG – blue). Genes in common from two analyses are presented in bold and totalled in each section of the Venn Diagram, while genes found upregulated in one analysis but not available in the others are included in small font and the number indicated in brackets.

The idea that imprinted genes converge on specific physiological or behavioural processes is not unprecedented. Specialisation of function is predicted when considering why genomic imprinting evolved at all (Keverne, 2014, Keverne et al., 1996b, Moore and Haig, 1991,

Trivers and Burt, 1999). Moreover, there is increasing evidence that the imprinted genes themselves appear to be co-expressed in an imprinted gene network (IGN) and have confirmed regulatory links between each other (Al Adhami et al., 2015, Varrault et al., 2006, Gabory et al., 2009). The idea of an IGN or, at the very least, heavily correlated and coordinated expression between imprinted genes adds further support to the idea that imprinted genes work in concert to influence processes, rather than in isolation, and that perturbating one may influence many others (Patten et al., 2016).

When it came to specific cell populations, strikingly, and suggestive of meaningful enrichment, I saw convergence across my distinct levels of analysis with several hypothalamic neuronal types identified in the whole brain, whole hypothalamus, and hypothalamic-region-level analysis. Enrichment was consistently found in *Agrp* expressing neurons within the arcuate nucleus. These neurons operate as feeding promotors and a few imprinted genes have previously been associated with their function (*Asb4*, *Magel2*, *Snord116*) (Vagena et al., 2022, Cassidy and Charalambous, 2018) but IGs have never be found to be enriched here. Circadian processes are controlled principally by the Suprachiasmatic Nucleus and here I found strong imprinted gene enrichment in *Avp/Nms* expressing neurons (an active circadian population). This population is of interest given the growing appreciation of the role imprinted genes play in circadian processes and the SCN suggested by studies of individual imprinted genes (Tucci, 2016). Pituitary endocrine regulation also emerged as a key function, considering the over-representation in the dopaminergic: *Th/Slc6a3/Prlr* neuron type (top hit in the arcuate nucleus and across dopaminergic neurons of the brain) and the *Th/Ghrh* subpopulation. These neuron populations can regulate prolactin (regulating lactation, stress, weight gain, parenting and more (Grattan et al., 2008, Grattan and Kokay, 2008)) and growth hormone (promoting growth and lipid/carbohydrate metabolism (Vijayakumar et al., 2011, Waxman and Frank, 2000)) release, respectively. Remarkably, I also found a matching enrichment in the lactotroph and somatotroph cells in the pituitary. A role for imprinted genes in pituitary function is well known (Davies et al., 2008, Ivanova and Kelsey, 2011) with pituitary abnormalities associated with imprinted disorders such as PWS (Miller et al., 2008) and recent sequencing work showing imprinted genes are amongst the highest expressed transcripts in the mature and developing pituitary (Scagliotti et al., 2021). Specific genes I found highly expressed here, such as *Dlk1* and *Nnat*, have been shown to alter somatotroph phenotypes (Charalambous et al., 2014, Huerta-ocampo et al., 2004). Outside the

hypothalamus, the midbrain dopamine neurons were enriched when contrasted to other dopamine neurons from the brain and the enriched serotonergic neurons were those that project to the subcortical regions of the brain known to be associated with feeding and other motivated behaviours (Donovan and Tecott, 2013), providing convergence with the functional hotspots seen in the hypothalamus.

Several sites of enrichment identified in this chapter are relevant to parenting behaviour. This includes the enrichments discussed above, in the subcortical projecting serotonin system which projects to the hypothalamus and the POA, the midbrain dopamine system which directly regulate parental motivation and the enrichment in the regulation of prolactin which acts on the POA to prime mothers during pregnancy and fathers to perform parenting behaviour (Brown et al., 2017, Larsen and Grattan, 2012, Smiley et al., 2022) . However, the most interesting site of enrichment I found was in the galanin neurons of the POA. These neurons are fundamentally important to parenting, ablation of these removes the ability of the animal to parent (Wu et al., 2014) and the optogenetic activation of these neurons is enough to produce this behaviour in animals which would otherwise not produce parenting behaviour (Kohl et al., 2018). The demonstration of an imprinted gene enrichment in galanin neurons at multiple levels of analysis, as well as in the specific population in the POA, is a powerful indicator that direct regulation of parenting is a function upon which imprinted genes may converge. If a gene were to modulate the effects on serotonin, dopamine and prolactin, it will have functional effects beyond parenting, and these other functions may be the intentional target of the gene's action. This galanin neuronal population in the POA is specifically for the control of parenting behaviour. Therefore, it was the imprinted genes expressed in these neurons that I put forward as candidates for having a role in parenting behaviour and which were my focus for the remainder of this thesis.

Figure 3.10 displays these imprinted gene candidates based on expression in multiple datasets or high expression in the POA dataset specifically (since these neurons are confirmed to be the parenting related neurons while the others could still include other types of galanin neurons in their populations). 21 imprinted genes demonstrated this form of expression in the POA^{Gal} population which works out to 1/6 of the imprinted genes in the list and 1/4 of the imprinted genes expressed in at least 20 cells of the POA. This expression hotspot is therefore not comprised of a few highly expressed genes but a substantial amount of the list. Importantly, the imprinted genes previously associated with parenting behaviour – *Peg3, Mest* & *Dio3* are all represented in this list (Peg13 was not sequenced). The mechanisms by

which these genes impact parenting behaviour has yet to be definitively established and so a role in POA^{Gal} could be an unappreciated mechanism for these genes.

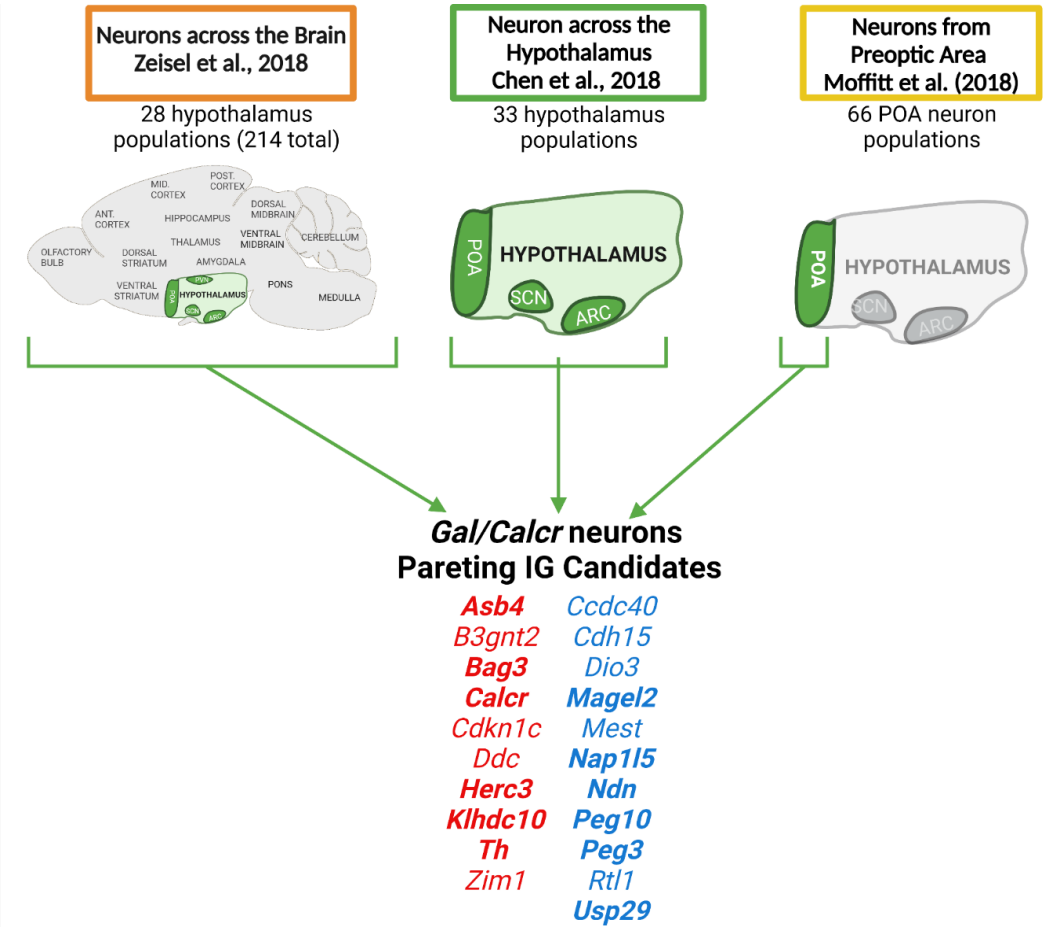


Figure 3.10. Imprinted gene parenting candidates based on gene expression in the POA^{Gal} neurons across different single-cell datasets. Imprinted genes are listed which show significant upregulation ($q \leq 0.05$ and Log2FC > 0) in the cell types. Genes in bold were enriched in galanin neurons in all 3 datasets (Moffitt et al., 2018, Chen et al., 2017, Zeisel et al., 2018). While the remaining genes either showed enrichment in 2 datasets or showed strong enrichment in the POA dataset (Moffitt et al., 2018). Parental-bias is indicated by colour (MEG - red, PEG - blue).

Since no imprinted gene has been associated with this neuronal population before, nor with galanin phenotypes, I thought it necessary to test the validity and utility of this bioinformatic approach to identity functional roles for genes based on expression patterns. When looking at this list of genes, I wanted to select one of the genes which had not been found to have a parenting deficit before and hence would aim to show that this gene would have implications for parenting behaviour when disrupted and would also display some form of phenotype in the POA/ POA^{Gal} neurons. Several genes (e.g., *Asb4, Bag3, Ndn, Nap1l5, Usp29)* showed significant enrichment in the POA neurons and galanin neurons in the wider contexts, however the gene I selected was *Magel2*. *Magel2* was one of the top enriched genes in

galanin neurons at each level of analysis and has never found to have impacts on parenting behaviour prior to the work of this thesis. Furthermore, *Magel2* presented a unique opportunity since this is an extensively well-studied gene and yet has not been shown to impact parenting, and so I believed that exposing a deficit in this gene would be more impactful to the validity of my approach as opposed to uncovering a new phenotype in a relatively un-studied gene such as the ones listed above. *Ndn* could be considered in the same situation as *Magel2* but the availability of a sophisticated mouse model (using Crispr-Cas9) of *Magel2* led me to focus on this gene for the rest of my thesis. The following chapter details step one in a three-step program that I would suggest any further parental characterisations of these imprinted genes goes through. Step 1 (Ch.4) – Validate target gene enrichment in target neuronal population, Step 2 (Ch.5) – Confirm suspected behavioural deficit in target gene mouse model, Step 3 (Ch.6) – Investigate target neuronal population phenotype/activity within target gene mouse model that demonstrates behavioural deficit.

3.4.1 Caveats

Analyses of these kind are always bound by the available data and therefore there are notable limitations and caveats to this chapter. The aim of this chapter was to generate information about 'hotspots' of imprinted gene expression. This approach, and the use of over-representation analysis and GSEA, therefore do not provide an exhaustive list of sites of expression, and non-differentially expressed genes could still be highly expressed genes but would not contribute to this analysis. There may also be functional effects occurring below the level of over-representation, or sites in which imprinted genes act during development and are not functionally enriched in adult cell types, or simply that compared to other cell types, some sites of expression are not 'hotspots' of imprinted expression. Specific sequencing of brain regions and cell types will be required to distinguish between these possibilities.

A second caveat relates to the nature of the clustering of cells. Datasets at larger scales inevitably produce more heterogenous clusters of cells as populations. There are many types of galanin positive neurons in the hypothalamus yet only two were resolved in Chen et al. (2017). Which one truly contains the POA^{Gal} cells is hard to know. Priority was given to the datasets with the smallest scales when identifying imprinted genes of interest. A third caveat is that, due to the nature of the datasets used, not all imprinted genes were included, and my analysis was missing a significant subset of imprinted genes encoding small RNAs or

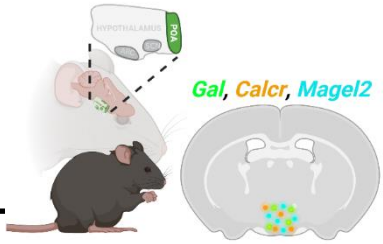
isoforms from the same transcription unit. Although it is worth noting that many non-coding RNA's (ncRNA) are exclusively expressed in the brain (Cavaillé et al., 2000, Andergassen et al., 2017) so would likely serve to strengthen the brain signal.

A final caveat is that I did not assess parent-of-origin expression for the 119 imprinted genes I included in the analysis. Previous expression profiling of imprinted genes has also not measured the POEs (Steinhoff et al., 2009, Gregg et al., 2010) but have restricted their gene selection to genes with reliable imprinting status. Consequently, I only included the canonical imprinted genes and genes with more than one demonstration of a POE when looking for enrichment. Furthermore, for the vast majority of these genes, a brain-based POE effect has also already been reported (Appendix Table A3.1). Although this does not replace validating the imprinting status of all 119 in the tissues and subregions examined, it does provide justification for looking at imprinted gene over-representation. To resolve this issue, scRNA-seq using tissues derived from reciprocal F1 crosses between distinct mouse lines will be key; for example, Laukoter et al. (2020) provide an example of the allelic specific single-cell expression measurements necessary to confirm the enrichments found in this study.

3.4.2 Summary of Findings

- Broadly, IG expression was shown to be enriched in neurons of the adult mouse brain, namely neurons from the hypothalamus, midbrain and hindbrain. Specifically, IG expression was shown to be convergently enriched in specific hypothalamic neurons of the POA, ARC and SCN as well as the monoaminergic system of the brain.
- Relating to parenting behaviour, imprinted genes show enrichment in the prolactin, serotonin and dopamine systems, and more fundamentally, the galanin neurons of the POA which directly govern parenting behaviour.
- 21 imprinted gene candidates have been put forward for having a potential relationship with parenting behaviour based on their expression profile in POA^{Gal}, 3 of which have already shown parenting deficits in mouse models.

4 *Magel2* Preoptic Area Expression



4.1 Overview

A gene that demonstrated enrichment in the POA^{Gal} hub might be expected to have an effect on the behaviour linked to that hub, specifically, parenting behaviour. As described in Chapter 3, *Magel2* was one of several imprinted genes identified as a candidate for an in-depth assessment into the role it plays in parenting behaviour.

Magel2 is known to have high expression in the hypothalamus and loss of expression has a resulting impact on hypothalamic mediated behaviours such as circadian regulation (Kozlov et al., 2007, Mercer et al., 2009) and suckling (Schaller et al., 2010). It has also been associated with oxytocin phenotypes in pups (Schaller et al., 2010, Reichova et al., 2021, Meziane et al., 2015, Da Prato et al., 2022) and in mothers (Ates et al., 2019), suggesting that *Magel2* is important for oxytocin functioning (Fountain and Schaaf, 2015). *Magel2*'s corresponding expression in the oxytocin neuron rich PVN and circadian regulating SCN of the hypothalamus is well known (Kozlov et al., 2007, Mercer et al., 2009). Chapter 3 highlighted that, in scRNA-seq data, *Magel2* demonstrates enriched expression in the *Avp/Nms* neurons of the SCN, the *Avp* neurons of the PVN, and, more surprisingly, the galanin neurons within the POA. However, to date, *Magel2* has not previously been associated with parenting deficits, nor has it been associated with galanin expressing neurons, which warranted a confirmatory investigation of *Magel2*'s expression in the POA^{Gal} hub.

The scRNA-seq analysis from Chapter 3 (Moffitt et al., 2018) found that *Magel2* specifically showed enriched expression in the i16:*Gal/Th* neurons of the POA. These neurons showed high expression of three marker genes - *Gal*, *Th* and *Calcr* and the Fluorescent In-situ Hybridization (FISH) work carried out by Moffitt et al. (2018) showed that two distinct populations of neurons made up this single cell population: the original *Gal/Th* descriptor, as well as a population of *Gal/Calcr* neurons. When assessed for immediate early gene activation following parenting behaviour, both these neuron groups significantly expressed *c-Fos*, but the response of *Gal/Calcr* neurons dwarfed the response of *Gal/Th* neurons, and this

was the case for mothers, fathers and virgin female mice. This distinguished *Gal/Calcr* as the principal markers of the POA^{Gal} hub.

The scRNA-seq analyses in Chapter 3 alone justify investigating the parenting behaviour of *Magel2*-null mice. However, this chapter aimed to definitively demonstrate an *in-situ* enrichment of *Magel2* not only in the POA, but in the POA^{Gal} hub. This would confirm *Magel2*'s candidacy as a gene potentially important for parenting behaviour as well as independently confirming the scRNA-seq expression profile. To do this, the spatial genomics technology RNAscope® was used to fluorescently label and visualize *Magel2* RNA molecules in the POA of the mouse brain alongside other gene targets (*Gal/Calcr* or *Gal/Th*) in the same sections of brain. In addition to confirming whether *Magel2* expresses meaningfully in the POA^{Gal} hub, and whether it is expressed differently in *Gal/Calcr* or *Gal/Th* cells, I also aimed to provide new information concerning *Magel2*'s expression in the murine brain by comparing this POA expression of *Magel2* to its expression alongside *Oxt* and *Avp* genes in neurons of the PVN, SCN and SON and the rest of the rostral hypothalamus. This allowed me to investigate the co-expression seen with *Magel2* and *Avp* in my scRNA-seq analysis as well as the literature-based association between *Magel2* and *Oxt* neurons.

4.2 Methods

4.2.1 FFPE Tissue Preparation

Wildtype (WT) mice (age 7-8 weeks, 3M & 3F) on a C57BL/6J background were transcradially perfused from the home cage with 10% NBF before whole brains were taken. 3 mm POA sections were embedded in paraffin and sectioned at a thickness of 10µm. Only sections containing the POA or containing the SCN/PVN were subsequently analysed. Sections from the brains were devoted towards three separate experiments. For the POA, sections from two brains (1M,1F) were assessed for *Gal/Th* neurons while sections from the remaining four brains (2M,2F) were assessed for *Gal/Calcr* neurons. For the SCN/PVN, sections post-POA from four brains (2M,2F) were assessed for *Oxt* and *Avp* positive neurons.

4.2.2 RNAscope® protocol

Three-plex RNA Scope was performed using RNAscope® Multiplex Fluorescent Reagent Kit v2 (ACD Bio-techne) on these FFPE brain sections. The manufacturer's protocol was

followed exactly following the 'standard' pre-treatment guidance (see Section 2.3.4). POA sections received a *Gal/Th/Magel2* probe mix, or a *Calcr/Gal/Magel2* and PVN/SCN sections received a *Oxt/Avp/Magel2* probe mix. The *Gal* Probe (Mm-Gal), the *Calcr* Probe (Mm-Calcr) and the *Oxt* probe (Mm-*Oxt*) were paired with fluorophore TSA Vivid™ 520. The *Th* Probe (Mm-Th-C2), the C2 *Gal* Probe (Mm-Gal-C2) and the *Avp* probe (Mm-*Avp*-C2) were paired with TSA Vivid™ 570. The *Magel2* Probe (Mm-*Magel2*-C3) was always assigned to TSA Vivid™ 650. All fluorophores were applied at a concentration of 1:1500. Slides were counterstained with DAPI (30 seconds) and mounted.

4.2.3 Image Acquisition and Analysis

Whole brain slides were imaged within one week of mounting using the Zeiss AxioScan Z1 at 20x magnification with the same light intensity/duration settings used for the same probes in each scan (see Section 2.3.5). Images were analysed with Zen Blue 3.6. Slides were pre-processed and thresholded as detailed in section 2.3.6 and Figure 2.5. For genes that displayed dots and clusters (*Gal, Th, Calcr, Magel2*), the clusters were resolved into molecule counts. For *Oxt* and *Avp* displaying an intense form of expression, percentage coverage metrics were applied at 20% as detailed in section 2.3.6. The finished data took the form of individual molecule counts for each of the channels for each DAPI identified cell.

4.2.4 Statistics and Figures

RNAscope® image data were analysed in two ways. Firstly, my quantitative analysis compared proportions of *Magel2* positive cells between areas of the brain and cell types using two-way Fisher's Exact Test with Bonferroni adjustment. I also compared molecule counts for *Magel2* between *Gal/Th* & *Gal/Calcr* vs. the rest of the preoptic area cells or within *Oxt* vs. *Avp* vs. the rest of cells using either Wilcoxon Ranked Sums Tests with Bonferroni correction or Kruskal Wallis Tests with post hoc Bonferroni corrected Dunn tests for multiple comparisons. Secondly, a semi-quantitative metric was also calculated by deriving a H-Score to compare distributions of cells expressing binned quantities of *Magel2* molecules (see Section 2.3.8).

4.3 Results

4.3.1 Magel2 is predominantly expressed in the hypothalamus and septal area and shows highest expression in the POA and SCN within the hypothalamus

Table 4.1 displays the *Magel2* expression data synthesized from the bioinformatic work from Chapter 3 of this thesis. A clear enrichment in the hypothalamus was found, as well as representation from neurons of the SCN expressing *Avp/Vip* and neurons of the PVN expressing *Avp/Oxt*. However, the neuron type with the strongest fold change difference for *Magel2* was the galanin expressing neurons of the hypothalamus, and based on the marker gene expression, this population most likely represented galanin neurons from the POA.

Table 4.1. *Magel2* expression in the scRNA-seq analysis performed in Chapter 3. *Magel2*'s expression in the Zeisel et al. (2018) and Chen et al. (2017) single cell datasets examining neurons across the whole mouse brain and whole mouse hypothalamus respectively is shown. *p*, *q*, FC values are reported for *Magel2*'s expression in a particular cell identity group vs. background as well as the percentage of cells in which *Magel2* is expressed in the identity group as well as in the background (Rest). *Magel2*'s expression is shown for neurons across different regions of the brain, across different neurons of the brain and across different neurons of the hypothalamus specific. POA^{Gal} related identity groups are boldened.

	Neurons with significantly higher expression of *Magel2*	*p*	*q*	FC	Proportion in	Proportion rest
Neurons across different brain regions (Zeisel et al., 2018)	**Hypothalamus**	**8E-303**	**2E-301**	**9.39**	**0.112**	**0.013**
	Medulla	6E-119	7E-118	5.48	0.094	0.015
	Pons	2E-72	1E-71	3.28	0.082	0.015
	Ventral Midbrain	2E-32	1E-31	3.05	0.055	0.016
	Spinal Cord	2E-16	7E-16	2.90	0.040	0.016
Neuron populations across the nervous system (Zeisel et al., 2018)	**TEINH3 - *Gal* - POA/BNST**	**4E-96**	**4E-94**	**17.47**	**0.234**	**0.016**
	HYPEP5 - *Avp/Oxt*, PVN	4E-83	2E-81	13.36	0.202	0.016
	HBSER4 - *5HT*, Raphe nuclei	7E-61	2E-59	12.32	0.264	0.016
	DEINH8 - *Nms/Vip* - SCN	1E-21	2E-20	12.12	0.125	0.016
	HBADR – *Adr* - Medulla	1E-15	1E-14	11.47	0.175	0.016
Neuron populations across the hypothalamus (Chen et al., 2017)	**GABA10 - *Brs3/Calcr/Gal***	**0.001111**	**0.018893**	**2.43**	**0.643**	**0.256**
	GABA8 - *Avp/Vip* - SCN	0.000002	0.000065	1.79	0.391	0.246
	GABA9 - *Vipr2/Lhx1* - SCN	0.002363	0.026780	2.05	0.478	0.256

In agreement with the literature and scRNA-seq data, RNAscope® analysis indicates that *Magel2* is only expressed strongly in the hypothalamus/POA and areas of the septum (Figure

4.1A). Images taken more caudally in order to capture the PVN and SCN, also show that the hypothalamus is the dominant locale for *Magel2* expression in the mouse brain (Figure 4.1B).

Magel2 was assessed using the same probe (Mm-*Magel2*-C3) and fluorophore (TSA Vivid 650) combination with the same light intensity (100%) and duration (150 ms) used for all the fluorescent images across all sections imaged in this chapter. This consistency in treatment of the *Magel2* probe allowed me to quantitatively compare the expression levels of *Magel2* between different regions from the rostral half of the hypothalamus sectioned in this chapter.

Table 4.2 displays the percentage of cells that were positive for *Magel2* (i.e., with 1 or more molecules) in the different ROIs. The surrounding background hypothalamus had ~54% of cells with molecules of *Magel2*, with an average of 1.45 molecules per cell. The POA in comparison had slightly more positive cells (~58%) and significantly more *Magel2* molecules per cell (1.86 molecules, $p = 7.5 \times 10^{-171}$, Mann-Whitney *U*-test) than the background hypothalamic expression. This suggested that the POA was a site of elevated *Magel2* expression and expressed significantly more *Magel2* molecules than the PVN (1.71 molecules, $p = 3.62 \times 10^{-8}$, Mann-Whitney *U*-test) which also had fewer *Magel2* positive cells (54.6%) and the SON which displayed fewer *Magel2* positive cells (~37%) and significantly fewer average molecules than even the background hypothalamus (1.21 molecules, $p = 9.8 \times 10^{-61}$, Mann-Whitney *U*-test). However, the clear hotspot of *Magel2* expression in the hypothalamus was the SCN with ~73% cells positive for *Magel2* and an average expression almost 2 times higher than the surrounding hypothalamus (2.88 molecules) and was significantly greater than the POA ($p = 1.18 \times 10^{-271}$, Mann-Whitney *U*-test). Statistics for all comparisons can be found in Appendix Table A6.1.

Table 4.2. *Magel2* expression metrics in the regions of interest of the hypothalamus from RNAscope® sections. For each region of interest, all cells were assessed for *Magel2* expression, the percentage of cells within the ROI with a *Magel2* molecule (i.e., *Magel2* positive) are reported as well as the average number of *Magel2* molecules per cell.

Region of Interest	Percentage of cells *Magel2* positive	Average number of *Magel2* molecules per cell
SCN	72.80 %	2.88
POA	57.52 %	1.86
PVN	54.66 %	1.71
Surrounding Hypothalamus	53.73 %	1.45
SON	36.82 %	1.21

4.3.2 In the Preoptic Area, Magel2 displays 2-3 times higher expression rates in Gal/Calcr, and Gal/Th cells compared to other POA cells

The cells of the preoptic area of the hypothalamus expressed more molecules of *Magel2* than the general surrounding hypothalamic cells. However, I wanted to know whether this POA expression was driven by any cell type in particular, specifically the POA^{Gal} cells. Using probes for the marker genes *Gal/Th* and *Gal/Calcr*, I quantified *Magel2* probe expression rates in cells tagged with these marker genes.

I first examined *Gal/Th* expression alongside *Magel2* in the POA (Figure 4.2A & 4.2B). Two WT brains (1M/1F) were sectioned through the POA with one section every 100 µm used with gene probes. This resulted in 18 sections and 69,560 POA cells analysed for this experiment. 3.7% of POA cells were identified as *Gal/Th* positive (2579 cells), representing 51% of the galanin positive cells. *Magel2* was found to be expressed in 88.29% of *Gal/Th* cells which was significantly more than the POA standard (57.52%, $p < 2.2x10\text{-}16$, Fisher's Exact Test). I then examined *Gal/Calcr* expression alongside *Magel2* in the POA (Figure 4.3A & 4.3B). Four WT brains (2M/2F) were sectioned through the POA with one section every 100 µm used with gene probes. This resulted in 33 sections and 156,169 POA cells analysed for these experiments. 2.5% of POA were identified as *Gal/Calcr* positive (3846 cells), representing 44% of the galanin positive cells. *Magel2* was expressed in 92.23% of *Gal/Calcr* cells which is significantly more than the POA standard (57.52%, $p < 2.2x10\text{-}16$, Fisher's Exact Test).

Quantitative analysis of *Magel2* molecules in these galanin cell types found that significantly more *Magel2* molecules were present in *Gal/Th* cells (5.28 molecules) than all other cells in the POA (2.11 molecules, FC = 2.5, $p = 0$, Mann-Whitney *U*-test). To further test the specificity of *Magel2* expression in the POA, I compared *Gal/Th* cells to all other POA *Gal* positive cells, with the former having significantly more *Magel2* molecules (4.11 molecules, FC = 1.3, $p(adj) = 3.8x10^{-11}$, Dunn's Test). Similarly, there were more *Magel2* molecules in *Gal/Th* cells than in other POA *Th* positive cells (3.22 molecules, FC = 1.64, $p(adj) = 2.1x10^{-87,}$ Dunn's Test) (See Figure 4.2C). Restricting the analysis to only *Magel2* positive cells indicated that there were still significantly more *Magel2* molecules in *Gal//Th* cells (5.99 molecules) compared to all other cells expressing *Magel2* (3.45 molecules, FC = 1.74, $p = 4.99x10^{-171}$, Mann-Whitney *U*-test). An identical finding was made when analysing

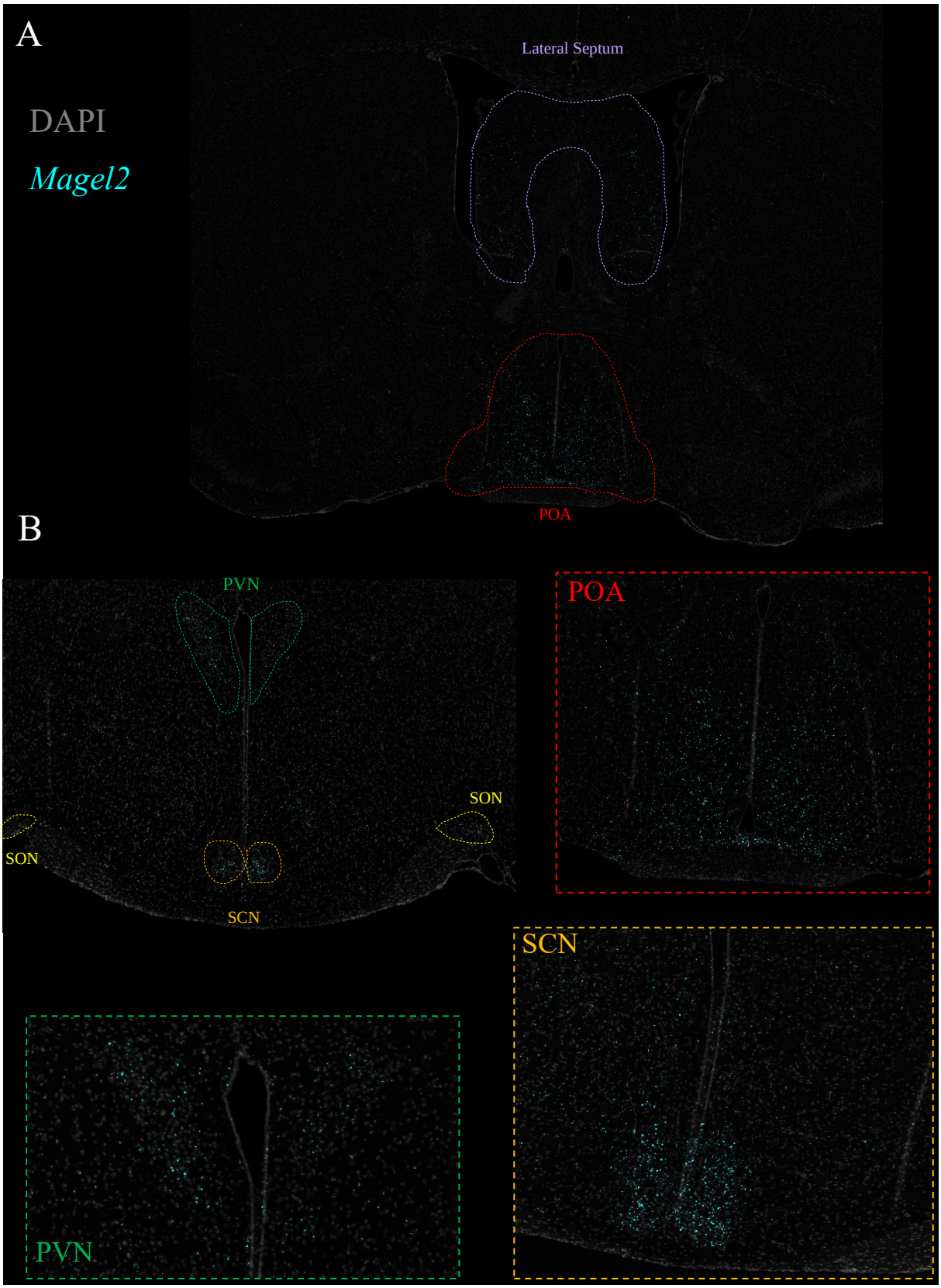
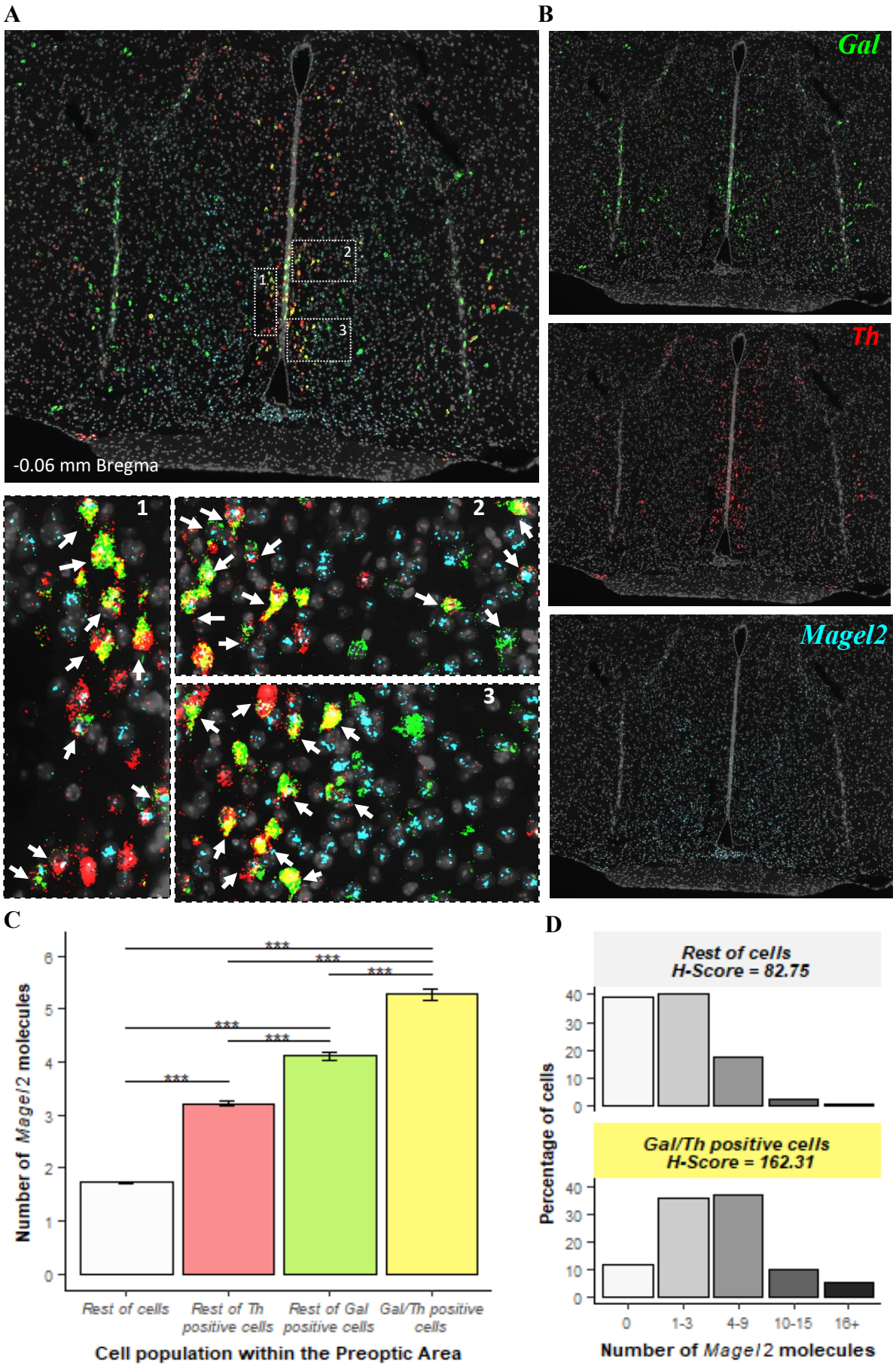


Figure 4.1. In situ expression of *Magel2* as visualised by RNAscope®. (A) Magel2 visualised in a whole brain slice with hotspots: POA and LS highlighted. Zoomed in view of the POA is highlighted in middle right. (B) Magel2 visualised in the central hypothalamus. PVN, SON and SCN are highlighted and a zoomed in view of the PVN is visible bottom left. Bottom right shows Magel2 in the SCN in a different central hypothalamic section with a larger part of the SCN.

Gal/Calcr cells. There were significantly more *Magel2* molecules in *Gal/Calcr* cells (5.81 molecules) than all other cells (1.59, FC = 3.66, $p = 0$, Mann-Whitney *U*-test), all other *Gal* cells (3.06, FC = 1.89, $p(adj) = 7.3x10^{-233}$, Dunn's Test), all other *Calcr* cells (3.3 molecules, FC = 1.76, $p(adj) = 2x10^{-80}$, Dunn's Test) (See Figure 4.3C) and when only using *Magel2* positive cells in the analysis (6.3 vs. 2.91 molecules, FC = 2.16, $p = 0$, Mann-Whitney *U*-test). Data for all comparisons can be found in Appendix Table A6.2. Semi-quantitative H-scores were also calculated for all the comparisons listed above and *Gal/Th* and *Gal/Calcr* consistently displayed higher H-scores in all comparisons and the highest H-score was observed in *Gal/Calcr* cells (Appendix Table A6.3). Figure 4.2D and 4.3D display histograms of these H-scores when comparing *Gal/Th* or *Gal/Calcr* cells to all of the rest of POA cells. These target POA cells have a much higher proportion of their cells with 4+ *Magel2* reads compared to the background *Magel2* expression and much fewer cells expressing zero *Magel2* molecules, both of which contributes to the H-scores being twice as high than the background ones. The reduction in cells with no *Magel2* molecules can be explained in part by the presence of non-neuronal cells in the background comparison which will not be included in cells expressing *Gal/Calcr* or *Gal/Th*. However, this does not fully explain the difference because when I removed cells in which *Magel2* is not expressed the H-scores for *Gal/Th* and *Gal/Calcr* were still much larger than the background values, indicating that these cells do express *Magel2* in larger quantities across a greater number of cells than the background POA. (Appendix Table A6.3)

Overall, these RNAscope® studies validated my findings from Chapter 3 by showing that *Magel2* is expressed significantly higher in *Gal/Th/Calcr* cells compared to other cell types in the POA region and is likely to be one of the hotspots of *Magel2* expression in the brain. I was underpowered to compare brains by sex, but values such as percentage of cells

Figure 4.2. (Next Page) In situ coexpression of *Magel2* in *Gal/Th* neurons. (A) *Top* Low magnification image of hypothalamic section after *in situ* amplification of *Gal* (green), *Th* (red), and *Magel2* (turquoise). *Bottom* High-resolution image of three open white dashed boxes numbered 1-3. Examples of co-expression of *Gal*, *Th* and *Magel2* in one cell are indicated with white arrows. (B) Low magnification image of hypothalamic section after *in situ* amplification of *Gal* (green), *Th* (red), and *Magel2* (turquoise) presented as single channel images in that order from top to bottom. (C) Number of *Magel2* RNA molecules detected in different cell types from all sections. *Gal/Th* cells expressed significantly more RNA molecules of *Magel2* than the other cell types, even including *Gal* expressing and *Th* expressing cells separately ($H(3) = 5313.6$, $p = 2.2x10\text{-}16$, ***$p<0.001$, post hoc Dunn test). (D) Histogram showing the percentage of cells with particular number of *Magel2* molecules in *Gal/Th* positive cells vs. all other cells. A larger percentage of *Gal/Th* cells expressed 4+ *Magel2* RNA molecules and a smaller percentage of cells expressed zero *Magel2* RNA molecules which contributed to the differences in H-Score.


A
B
Gal
Th
Magel2
-0.06 mm Bregma
1
2
3
C
D
Number of Magel2 molecules
Rest of cells
Rest of Th positive cells
Rest of Gal positive cells
Gal/Th positive cells
Cell population within the Preoptic Area
Rest of cells
H-Score = 82.75
Gal/Th positive cells
H-Score = 162.31
Percentage of cells
0
1-3
4-9
10-15
16+
Number of Magel2 molecules

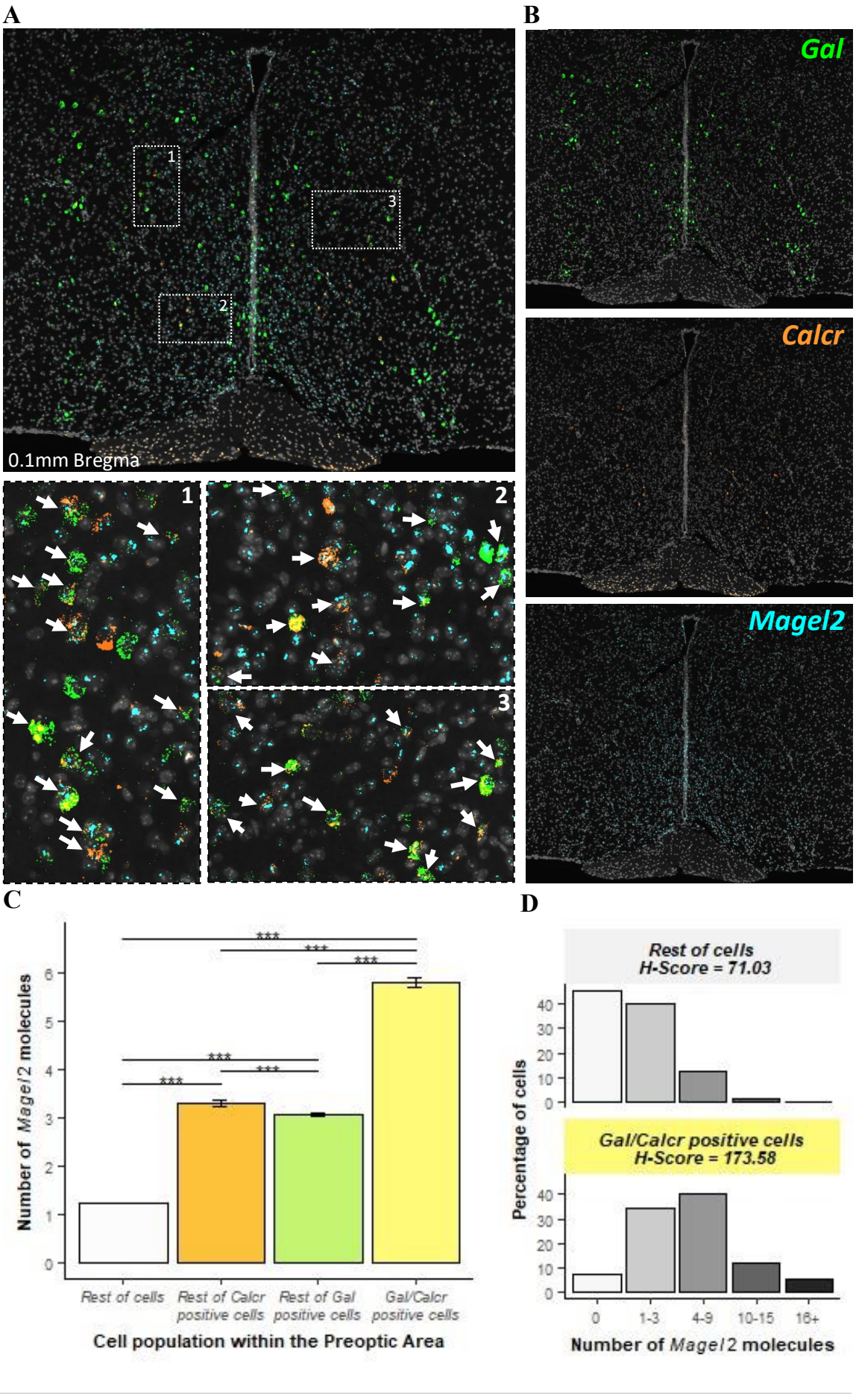

A
B
Gal
Calcr
Magel2
1
2
3
0.1mm Bregma
C
Number of Magel2 molecules
Rest of cells
Rest of Calcr positive cells
Rest of Gal positive cells
Gal/Calcr positive cells
Cell population within the Preoptic Area
***
D
Rest of cells
H-Score = 71.03
Gal/Calcr positive cells
H-Score = 173.58
Percentage of cells
0
1-3
4-9
10-15
16+
Number of Magel2 molecules

Figure 4.3. (Previous Page) In situ coexpression of *Magel2* in *Gal/Calcr* neurons. (A) *Top* Low magnification image of hypothalamic section after *in situ* amplification of *Gal* (green), *Calcr* (orange), and *Magel2* (turquoise). *Bottom* High-resolution image of the three open white dashed boxes numbered 1-3 from the top image. Examples of co-expression of *Gal*, *Calcr* and *Magel2* in one cell are indicated with white arrows. (B) Low magnification image of hypothalamic section after *in situ* amplification of *Gal* (green), *Calcr* (orange), and *Magel2* (turquoise) presented as single channel images in that order from top to bottom. (C) Number of *Magel2* RNA molecules detected in different cell types from all sections. *Gal/Calcr* cells expressed significantly more RNA molecules of *Magel2* than the other cell types, even including *Gal* expressing and Calcr expressing cells separately (H(3) = 17152, p = 2.2x10-16, ***p<0.001, post hoc Dunn test). (D) Histogram showing the percentage of cells with particular number of *Magel2* molecules in *Gal/Th* positive cells vs. all other cells. A larger percentage of *Gal/Calcr* cells expressed 4+ *Magel2* RNA molecules and a smaller percentage of cells expressed zero *Magel2* RNA molecules which contributed to the differences in H-Score.

expressing *Magel2*, and average molecule counts were consistent from the two sexes suggesting no substantial differences between males and females.

4.3.3 Magel2 is enriched in Avp cells of the PVN but not Oxt cells

The final analyses in this chapter centred around the *Oxt* and *Avp* probes used alongside *Magel2* in sections of the brain containing the PVN. The PVN showed higher *Magel2* expression than background hypothalamus but not as high as the POA and SCN. Previous reports of *Magel2* expressing in and influencing oxytocin neurons, as well as the scRNA analysis showing *Magel2* highly co-expressing with vasopressin, suggested that the PVN (as a major source of oxytocin and vasopressin neurons) would be an interesting place to examine *Magel2* expression. The oxytocin and vasopressin neurons also regulate parental motivation so could be additional targets via which *Magel2* could affect behaviour.

Figure 4.4A displays exemplar images for the *Oxt/Avp/Magel2* RNAscope® experiment. A total of 115,081 hypothalamic cells were identified across 33 sections from 4 mouse brains (2M, 2F). For the PVN (Figure 4.4B), there were 9780 cells identified with 9.4% of cells *Oxt* positive and 7% of cells *Avp* positive. 69% of oxytocin positive cells were *Magel2* positive compared to 84% of vasopressin positive cells which was significantly greater than the background PVN rate of 54.7% of cells ($p(adj)$ = $6.98x10^{-27}$ & $4.35x10^{-63}$, Fisher's Exact Test). *Avp* cells were significantly more positive than the *Oxt* cells (p(adj) = $5{,}95x10^{-11}$, Fisher's Exact Test). While *Magel2* had higher rates of expression in oxytocin (FC = 1.26, $p(adj)$=$2.6x10^{-17}$, Dunn's Test) and vasopressin neurons (FC = 2.82, $p(adj)$ = $4.5x10^{-106}$, Dunn's Test) compared to the rest of the PVN (Figure 4.4C), oxytocin cells only had an

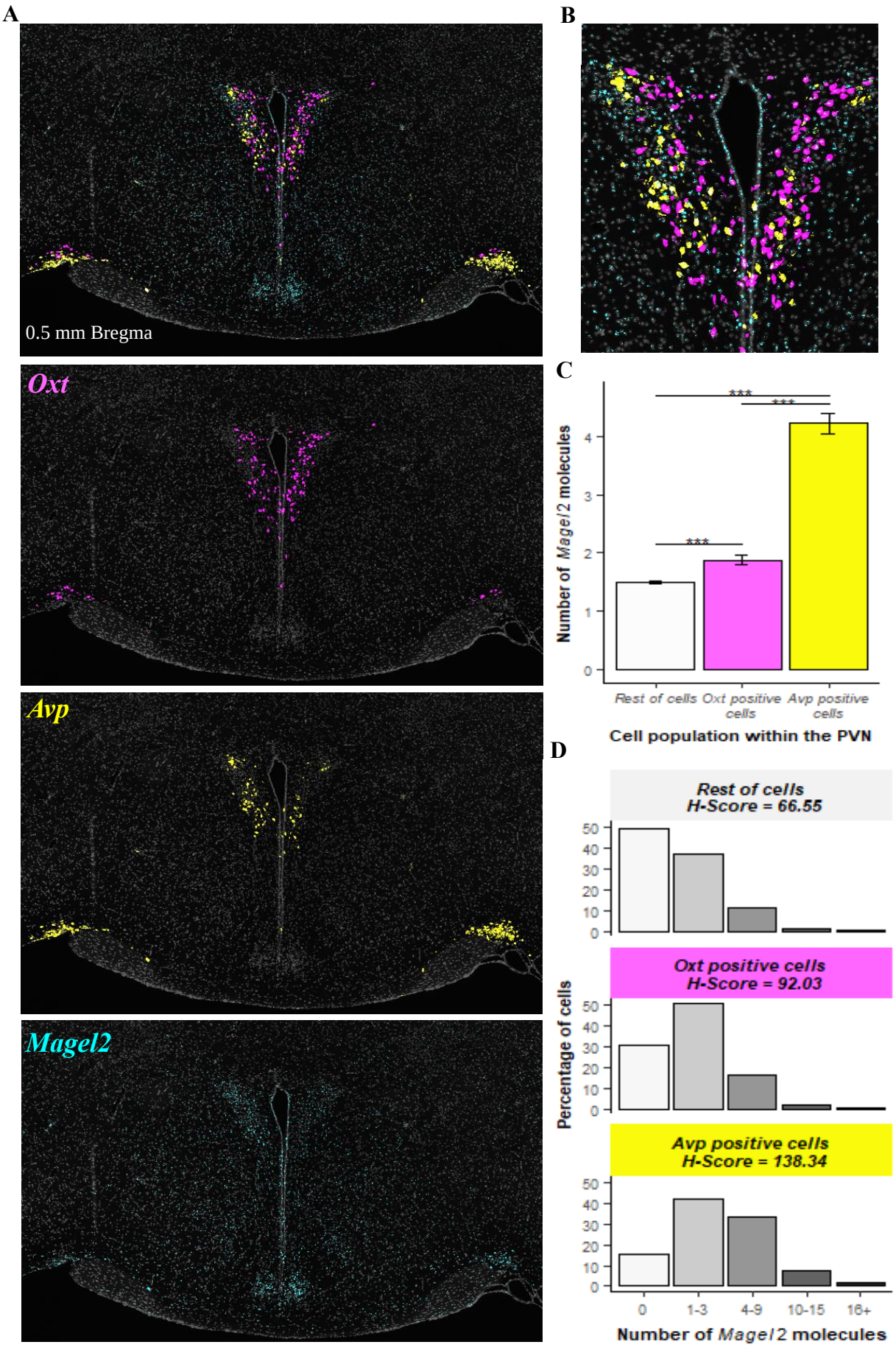

A
0.5 mm Bregma
Oxt
Avp
Magel2
B
C
***
***
***
Number of Magel2 molecules
0
1
2
3
4
Rest of cells
Oxt positive cells
Avp positive cells
Cell population within the PVN
D
Rest of cells
H-Score = 66.55
Oxt positive cells
H-Score = 92.03
Avp positive cells
H-Score = 138.34
Percentage of cells
0
10
20
30
40
50
0
1-3
4-9
10-15
16+
Number of Magel2 molecules

Figure 4.4. (Previous Page). In situ coexpression of *Magel2* in *Oxt* and *Avp* PVN neurons. (A) *Top* Low magnification image of hypothalamic section after *in situ* amplification of *Oxt* (pink), *Avp* (yellow), and *Magel2* (turquoise). *Bottom* Low magnification image of hypothalamic section presented as single channel images for *Oxt*, *Avp* and *Magel2*, from top to bottom. (B) Higher resolution image of just the PVN from the hypothalamic section (C) Number of *Magel2* RNA molecules detected in different cell types from all sections. *Avp* cells expressed significantly more RNA molecules of *Magel2* than the other cell types, yet *Oxt* cells still expressed more *Magel2* than the background ($H(2) = 499.46$, $p = 3.5 \times 10^{-109}$, ***$p < 0.001$, post hoc Dunn test). (D) Histogram showing the percentage of cells with particular number of *Magel2* molecules in *Avp* positive cells vs. *Oxt* positive cells vs. all other PVN cells. A larger percentage of *Avp* cells expressed 4+ *Magel2* RNA molecules and a smaller percentage of cells expressed zero *Magel2* RNA molecules which contributed to the differences in H-Score.

average of 1.88 *Magel2* molecules per cell whereas vasopressin positive cells had significantly more (4.22) molecules on average (FC = 2.25, $p(adj) = 4.8 \times 10^{-30}$, Dunn's Test).

As above, when using only *Magel2*-positive cells in the analysis curiously, I find that *Avp* neurons still express *Magel2* significantly more than *Oxt* (FC = 1.80, $p(adj) = 1.1 \times 10^{-24}$, Dunn's Test) and other PVN neurons (FC = 1.69, $p(adj) = 7.7 \times 10^{-38}$, Dunn's Test), but *Oxt* neurons no longer express more *Magel2* molecules than the rest of the PVN (FC = 0.93, $p(adj) = 1$, Dunn's Test). This suggest that *Magel2* is not highly expressed in mature *Oxt* neurons, it is just expressed in more of them than the background, something that is reinforced when looking at the H-scores and histograms (Figure 4.4D, Appendix Table A6.3) which show that while oxytocin neurons do display higher H-scores than the PVN background cells, most oxytocin cells have 0, 1, 2, or 3 molecules of *Magel2* which is a similar split to the background cells. *Avp* cells on the other hand have a higher number of cells with 4+ *Magel2* molecules. All this suggests that *Avp* cells, rather than *Oxt* cells, are the hotspot of expression for *Magel2* in the adult PVN.

4.4 Discussion

This chapter aimed to explore *Magel2* expression by quantifying RNA molecules in cells using RNAscope® in sections of the mouse brain containing the rostral half of hypothalamus. This chapter was principally a validation exercise for the results that the transcriptomics of Chapter 3 revealed about *Magel2* and hence the limited number of replicates did not fundamentally weaken the following conclusions. I found that *Magel2* was expressed predominantly in the hypothalamus and the septal areas of the brain. Within the hypothalamus the suprachiasmatic nucleus (SCN) was the clear hotspot of expression,

followed by the preoptic area (POA) and then the paraventricular nucleus (PVN), with all expressing more *Magel2* than the general hypothalamus. Interestingly, little is reported in the literature on *Magel2*'s expression in the POA whereas there are many reports of *Magel2* high expression in the PVN, yet *Magel2* was expressed higher in the former. When looking within the POA, I found that *Gal/Th* and *Gal/Calcr* cells, both markers for neurons important in parenting, expressed 2-3 times more *Magel2* than the background POA expression, strongly suggesting that the POA^{Gal} hub is a hotspot for *Magel2* expression within the mouse brain. Additionally, *Magel2* was found to be highly expressed in vasopressin cells of the PVN, much more than the oxytocin cells. Overall, SCN cells, *Avp* expressing PVN cells and *Gal/Calcr* cells of the POA show elevated expression of *Magel2*. These are the same sites of enrichment suggested by the single-cell sequencing data in Chapter 3, and these sites suggest a role for *Magel2* in circadian regulation as well as parenting and social behaviour.

The high expression of *Magel2* in the hypothalamus and the septal area is well known (Mercer et al., 2009, Kozlov et al., 2007, Tacer and Potts, 2017, Lee et al., 2000). The septal area (medial olfactory area) plays a role in reward and reinforcement and is seen as having top-down control of motivated behaviour such as social interactions (Kolb and Nonneman, 1974), parenting behaviour (Stack et al., 2002, Carlson and Thomas, 1968) and sexual behaviour (Wirtshafter and Wilson, 2021). The expression of *Magel2* here highlights the functionality of this area as another inter-related mechanism by which *Magel2* can impact fundamental social behaviours such as parenting, particularly as the lateral septum is one of the major inputs and outputs of the POA^{Gal} hub (Kohl et al., 2018) and oxytocin and vasopressin activity in this area has been shown to impact social behaviour in *Magel2*-null mice (Borie et al., 2020).

When looking at hypothalamic regions, the clear hotspot of *Magel2* expression in the hypothalamus was the SCN. *Magel2*'s expression here is already known and, linking expression to function, *Magel2* has been associated with circadian abnormalities such as impacts on circadian regulated movement (Kozlov et al., 2007). Also of interest is that Kozlov et al. (2007) found *Magel2* to be highly expressed in vasopressin positive cells of the SCN, the same cell type found to be enriched for imprinted gene expression in Chapter 3. When examining cell-type specific PVN populations, *Magel2* did display higher expression in oxytocin neurons (slightly) and in vasopressin neurons of the PVN (more dramatically); yet the expression in the SCN and POA were more pronounced. Oxytocin phenotypes have been noted in mice such as 36% reduction in oxytocin levels in the P0 hypothalamus

(Schaller et al., 2010) but clearly this is not replicated by *Magel2* expression in oxytocin neurons in the mature brain seen here and indeed Kozlov et al. (2007) also did not see oxytocin expression differences in *Magel2* mutant adult mice. Instead, vasopressin appears to be a more prominent site of *Magel2* enrichment and perhaps more important when it comes to the functional impacts of *Magel2* in the adult brain. Vasopressin neurons of the PVN are known for their role in social behaviour and motivation as well as being linked to regulation of parenting behaviour, particularly in male mice (Bendesky et al., 2017, Namba et al., 2016, Stohn et al., 2018, Kessler et al., 2011, Caldwell, 2017, Stevenson and Caldwell, 2012). Interestingly Kohl et al. (2018) demonstrated that the POA^{Gal} hub received heavy inputs from the PVN and SON *Avp* neurons yet not the *Oxt* neurons, further tying *Magel2* expression back to the POA^{Gal} hub.

Finally, most interestingly for this thesis, I found an enrichment of *Magel2* in the POA which has seemingly gone under-appreciated in the literature. Additionally, two particular cell types in the POA showed remarkable enrichment - the *Gal/Th* neurons and an even stronger enrichment in the *Gal/Calcr* neurons. The multiplexed error-robust FISH (MERFISH) data from Moffitt et al. (2018) showed that POA neurons expressing *Gal* and *Calcr* had the highest *c-Fos* response following parenting behaviour in mothers, fathers and virgin females. These cells effectively constitute the POA^{Gal} hub in mice. In Chapter 3 I found that imprinted genes demonstrated enrichment in this cell type compared to other POA cells and *Magel2* was one of these genes. Here, I have found that *Gal/Calcr* neurons of the POA display the strongest signal for *Magel2* out of all other cell types examined here, with 92% of cells expressing *Magel2* and the average *Gal/Calcr* cells contained ~ 6 *Magel2* molecules compared to a background expression in the preoptic area of ~1.5 molecules. Considering this exceeds *Magel2*'s expression in other known important neuron types such as the SCN cells and the *Oxt/Avp* cells, it strongly suggests the galanin expressing cells of the POA are an enrichment site for *Magel2* and since the main behavioural output of these particular galanin neurons is parental motivation, it is strongly suggestive that disrupting *Magel2* expression will have implications for parenting behaviour, which is explored in detail in Chapter 5.

4.4.1 Caveats

The primary caveat to this chapter is the limited scope of my analysis. RNAscope requires many resources and so the scope of the experiment must be carefully curtailed. All experiments were underpowered to explore sex differences. The *Gal/Th/Magel2* experiment only utilised two animals which is useful for the cell type specific analysis since I had 1000's

of cells for each of the experiments but one animal per group was not enough to compare male and female differences and to work as a validation study for the transcriptomics.

Another drawback was that I have tried to draw insights about *Magel2*'s expression in the brain and yet I only analysed sections from a 3 mm section from the centre of the brain. This provides insight on the hypothalamus but even then, this does not include the whole hypothalamus and so the hotspots I have identified could yet be dwarfed by something more caudal. *Magel2* was suggested to be expressed in serotonin and midbrain dopamine cells also which were outside of the 3 mm range of the sections.

Finally, the conclusions about *Magel2*'s expression in *Gal/Calcr*, *Gal/Th* and *Oxt* or *Avp* cells would be strengthened by using different probes and examining *Magel2* expression in many other cell types. The sections I have used will have contained many non-neuronal cells and so comparisons of *Magel2* expression in the *Gal/Calcr* cells vs. the background cells of the POA may be an unfair comparison as the former will be predominantly neurons and the latter will contain many non-neuronal cells. However, I did try to address this, and I showed that *Magel2* is expressed 2-fold higher in *Gal/Calcr* cells even when compared to *Gal* expressing cells and *Calcr* expressing cell which should also contain predominantly neuronal cells. Having other cell types would confirm the enrichment in POA^{Gal} is genuine and not overly enhanced by comparing neurons to non-neurons

4.4.2 Summary of Findings

- *Magel2* was shown to be highly expressed in the rostral hypothalamus and lateral septum and hotspots of *Magel2* expression in the rostral hypothalamus included the cells of the suprachiasmatic nucleus, paraventricular nucleus and the cells of the preoptic area.
- Using probes in the PVN, *Magel2* was shown to be expressed highly in the *Avp* positive neurons and moderately in *Oxt* positive neurons.
- The strongest signal for *Magel2* was in the POA^{Gal} hub, specifically *Magel2* showed its highest expression and highest number of positive cells in the *Gal/Th* and particularly the *Gal/Calcr* cells of the POA.

5 *Magel2*-null Parenting Assessment

5.1 Overview

In Chapters 3 and 4, *Magel2*, an imprinted gene not previously shown to directly affect parenting behaviour, was shown to have elevated expression in the POA^{Gal} neurons in mice. This was seen when looking at *Magel2*'s expression in single-cell clustered populations from the POA (Ch.3) as well as a localized, enriched expression of *Magel2*'s in the *Gal/Calcr* neurons (Ch.4), found to be essential to parenting in mothers, fathers and virgin female mice by Moffitt et al. (2018). The aim of this chapter was to carry out an assessment of parenting behaviour in *Magel2*-null mice, to confirm, based on molecular findings from Chapters 3 and 4, whether the expression of *Magel2* would be predictive of a parenting deficit, a behaviour which as yet, is unexplored in this genetic model.

Parenting is a suite of behaviours in mice, comprising explicit actions such as pup-retrieval and high-quality nest building, but more broadly, is characterized by the heightened motivation and diligence in a mouse directing its attention to pups (Kohl and Dulac, 2018). Traditionally only maternal behaviour was assessed in animal models, and this has also been the case in imprinted gene models associated with parenting, however, other conspecifics are capable of parenting beyond mothers. The paternal role of mice has only recently been appreciated. Male mice are hostile to pups as a general rule, but cohabitation with pregnant females causes a behavioural switch in males, making them diligent and caring parents, capable of licking, grooming, crouching and even performing more demanding behaviours such as retrieval and nest building (Tachikawa et al., 2013). Virgin female mice are another group of animals capable of parenting. Pups seem intrinsically rewarding to female mice, and repeat exposures to pups will lead to virgin females performing standard maternal behaviours, including retrieval, grooming and crouching (Alsina-Llanes et al., 2015). What makes mothers special is their interaction with the offspring while carrying them pre-birth. Their parental investment begins at the day of conception having to provide nutrients and warmth to the offspring but the hormonal events of pregnancy also co-ordinate the maternal

brain to produce the heightened motivation necessary to parent their offspring as soon as they leave the womb (Bridges, 2015).

In this chapter, I used a hybrid Pup Retrieval/Nest Building assessment alongside a Three Chambers pup preference assessment to compare the quality of parenting behaviour between *Magel2*-null mice and WT comparisons. This assessment was carried out with mothers but also fathers and virgin females to provide a more comprehensive assessment of the parenting capacity of the *Magel2*-null mice.

5.2 Methods

5.2.1 Mice & Cohorts

Mice were either Wildtype (WT) or *Magel2*-null on a C57BL/6J background (see Section 2.2.1). Figure 5.1A demonstrates the various experimental cohorts that underwent a parenting behaviour assessment. I tested the parenting behaviour of *Magel2*-null$^{(m+/p-)}$ mice with their own offspring produced from pairing with WT mice. 25 male and female *Magel2*-null$^{(m+/p-)}$ mice were paired with WT mice generated in the same set of litters (but never siblings) making up the experimental cohort. I additionally tested a WT control cohort, comprised of 20 male and female WTs, paired together to produce WT litters. The prerequisite for inducing parenting behaviour in murine fathers is an extended cohabitation phase post-coitus with a pregnant female, and so all males used in my experiments were permanently co-housed with the females that would produce their litter. Finally, a separate cohort of virgin/naïve females, comprising 20 *Magel2*-null$^{(m+/p-)}$ and 20 WT virgin females, was also assessed.

5.2.2 Power Calculation

The optimization experiment (see Section 2.2.4 & Appendix A4) performed with first-time and second-time mothers and fathers allowed an effective power calculation for this main behavioural experiment. Since *Magel2* has never been associated with parenting behaviour before, it was unlikely that the effect on parenting would be large and so I anticipated a subtle reduction in parental motivation. A good approximation for this small effect was the difference in retrieval times between primiparous females (mean = 186 seconds, SD = 66 seconds) and multiparous females (mean = 117 seconds, SD = 83 seconds) in my pilot. To detect differences between these groups with an alpha of 0.05 and power of 0.9, a sample size of 19 animals per group was required. Due to the detrimental effects that *Magel2* was

purported to have on pregnancy rates and litter survival (Mercer and Wevrick, 2009) and the need for viable litters to have 3 or more pups, I paired an excess of mutant animals with WT (25 each) as well as 20 WT x WT pairings to ensure the number of 19 was reached for each group. Mutant male reproduction was slightly poorer with several pairings failing to get pregnant but mutant females were surprisingly unaffected and, consequently, I was overpowered in this group as all animals that produced viable litters and were assessed.

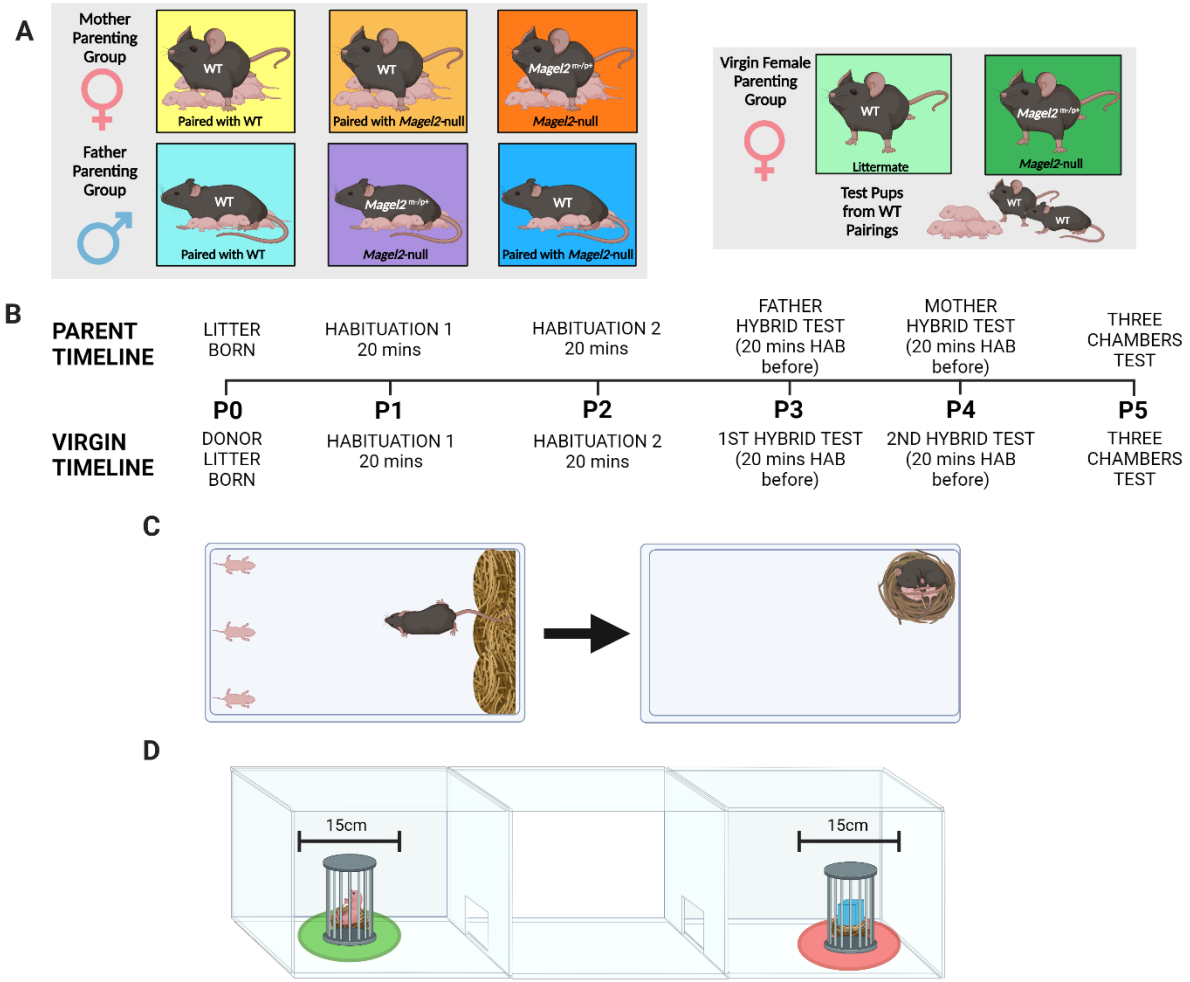


Figure 5.1. Chapter 5 Behavioural Paradigm and Set up. (A) Behavioural cohorts for parenting assessment in mothers & fathers (Left) and virgin females (Right). The maternal and paternal cohorts were paired with each other as indicated and tested with their own litters; virgin females were tested with donor pups from separate WT x WT pairings. (B) Timeline indicating the order in which tests and habituations were carried out based on the day the test litter was born. Above the timeline are events occurring for mothers and fathers while events occurring for virgin females are indicated below. (C) Retrieval/Nest Building task. *Left* – set up (pre-recording) with 3 pups displaced to one short side of the cage and the home nest deconstructed against the opposite short side. *Right* – example of finished behavioural test with all three pups retrieved and visible nest re-constructed from the scattered material. (D) Three Chambers Test. Three pups used in retrieval were placed under a protective cage in one side chamber and a novel object was placed in an identical cage in the opposite side chamber. Time spent in a 15 cm zone around the pup and novel object cage was measured for the pup-preference and pup-aversion scores.

5.2.3 Behavioural Testing

Behavioural testing occurred as detailed in Sections 2.2.3 - 2.2.5 and summarized in Figure 5.1B. Pairs producing viable litters and virgin females with viable litters produced for them, were habituated to the testing apparatus for two days following birth of the litter. On P3, fathers and virgin females (first exposure) were tested and on P4, mothers and virgin females (second exposure) were tested on the Retrieval/Nest Building assessment (Figure 5.1C). On P5 all animals carried out a Three Chambers assessment (Figure 5.1D). The same three pups were used for the two assessments, selected at P2 for being the heaviest. All videos were recorded on a standard HD webcam positioned directly above the testing apparatus.

5.2.4 Metrics

The Retrieval/Nest Building assessment was scored at the millisecond level using BORIS and enabled quantification of several metrics (see Section 2.2.4). The hour time period began from the instance that the test mouse first sniffed any of the pups in the trial. From here, I recorded the time taken to retrieve each of the pups to the nest area, the time taken to construct a nest of sufficient quality and the time taken to complete the trial (defined as the mouse having retrieved all three pups to a suitable quality nest). I scored the quality of the final nest built on a 1-5 scale, Deacon (2006). Nests were scored on completion of the trail, or in the event that all pups were not successfully retrieved, after 60 minutes had expired. I also scored the amount of time that the animals spent performing pup-directed behaviour, defined as any of the following: sniffing pups, licking pups, grooming pups, carrying pups, nest building while pups are inside the nest, crouching/sitting in the nest while pups are inside. This was then scored as a proportion of the total time it took animals to retrieve all three pups and to finish the task. All metrics were scored blind of genotype by the primary scorer (MJH) and 80/210 videos were also second scored by a second blind researcher (Miss. Anna Webberley). Interclass correlations coefficients on all metrics were greater than 0.75 with most scoring greater than 0.9 indicating a good/excellent level of agreement between the primary and secondary scorer. (Table 5.1)

The Three Chambers assessment was scored using Ethovision (see Section 2.2.5). I assessed the number of seconds that the test mice spent in the pup chamber compared to the object chamber, but more indicative of interest in pups, I recorded the time the animal spent within a 15 cm diameter of the pup cage as well as the time spent within a 15 cm diameter of the novel object cage.

Motility analysis was also carried out using Ethovision. Velocity was calculated for mothers, fathers and virgin females during the Retrieval/Nest Building and Three Chambers assessment as well as the number of chamber crosses that each cohort performed in the latter. Whether animals with mutant test pups retrieved WT pups preferentially was also assessed.

Table 5.1 Intra class correlation coefficient (ICC) Kappa and *p* values as measures of the primary and second scorer inter-rater reliability. Statistics were calculated from the 80/210 videos that were second scored.

Metrics from Retrieval/Nest Building Assessment	ICC Kappa (0-1)	ICC *p* value
Time to Retrieve Pup 1	0.915	$1.89x10^{-33}$
Time to Retrieve Pup 2	0.936	$3.05x10^{-38}$
Time to Retrieve Pup 3	0.958	$3.05x10^{-45}$
Number of Pups Retrieved	0.935	$6.12x10^{-38}$
Time to Build a Level 3 Nest	0.894	$7.44x10^{-30}$
Final Nest Quality Rating	0.778	$4.61x10^{-18}$
Time Until Task Finished	0.872	$9.33x10^{-27}$
PDB Until Retrieval	0.853	$2.36x10^{-24}$
PDB Until Task Finished	0.928	$3.46x10^{-36}$

5.2.5 Statistics and Figures

For behavioural measures for mothers and fathers, all continuous variable analyses were performed using one-way ANOVAs and post-hoc pairwise t tests if the data met normality assumptions while virgin females were analysed using two-way Mixed ANOVAs with Genotype and Exposure as variables. If normality assumptions were not met, then log transformations were used. If date were deemed non-normal/categorical, analysis was performed using the Kruskal Wallis Test followed by pairwise Dunn Tests or via the R package nparLD (Noguchi et al., 2012) which provides a rank-based alternative for analysing longitudinal data in factorial settings, Proportion variables were corrected using an arc sine correction. Three Chamber assessments were analysed using two-sided one sample t tests.

5.3 Results

5.3.1 Magel2$^{m+/p}$ mice parenting assessment and confounders

I assessed parental behaviour in *Magel2*-null mice (paternal transmission of ablated allele), using three distinct groups of mice capable of parenting behaviour: primiparous mothers, first-time fathers and naïve virgin females. These groups were first tested using a combined Retrieval/Nest Building paradigm (e.g. Stagkourakis et al., 2020) in which each animal had

one hour to retrieve 3 scattered pups alongside reconstructing their deconstructed home nest. This was followed on a subsequent day by a Three Chambers pup-preference test (Liu et al., 2013) in which the same 3 pups were placed in one side chamber and a novel object placed in the other and the time spent in proximity of these across a 10-minute span was recorded.

Several factors can influence parenting behaviour indirectly such as litter size, parent motility and olfaction. I saw no significant differences in litter size recorded at P2 (Appendix Figure A7.A) between the cohorts of mothers and fathers and litter loss was equivalent between the three pairings (13 from *Magel2*$^{m+/p-}$ female pairings, 20 from *Magel2*$^{m+/p-}$ male pairings and 21 from WT x WT pairings). There were no significant differences in time taken to first sniff and investigate the pups (Appendix Figure A7.B) indicating no overt olfactory deficit. There were also no overt motility disadvantages between the *Magel2*$^{m+/p-}$ and WT individuals in each group (mothers, fathers, virgins), with no significant differences in velocity in the retrieval task (Appendix Figure A7.C), and no differences in number of times moving between the chambers in the Three Chambers assessment (Appendix Figure A7.D).

5.3.2 Magel2$^{m+/p-}$ mothers displayed poorer nest building and less pup-directed motivation

The three maternal cohorts were as follows: **WT(WT)** - WT female paired with WT male, mothering WT pups, **WT (*Magel2*)** - WT female paired with *Magel2*$^{m+/p-}$ male, mothering WT and mutant pups (*Magel2*$^{m+/p-}$) and ***Magel2*$^{m+/p-}$** - *Magel2*$^{m+/p-}$ female paired with WT male, mothering WT and functionally WT pups (*Magel2*$^{m-/p+}$).

Success rate in the task differed between the three maternal cohorts (Figure 5.2B, $H(2) = 20.86$, $p = 2.95 \times 10^{-5}$). *Magel2*$^{m+/p-}$ mutant mothers paired with WT studs displayed a significantly worse performance during the Retrieval/Nest Building task than both WT(WT) ($p = 0.0004$) and WT(*Magel2*) ($p = 0.0003$). Both WT maternal cohorts successfully retrieved all 3 pups and rebuilt their nest in the one-hour time frame whereas only 56% of *Magel2*$^{m+/p-}$ mothers achieved the same. The time taken to complete the task differed between the maternal cohorts ($F(2, 62) = 21.48$, $p = 8.16 \times 10^{-8}$) with both WT(WT) ($p = 1.9 \times 10^{-6}$) and WT(*Magel2*) ($p = 1.6 \times 10^{-6}$) completing the task faster than the *Magel2*$^{m+/p-}$ mothers (Figure 5.2C).

During the Retrieval/Nest Building task, 100% of the *Magel2*$^{m+/p-}$ mothers successfully retrieved the three pups to the nest area (Figure 5.2F). There were no significant differences

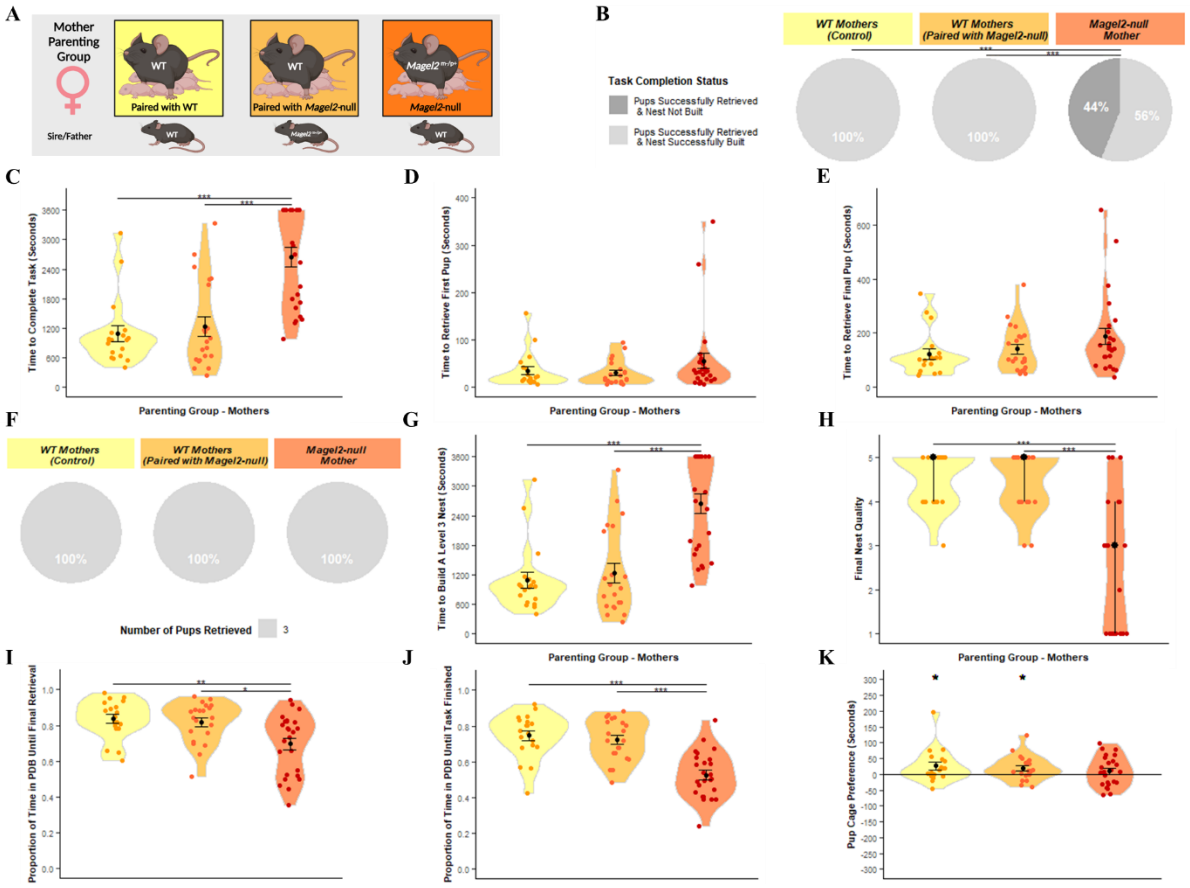


Figure 5.2. Mother Parenting Assessment. (A) Schematic of behavioural paradigm with mothers. WT (Paired with WT) *n = 19*, WT (Paired with *Magel2*-null male) *n* = 21, *Magel2*-null *n* = 25 (B) Task Completion Status at conclusion of Retrieval/Nest Building Task. Mothers were categorised on their ability to rebuild their nest to a level 3 quality and to retrieve the pups into the nest within the one-hour time limit. Percentages of mothers falling within those categories are shown. (C) Time taken to complete the Retrieval/Nest Building Task. Time taken to retrieve all three pups to the area where the nest was rebuilt and to rebuild the nest to a level 3 quality or higher. (D) Time taken to retrieve the first pup to the nest in the Retrieval/Nest Building Task. (E) Time taken to retrieve the final/third pup to the nest in the Retrieval/Nest Building Task and hence completing the retrieval portion of the task. (F) Number of pups retrieved at conclusion of Retrieval/Nest Building Task. Mothers were categorised on the number of pups they successfully retrieved and percentages of mothers falling within those categories are shown. (G) Time taken to re-build the nest within the Retrieval/Nest Building task to a level 3 quality or higher. Time was recorded for when the nest being constructed by the mothers scored a level 3 quality score (the point when the nest takes functional shape). (H) Nest Quality score at conclusion of Retrieval/Nest Building Task. Mother's rebuilt nests were scored from 0 – 5 upon completion of the test. (I) Proportion of time spent engaged in pup-directed behaviour until the final/third pup was retrieved to the nest/until task time expires in the Retrieval/Nest Building Task. Mother's behaviours were scored continuously through the one-hour trial. Pup-directed behaviour included time spent engaging in licking, grooming, sniffing, retrieving pups, alongside nest building and crouching in nest (only while pups were present in the nest). (J) Proportion of time spent engaged in pup-directed behaviour until the final/third pup was retrieved to the nest and the nest was rebuilt to a level 3 standard/until task time expires in the Retrieval/Nest Building Task. (K) Pup preference score in Three Chambers Assessment. Pup preference scores was calculated as time the mother spent within a 15 cm zone around the pups minus time spent within a 15 cm zone around the novel object. Positive values indicate a preference for proximity to pups. Significance for continuous variables determined using one-way ANOVA and Bonferroni-corrected pairwise t tests. Significance for categorical variables determined using Kruskal-Wallis test and Bonferroni-corrected Dunn test. Statistical significance: *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$

in time taken to retrieve the first pup (Figure 5.2D; F(2, 62) = 1.35, p = 0.271) and final pup (Figure 5.2E; F(2, 62) = 1.98, p = 0.07), indicating that *Magel2*$^{m+/p-}$ mothers have comparable retrieval ability to their WT comparisons. However, 46% of *Magel2*$^{m+/p-}$ mothers failed to build a suitable quality nest and the maternal cohorts differed in both the time taken to rebuild the home nest to a Level 3 state (Figure 5.2G; F(2, 62) = 21.48, $p = 8.16 \times 10^{-8}$) and the final quality of the rebuilt nest (Figure 5.2H, H(2) = 20.06, $p = 4.40 \times 10^{-5}$), *Magel2*$^{m+/p-}$ mothers were slower to build a level 3 nest than WT(WT) ($p = 1.90 \times 10^{-6}$) and WT(*Magel2*) ($p = 1.6 \times 10^{-6}$) and had significantly poorer quality nests than WT(WT) (p = 0.0003) and WT(*Magel2*) (p = 0.0007).

In addition to differences in nest building, there was a difference between the maternal cohorts in the proportion of time that mothers spent in pup-directed behaviour up until that successful final retrieval (Figure 5.2I; F(2, 62) = 7.12, p = 0.002). *Magel2*$^{m+/p-}$ mothers spent a significantly smaller proportion of their time leading up to the successful final retrieval engaging in pup-directed behaviour compared to WT(WT) (p = 0.0035) and WT(*Magel2*) (p = 0.013). This difference was also found when considering the proportion of time spent in pup-directed-behaviour until the task was finished (either upon completion of the task or upon the expiration of the one-hour testing time; Figure 5.2J; F(2, 62) = 21.02, $p = 1.07 \times 10^{-7}$) with *Magel2*$^{m+/p-}$ mothers spending a smaller proportion of their time compared to WT(WT) ($p = 18.2 \times 10^{-7}$) and WT(*Magel2*) ($p = 6.80 \times 10^{-6}$). The Three Chambers assessment was used as a second independent measure of pup affiliation and parental motivation and WT(WT) (t(17) = 2.15, p = 0.045) and WT(*Magel2*) (t(20) = 2.37, p = 0.028) both spent significantly more time in vicinity of the pups than the novel object and hence demonstrated a pup-preference (Figure 5.2K). *Magel2*$^{m+/p-}$ mothers did not demonstrate a significant pup-preference score (t(24) = 1.07, p = 0.29).

5.3.3 Magel2$^{m+/p-}$ fathers performed poorly on all measures of parenting behaviour

The three paternal cohorts were produced from the same pairing as the maternal cohorts and were as follows: **WT(WT)** (WT male paired with WT female fathering WT pups), **WT(*Magel2*)** (WT male paired with *Magel2*$^{m+/p-}$ female fathering WT and functionally WT pups (*Magel2*$^{m-/p+}$) and ***Magel2*$^{m+/p-}$** (*Magel2*$^{m+/p-}$ male paired with WT female fathering WT and mutant pups (*Magel2*$^{m+/p-}$))

Success rate in the task differed between the three paternal cohort (Figure 5.3B; H(2) = 26.86, $p = 1.47 \times 10^{-6}$). *Magel2*$^{m+/p-}$ fathers displayed a significantly worse performance during the Retrieval/Nest Building task than both WT(WT) ($p = 5.00 \times 10^{-6}$) and WT(*Magel2*) ($p = 0.0001$). 79% of WT(WT) completed the test successfully, 60% of WT(*Magel2*) and only 9% of *Magel2*$^{m+/p-}$ fathers successfully completed the task. All paternal cohorts had a percentage of failures, but the time taken to complete the task differed between the paternal cohorts (F(2, 63) = 13.24, $p = 1.59 \times 10^{-5}$) with *Magel2*$^{m+/p-}$ fathers completing the task slower than both WT(WT) ($p = 2.00 \times 10^{-5}$) and WT(*Magel2*) ($p = 0.001$). (Figure 5.3C).

Focusing on the retrieval component, the paternal cohorts differed in the number of pups they retrieved during the task (Figure 5.3F; H(2) = 23.06, $p = 9.83 \times 10^{-6}$). *Magel2*$^{m+/p-}$ fathers retrieved significantly fewer pups than the WT(WT) ($p = 8.30$x10-5) and WT(*Magel2*) ($p = 0.0001$). 64% of *Magel2*$^{m+/p-}$ fathers failed to retrieve any pups while only 20% and 16% of WT fathers failed to retrieve no pups. There were significant differences between the paternal cohorts in the time taken to retrieve both the first pup (Figure 5.3D; F(2, 63) = 11.52, $p = 5.44 \times 10^{-5}$) and last pup (Figure 5.3E; F(2, 63) = 12.86, $p = 4.59 \times 10^{-5}$). *Magel2*$^{m+/p-}$ fathers were significantly slower to retrieve the first pup compared to WT(WT) ($p = 0.0001$) and WT(*Magel2*) ($p = 0.001$), and the final pup compared to WT(WT) ($p = 0.0001$) and WT(*Magel2*) ($p = 0.0005$). Within the nest building component, the paternal cohorts differed in both the time taken to rebuild the home nest to a Level 3 state (Figure 5.3G; F(2, 63) = 18.72, $p = 4.15 \times 10^{-7}$) and the final quality of the rebuilt nest (Figure 5.3H, H(2) = 12.02, $p = 1.47 \times 10^{-6}$), *Magel2*$^{m+/p-}$ fathers were slower to build a level 3 nest than WT(WT) ($p = 3.60 \times 10^{-7}$) and WT(*Magel2*) ($p = 3.2 \times 10^{-4}$) and had significantly poorer quality nests than WT(WT) ($p = 0.002$) but not WT(*Magel2*) ($p = 0.66$).

There were differences between the paternal cohorts in the proportion of the time devoted to pup-directed behaviour for the time until the final pup was retrieved (Figure 5.3I; F(2, 63) = 5.90, $p = 0.004$) and until the task was finished/one hour expired (Figure 5.3J; F(2, 63) = 9.52, $p = 0.0002$). *Magel2*$^{m+/p-}$ fathers dedicated a smaller proportion of their time to pup-directed behaviour than the WT(WT) until final retrieval ($p = 0.006$) and until task finished ($p = 0.031$) and, than the WT(*Magel2*) until final retrieval ($p = 0.0002$) and until task finished ($p = 0.011$). None of the fathers demonstrated a significant pup-preference score in the Three Chambers test (Figure 5.3K). However, *Magel2*$^{m+/p-}$ fathers ($t(20) = -3.18$, $p = 0.005$) and WT(*Magel2*) fathers ($t(24) = -2.58$, p = 0.016) demonstrated a significant preference for the object zone compared to the pup zone, so a significant pup-avoidance score.

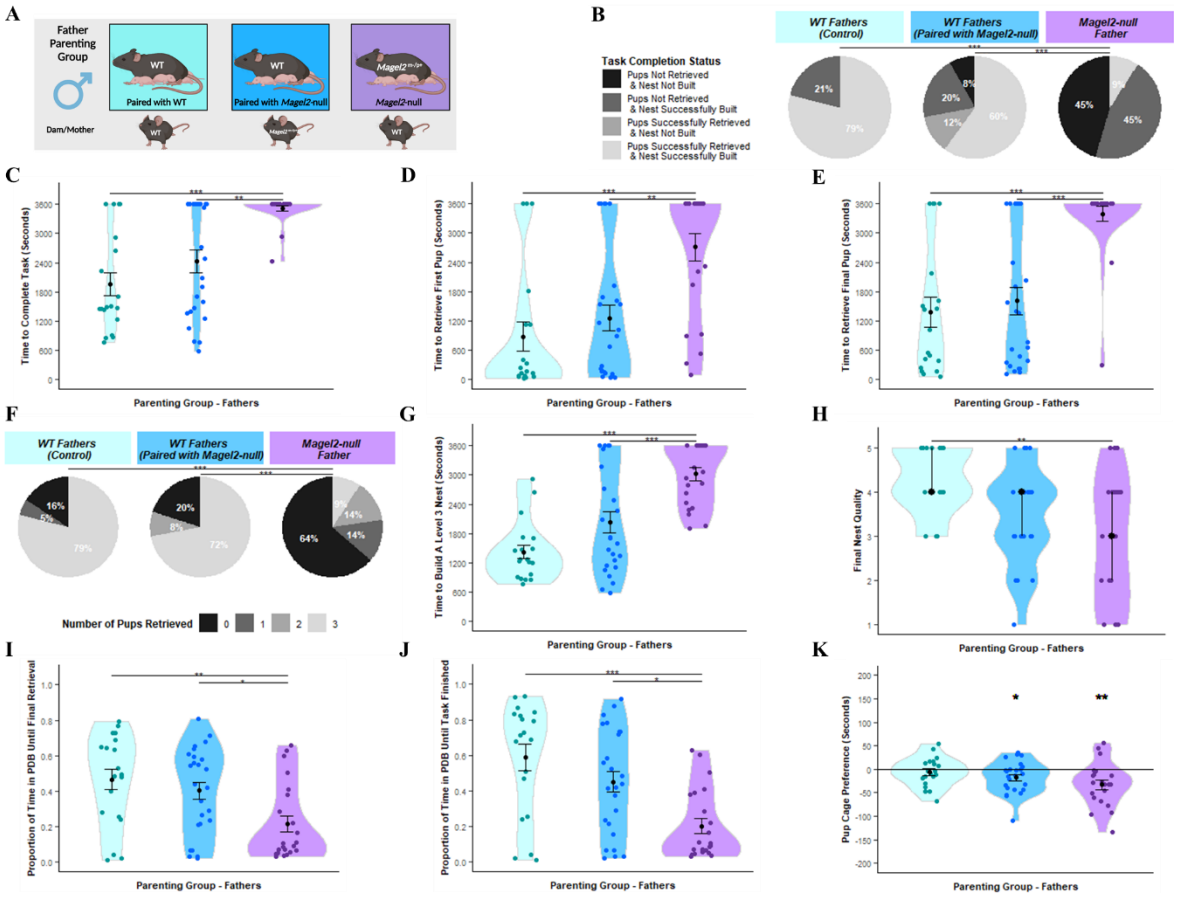


Figure 5.3. Father Parenting Assessment. (A) Schematic of behavioural paradigm with fathers. WT (Paired with WT) *n = 19*, WT (Paired with *Magel2*-null female) *n* = 25, *Magel2*-null *n* = 22 (B) Task Completion Status at conclusion of Retrieval/Nest Building Task. Fathers were categorised on their ability to rebuild their nest to a level 3 quality and to retrieve the pups into the nest within the one-hour time limit. Percentages of fathers falling within those categories are shown. (C) Time taken to complete the Retrieval/Nest Building Task. Time taken to retrieve all three pups to the area where the nest was rebuilt and to rebuild the nest to a level 3 quality or higher. (D) Time taken to retrieve the first pup to the nest in the Retrieval/Nest Building Task. (E) Time taken to retrieve the final/third pup to the nest in the Retrieval/Nest Building Task and hence completing the retrieval portion of the task. (F) Number of pups retrieved at conclusion of Retrieval/Nest Building Task. Fathers were categorised on the number of pups they successfully retrieved and percentages of fathers falling within those categories are shown. (G) Time taken to re-build the nest within the Retrieval/Nest Building task to a level 3 quality or higher. Time was recorded for when the nest being constructed by the fathers scored a level 3 quality score (the point when the nest takes functional shape). (H) Nest Quality score at conclusion of Retrieval/Nest Building Task. Father's rebuilt nests were scored from 0 – 5 upon completion of the test. (I) Proportion of time spent engaged in pup-directed behaviour until the final/third pup was retrieved to the nest/until task time expires in the Retrieval/Nest Building Task. Father's behaviours were scored continuously through the one-hour trial. Pup-directed behaviour included time spent engaging in licking, grooming, sniffing, retrieving pups, alongside nest building and crouching in nest (only while pups were present in the nest). (J) Proportion of time spent engaged in pup-directed behaviour until the final/third pup was retrieved to the nest and the nest was rebuilt to a level 3 standard/until task time expires in the Retrieval/Nest Building Task. (K) Pup preference score in Three Chambers Assessment. Pup preference scores was calculated as time the father spent within a 15 cm zone around the pups minus time spent within a 15 cm zone around the novel object. Positive values indicate a preference for proximity to pups. Significance for continuous variables determined using one-way ANOVA and Bonferroni-corrected pairwise t tests. Significance for categorical variables determined using Kruskal-Wallis test and Bonferroni-corrected Dunn test. Statistical significance: *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$

5.3.4 Magel2$^{m+/p-}$ pups had no significant effect on task completion

It has already been suggested that pups inheriting the paternally mutated allele of *Magel2* have behavioural differences that can influence maternal preference during retrieval (Bosque Ortiz et al., 2022). Additionally, *Magel2*$^{m+/p-}$ pups are known to have growth deficits (Mercer and Wevrick, 2009, Bischof et al., 2007). Hence, in an attempt to minimize the number of *Magel2*$^{m+/p-}$ pups that made it into the test litters of *Magel2*$^{m+/p-}$ fathers and the paired WT(*Magel2*) mothers, I weighed all pups at P2 and marked the three heaviest as the test pups. This was done for all maternal and paternal cohorts. This approach proved successful as genotyping of all test pups showed that while 45% of maternal-inherited test pups were *Magel2*$^{m-/p+}$ (from *Magel2*$^{m+/p-}$ mothers paired with WT males), only 24% of paternally inherited test pups were *Magel2*$^{m+/p-}$ (from *Magel2*$^{m+/p-}$ fathers paired with WT females). This meant that half of the *Magel2*$^{m+/p-}$ fathers and their associated WT mothers had no mutant pups in their test litters.

However, 11/22 of the *Magel2*$^{m+/p-}$ fathers and their associated WT mothers still had at least one mutant pups in their assessment. To assess whether mutant pups were influencing the outcome of my assessment, I first compared the average retrieval times for a mutant pup in these 11 litters compared to the WT littermates (Figure 5.4), I found no significant differences in the time for WT(*Magel2*) mothers to retrieve a mutant pup compared to WT pups (W = 117, p = 0.532, Wilcoxon Ranked-Sum Test) and the same was seen for *Magel2*$^{m+/p-}$ fathers (W = 117, p = 0.42, Wilcoxon Ranked-Sum Test). Secondly, I reran the analyses of the previous sections (mothers and fathers) while excluding the litters containing mutant pups and produced the same statistical disparities as previous,

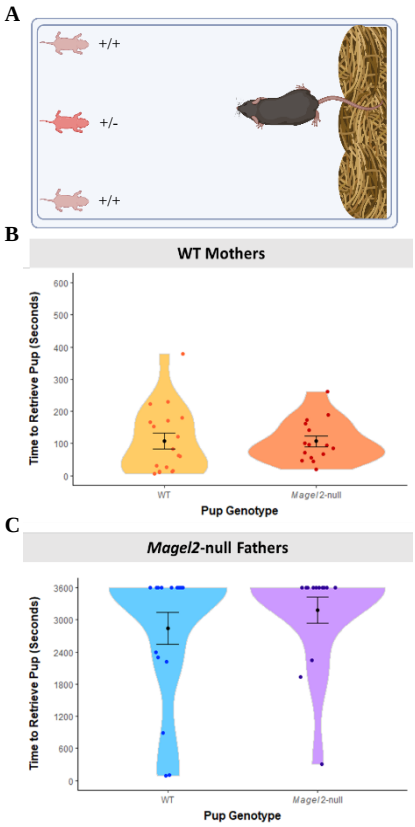


Figure 5.4. Pup retrieval times for WT (n = 18) and *Magel2*-null pups (n = 15) within mixed genotype litter retrievals (n = 11). Mixed litters could only result from pairings with WT females and *Magel2*-null males. (A) Schematic showing retrieval set up with a mutant pup present as one of the three animals to be retrieved. Animals had a maximum of 3600 seconds to retrieve pups. (B) Time to retrieve WT and mutant pups for WT mothers (C) Time to retrieve WT and mutant pups for *Magel2*-null fathers.

further suggesting that the presence of *Magel2*$^{m+/p-}$ pups was not influencing the parental behaviour I observed at P3-P5 (Appendix A8).

5.3.5 Magel2$^{m+/p-}$ *virgin females displayed poorer retrieval behaviour and less pup-directed motivation*

Virgin females display parenting behaviour spontaneously, with less reliability than mothers, but subsequent exposures to pups improves the reliability of parenting behaviour manifesting (Alsina-Llanes et al., 2015, Martín-Sánchez et al., 2015). To incorporate this improvement effect, I had the virgin females undergo two Retrieval/Nest Building tests on subsequent days before performing the Three Chambers test on the following day. Cohorts consisted of ***Magel2***$^{m+/p-}$ mutant females and their **WT** littermates, tested with a unique set of three WT pups derived from WT x WT pairings.

For success rate in the task (Figure 5.5B), there was a main effect of Genotype ($H(1) = 12.36$, $p = 4.39 \times 10^{-4}$) with *Magel2*$^{m+/p-}$ virgin females displayed a significantly worse performance during the hybrid-retrieval task than the WTs and a main effect of Exposure ($H(1) = 12.80$, $p = 3.46 \times 10^{-4}$) with a higher task success in the second exposure. The WT success rate was 75% on first exposure and 90% on the second exposure, while for *Magel2*$^{m+/p-}$ virgin females they had a 30% success rate followed by a 65% success rate. For time to complete the task (Figure 5.5C), there was a significant interaction effect ($F(1,38) = 6.22$, $p = 0.017$) and simple main effects analysis revealed that *Magel2*$^{m+/p-}$ virgin females took longer to finish the task compared to WT virgin females only in the first exposure ($p = 0.0078$) but not in the second. *Magel2*$^{m+/p-}$ virgin female also saw significant improvement between the first and second exposure ($p = 0.0064$) while the WT virgin females did not.

Focusing on the retrieval component, for the number of pups retrieved (Figure 5.5F), there was a significant main effect of Genotype ($H(1) = 9.62$, $p = 0.0019$) with *Magel2*$^{m+/p-}$ virgin females having retrieved fewer pups than WTs. There was a main effect of Exposure ($H(1) = 11.68$, $p = 0.0006$) with more pups retrieved in the second exposure. 80% of WT virgin females retrieve all 3 pups in the first exposure whereas on 35% of *Magel2*$^{m+/p-}$ virgin females did.

Figure 5.5D and 5.5E displays the time taken to retrieve the first and last pups, respectively. For the time to retrieve the first pup, there was a significant interaction effect ($F(1,38) = 4.995$, $p = 0.031$) and simple main effects analysis revealed that *Magel2*$^{m+/p-}$ virgin females

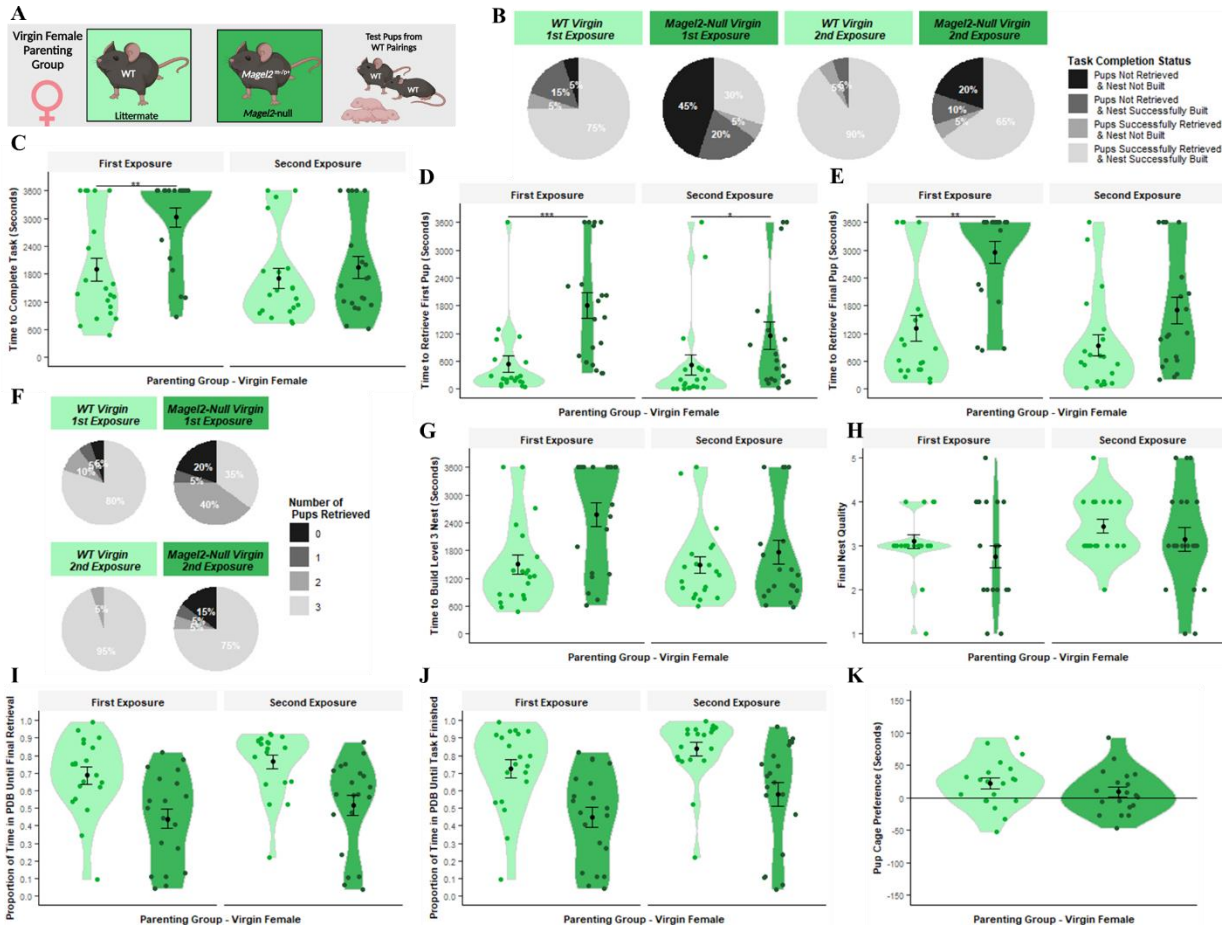


Figure 5.5. Virgin Female Parenting Assessment. (A) Schematic of behavioural paradigm with virgin females. WT (Littermate) *n = 20, Magel2*-null *n* = 20. Each female was tested with 3 unique pups acquired from WT x WT pairings. The Retrieval/Nest Building Task was carried out twice for each female (First Exposure and Second Exposure). (B) Task Completion Status at conclusion of the first and second Retrieval/Nest Building Task. Virgin females were categorised on their ability to rebuild their nest to a level 3 quality and to retrieve the pups into the nest within the one-hour time limit. Percentages of virgin females falling within those categories are shown. (C) Time taken to complete the Retrieval/Nest Building Task. Time taken to retrieve all three pups to the area where the nest was rebuilt and to rebuild the nest to a level 3 quality or higher. (D) Time taken to retrieve the first pup to the nest in the Retrieval/Nest Building Task. (E) Time taken to retrieve the final/third pup to the nest in the Retrieval/Nest Building Task and hence completing the retrieval portion of the task. (F) Number of pups retrieved at conclusion of Retrieval/Nest Building Task. Virgin females were categorised on the number of pups they successfully retrieved and percentages of virgin females falling within those categories are shown. (G) Time taken to re-build the nest within the Retrieval/Nest Building task to a level 3 quality or higher. Time was recorded for when the nest being constructed by the virgin females scored a level 3 quality score (the point when the nest takes functional shape). (H) Nest Quality score at conclusion of Retrieval/Nest Building Task. Virgin female's rebuilt nests were scored from 0 – 5 upon completion of the test. (I) Proportion of time spent engaged in pup-directed behaviour until the final/third pup was retrieved to the nest/until task time expires in the Retrieval/Nest Building Task. Virgin female's behaviours were scored continuously through the one-hour trial. Pup-directed behaviour included time spent engaging in licking, grooming, sniffing, retrieving pups, alongside nest building and crouching in nest (only while pups were in nest). (J) Proportion of time spent engaged in pup-directed behaviour until the final/third pup was retrieved to the nest and the nest was rebuilt to a level 3 standard/until task time expires in the Retrieval/Nest Building Task. (K) Pup preference score in Three Chambers Assessment. Pup preference scores was calculated as time the virgin female spent within a 15 cm zone around the pups minus time spent within a 15 cm zone around the novel object. Positive values indicate a preference for proximity to pups. Significance for continuous variables determined using two-way ANOVA and Bonferroni-corrected pairwise t tests. Significance for categorical variables determined using Kruskal-Wallis test and Bonferroni-corrected Dunn test. Statistical significance: $*p < 0.05$, $**p < 0.01$, and $***p < 0.001$

took longer to retrieve the first pup in both the first exposure (p = 0.00004) and in the second exposure (p = 0.025). Neither *Magel2*$^{m+/p-}$ nor WT virgin females saw significant improvement upon second exposure. For the time to retrieve the final pup, there was a significant interaction effect (F(1,38) = 4.828, p = 0.034) and simple main effects analysis revealed that *Magel2*$^{m+/p-}$ naïve females took longer to retrieve the first pup in the first exposure (p = 0.00027) but not in the second (p = 0.071). *Magel2*$^{m+/p-}$ naïve female also saw significant improvement between the first and second exposure (p = 0.0086) while the WT naïve females did not.

In the nest building component, there was a main effect of Genotype for the time taken to construct a Level 3 nest (Figure 5.5G; F(1,38) = 6.76, p = 0.013) with WTs building level 3 nests faster, as well as a main effect of Exposure (F(1,38) = 4.41, p = 0.043) with level 3 nests built faster in the second exposure. For nest quality at the end of the assessment (Figure 5.5H), there was a main effect of Exposure (H(1) = 4.61, p = 0.032) with higher quality nests built in the second exposure but no main effect of Genotype, which indicates that while virgin mice of both genotypes don't tend to build high quality nests within the hour, the WTs are still quicker to build a suitable nest for the pups.

When considering pup directed behaviour, there was a main effect of Genotype for the pup-directed behaviour up to final retrieval (Figure 5.5I; F(1,38) = 15.28, p = 0.0004) and for the pup-directed behaviour until task finished (Figure 5.5J; F(1,38) = 15.90 , p = 0.0003),with *Magel2*$^{m+/p-}$ naïve females dedicating a smaller proportion of their time to pup-directed behaviour than the WTs. There was a main effect of Exposure for the pup-directed behaviour until retrieval (F(1,38) = 4.311 , p = 0.045) and task finished (F(1,38) = 11.14, p = 0.002) with more PDB displayed in the second exposure. Additionally, the WT naïve females (following their two exposures) demonstrated a significant pup-preference score during the Three Chambers assessment (Figure 5.5K; t(19) = 2.70, p = 0.014). This was not true for the *Magel2*$^{m+/p-}$ naïve females who failed to show a preference for either the pups or the novel object (t(18) = 1.15, p = 0.26).

5.4 Discussion

Magel2 is a paternally expressed gene and an integral member of the Prader-Willi syndrome locus, capable of producing its own syndrome in humans when lost – Schaaf-Yang Syndrome (Fountain and Schaaf, 2016). Mice null for paternal *Magel2* have been extensively

characterised, with phenotypes seen in metabolism (Kamaludin et al., 2016), feeding (Bischof et al., 2007), and several deficits in neonates including suckling (Schaller et al., 2010) and USV production (Bosque Ortiz et al., 2022). Due to its clinical relevance for development, *Magel2* might not be suspected of having a role in parenting behaviour. However, here I have shown that loss of *Magel2* results in deficits in parenting behaviour in mothers, fathers and virgin female mice in retrieval, nest building and motivation to interact with pups.

Mothers, fathers and virgin female mice have significant overlap in the neural circuitry necessary to produce parenting behaviour; they have the same POA^{Gal} hub (Kohl et al., 2018) and specifically the same high activity in *Gal/Calcr* of the POA (Moffitt et al., 2018). Mothers are unique in being primed by the hormonal events of pregnancy in advance of experience (Bridges, 2015, Rilling and Young, 2014), whereas fathers are dependent on social cohabitation with pregnant females for the duration of the pregnancy to transition their default aggressive behaviour towards pups to reliable parental care while their pups mature via vomeronasal suppression (Tachikawa et al., 2013, Rogers and Bales, 2019). Virgin females on the other hand display 'spontaneous maternal behaviour' in which a certain proportion will display full maternal behaviour towards pups when first exposed (Alsina-Llanes et al., 2015, Martín-Sánchez et al., 2015). This proportion steadily increases upon subsequent exposures until 100% of these animals will display some parenting behaviour, although not to the same level of motivation and reliability as mothers. Finding deficits in one group but not others would suggest *Magel2* was impacting a group specific mechanism. For instance, deficits only in mothers would suggest a disruption to the priming effect of pregnancy hormones or to the hormonal output of the placenta. Here I do find an insult to all three groups however fathers and virgins are more severely affected as opposed to mothers (who did not have altered retrieval behaviour). This suggests that either, *Magel2* is affecting the learning/improvement nature of parenting (and so affecting the un-primed groups more) or that a core mechanism, or multiple overlapping mechanisms, are affected in the *Magel2*-null mice but the hormonally primed nature of murine mothers compensates for this global insult, protecting them from the full impact of the mutation on their performance on essential tasks like retrieval which we see in fathers and virgins.

Several indirect mechanisms could account for some of this parenting deficit. Common issues such as motility and olfactory/sensory issues could cause parenting deficits indirectly. Indeed, olfactory deficits have been previously reported for *Magel2*-null mice in a reproductive setting (Mercer and Wevrick, 2009). Nevertheless, here *Magel2*-null parents displayed

comparable pup sniffing behaviour to WTs, as well as velocity, suggesting impaired olfaction or motility is unlikely to explain the parenting deficits. Pup behaviour can worsen parenting behaviour but here I minimised mutant pup presence in the test and showed that these pups were not impacting the parenting deficits I observed. Hormonal regulation issues could cause parenting deficits and *Magel2* has been associated with endocrine insufficiency and improper ovary and pituitary function (Tennese and Wevrick, 2011). This could indeed explain the deficits in mothers and fathers (who rely on prolactin) but would not explain the results of the virgin females. Given *Magel2*'s hypothalamic expression, a core neuronal mechanisms could explain these deficits. One predicted issue could be some reduction or irregular functioning in oxytocin neurons, a crucial neural population when it comes to promoting pup-directed engagement and bonding. This has already been identified in *Magel2*$^{+/-}$ mice (Ates et al., 2019, Reichova et al., 2021, Schaller et al., 2010) and also seen in *Peg3* mice whose parenting deficit has been established previously (Li et al., 1999, Frey et al., 2018). However, a previously unconsidered mechanism, implicated by the findings in Chapter 3 and 4, is some difference in number and/or performance of galanin neurons in the POA, an investigation that makes up Chapter 6 of this thesis.

5.4.1 Caveats

There are a few caveats to this chapter. Firstly, parenting behaviour is a delicate behaviour which means it is easy to produce a deficit indirectly. This is particularly true considering I was using a global knockout, as loss of *Magel2* in other tissues and/or systems in the body could impact parenting indirectly. Several of these indirect effects on parenting have been explored (including olfaction, motility, litter size, pup behaviour) but others yet remain unexplored such as the impact of physiological insult, such as sleep issues, metabolism issues or mood, affecting behaviour. For mood, higher anxiety is known to worsen maternal performance (Kessler et al., 2011) however *Magel2*-null mice have been reported to have lower anxiety (Fountain et al., 2017b) suggesting this would not impact the results. Models with more targeted KO's (such as brain or gal-specific KO's) could determine which organ system was primarily responsible for the deficit (e.g., brain or hormones).

Secondly, although I used two assessments of parenting, most of the metrics came from a single one-hour test. This is standard for parenting assessments in rodents, and this chapter is strengthened with the results of the Three Chambers assessment. However, valid questions can be asked about the translatability of the assessments chosen. In order to confirm *Magel2*

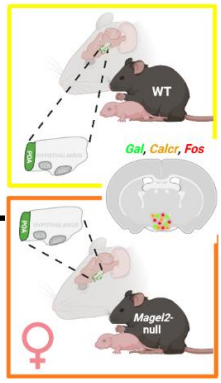
is implicated in parenting behaviour in the undisturbed home nest environment, pre and post birth 24-hour monitoring would highlight the impact *Magel2* KO has on mice preparing for their pups to be born and in the first few days when parenting behaviour is most essential. Any reduction in nursing, poorer quality nest building or reduction in crouching and grooming could be assessed without physically moving parent or pups.

Mutant mothers failed to show a retrieval deficit here. Many studies of parenting assess retrieval behaviour alone, and so to their standards, these mice would not have a parenting deficit. My collection of other metrics do show however, that the mutant mothers display altered parenting behaviour in all other measures of parenting I had available. Retrieval behaviour is an essential behaviour in mice and is the only parenting metric being measured here in which the pups are effectively in danger, exposed to the open. In some ways this would be the behaviour one might expect the most severe deficits to be seen, given its demanding nature. However, given the life-or-death implications of retrieval, this is also the behaviour with the highest motivational component and perhaps the behaviour most likely to be still performed despite some insult to parenting. The mutant maternal success at retrieval (compared to the fathers and virgin females) may reflect the higher level of motivation this group have to perform parenting behaviour in the first place, as well as the rescuing effect of hormonal priming during pregnancy. Therefore, it takes something less essential (such as the time spent in PDB or time to build a nest) to expose the parenting insult in mothers.

5.4.2 Summary of Findings

- Mothers, fathers and virgin female mice null for *Magel2*, showed deficits in retrieval behaviour (fathers and virgins) and nest building (mothers and fathers) which could not be explained by motility or olfactory issues.
- Mothers, fathers and virgin female mice null for *Magel2* also showed a reduction in the proportion of their time they spent engaged in pup-directed behaviour, which was mirrored by their lack of pup-preference during the Three Chambers test.
- *Magel2*-null pups were not less preferentially retrieved and did not impact the retrieval behaviour of their parents.

6 *Magel2*-null Preoptic Area Phenotyping



6.1 Overview

Magel2 mice displayed a significant reduction in parental motivation and ability (Chapter 5). The final chapter of this thesis aimed to shed light on the mechanism by which *Magel2* is affecting parenting behaviour by quantifying the *c-Fos* response of the POA upon exposure to and interaction with pups. Given the enrichment of *Magel2* expression in the POA seen in Chapter 3 and Chapter 4, specifically the dramatic enrichment in *Gal/Calcr* expressing cells of the POA seen in the latter, these neurons seemed like a logical target to assess activity in.

C-Fos expression is a well-known marker of neuronal activity (Bullitt, 1990, Kovács, 2008). Enhanced *c-Fos* response in the POA in response to parenting behaviours has been observed multiple times (Wu et al., 2014, Zhong et al., 2014, Matsushita et al., 2015, Numan and Numan, 1994, Stack and Numan, 2000, Mathieson et al., 2002), and the strongest response was seen in the *Gal/Calcr* neurons following exposure to pups in mothers and fathers (Moffitt et al., 2018). Approaches exist to quantify *c-Fos* activity using RNAscope® (Cosi et al., 2021), and the approach utilised in this chapter (RNAscope with *Gal/Calcr/Fos*) allows both a quantification of *c-Fos* RNA in POA cells generally as well as in POA *Gal/Calcr* cells specifically, but it also allowed me to assess any altered pathology in *Gal* or *Calcr* expression in this parenting-critical brain region.

Overall, this approach aimed to identify the mechanism by which *Magel2* is affecting parenting behaviour in the maternal behavioural group and may further suggest new mechanisms by which imprinted genes could be influencing parenting directly, by modulating the galanin neurons of the POA.

6.2 Methods

6.2.1 Maternal cohorts

Figure 6.1 displays an overview of the maternal cohorts used in this chapter. Maternal mice were either *Magel2*-null (*Magel2*$^{(m+/p-)}$, N=8) or Wildtype littermates (WT, N=8) on a C57BL/6J background. Four mice from each group (WT and *Magel2*-null) were permanently cohoused with a WT non-sibling male to produce litters. All mice produced litters with 3+ pups. When pups were P2-P3, mothers were habituated to the test room each day. On P4, pups and dads were removed from the home cage and housed in a new clean cage. Mothers were left for a one-hour isolation period in the test room to standardize the *c-Fos* activity in the POA. After the hour, 3 pups were returned to the home cage with the mother, placed on the opposite side of the cage from the undisturbed home nest. Time began when the mother first investigated the pups. 30 minutes later, pups were removed, and maternal pup-exposed mice were transported for perfusion and tissue harvest. Non-pup-exposed WTs (N=4) and *Magel2*-nulls (N=4) were habituated in the same manner, isolated for one hour in the test room and transported directly from the home cage for tissue harvest.

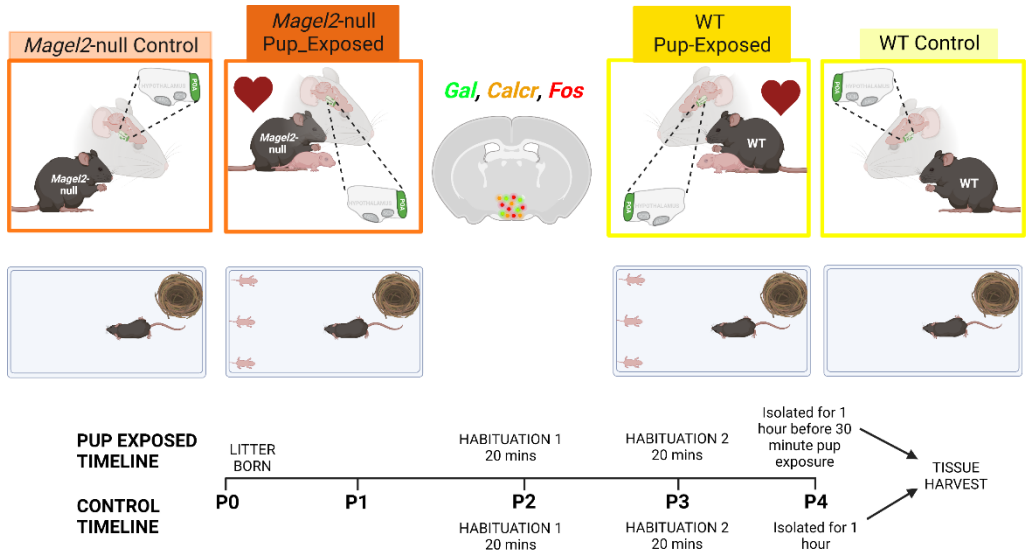


Figure 6.1. Chapter 6 Experimental overview. *Magel2*-null and WT mice were either paired to produce litters and then used as the Pup-Exposed group (N=4 per genotype) or were left undisturbed to act as Controls (N=4 per genotype). Once litters were born, mice were introduced to the test room each day for two days prior to tissue harvest. On tissue harvest day (P4) mothers were isolated from their pups for 1 hour. Mothers then received a pup exposure during which three of her own pups were returned to the opposite side of the cage and she was allowed 30 minutes to interact with them (timed from her first olfactory investigation). After 30 minutes these Pup-Exposed mice were culled. Control animals were culled following a 1-hour isolation period.

6.2.2 FFPE Tissue Preparation

WT and *Magel2*-null mice were transcradially perfused from the home cage with 10% NBF before whole brains were taken. 3 mm POA sections were embedded in paraffin and sectioned at a thickness of 10 µm. Only sections reliably containing high numbers of *Gal/Calcr* cells (identified from sections in Chapter 4) were subsequently analysed.

6.2.3 RNAscope® protocol

Three-plex RNA Scope was performed using RNAscope® Multiplex Fluorescent Reagent Kit v2 (ACD Bio-techne) on these FFPE brain sections. The manufacturer's protocol was followed exactly following the 'standard' pre-treatment guidance (see Section 2.3.4). *Calcr* Probe (Mm-*Calcr*) was paired with fluorophore TSA Vivid™ 570, the *Gal* Probe (Mm-*Gal*-C2) was paired with TSA Vivid™ 520 and the *Fos* Probe (Mm-*Fos*-C3) was paired with TSA Vivid™ 650. All fluorophores were applied at a concentration of 1:1500. Slides were counterstained with DAPI (30 seconds) and mounted.

6.2.4 Image Acquisition and Analysis

Whole brain slides were imaged within one week of mounting using the Zeiss AxioScan Z1 at 20x magnification with the same light intensity/duration settings used for the same probes in each scan (detailed in 2.3.5). Images were analysed with Zen Blue 3.6. Slides were preprocessed and thresholded as detailed in 2.3.6 and Figure 2.5. All the genes (*Gal, Calcr, Fos*) displayed dots and cluster style expression and the clusters were resolved into molecule counts.

6.2.5 Statistics and Figures

RNAscope® image data were analysed in two ways. Firstly, my quantitative analysis compared proportions of positive cells for a particular gene (*Gal/Calcr* – 2+ molecules, *Fos* – 5+ molecules) between areas of the brain and cell types. Variability in sections/POA position was accounted for by normalizing *Fos/Gal/Calcr* positive cell counts per animals to counts per 1000 POA cells. WT and *Magel2*-nulls, pup-exposed and non-pup-exposed were compared using two-way ANOVA. Additionally, the average number of gene molecules was also compared between *Magel2*-null mice and WTs and analysed with two-way ANOVA also. Secondly, a semi-quantitative metric was also calculated by deriving a H-Score (see Section 2.3.8).

6.3 Results

6.3.1 Magel2-null mothers have reduced c-Fos activity in the POA

I first wanted to know whether there would be any significant differences in *c-Fos* activity in the POA of *Magel2*-null and WT mothers (Figure 6.2). When counting the total number of *c-Fos* molecules per 1000 POA cells between my groups, in a two-way ANOVA, I found a main effect of Exposure ($F(1,12) = 25$, $p = 0.0003$) showing that mice exposed to pups had 40.8% more *c-Fos* molecules produced in the POA compared to the controls and I also found a main effect of Genotype ($F(1,12) = 10.79$, $p = 0.007$), showing that *Magel2*-null mice had 19% fewer *c-Fos* molecules in the POA than WTs (Figure 6.3A). This was mirrored when counting the number of *c-Fos*-positive cells per 1000 POA cells (Cells expressing at least 5 molecule of *c-Fos*) with a main effect of Exposure ($F(1,12) = 31.7$, $p = 0.0001$, Pup exposed = 54% more *c-Fos* positive cells) and Genotype ($F(1,12) = 9.95$, $p = 0.008$, *Magel2*-null = 20% fewer POA *c-Fos* positive cells) (Figure 6.3B).

I did not see an interaction effect. This could potentially suggest that *Magel2*-null mice displayed a global *c-Fos* deficit (perhaps due to some neural deficiency) and genotype differences were not based on a specific reduced activity following pup-exposure. However, two things suggested this not to be the case. The size of the difference between the *Magel2*-null control and WT control groups was much smaller than the Pup-exposed groups, suggesting that, although the WT brain may be slightly more *c-Fos* responsive, the pup-exposure was exposing a larger deficit. Secondly, in order to see how another area of the brain was effected, I performed the same *c-Fos* analysis for the cortical areas of the same brain slices showing the POA deficit, (the cortex was generally more active than the POA, expressing over twice as many *c-Fos* molecules per 1000 cells (3363 molecules - Cortex vs. 1603 molecules – POA). Here I no longer found a main effect of Genotype for number of *c-Fos* positive cells ($F(1,12) = 0.04$, $p = 0.84$) (Figure 6.3C) or for number of *c-Fos* molecules ($F(1,12) = 0.01$, $p = 0.92$). This suggested that *c-Fos* was specifically less expressed in the POA of *Magel2-null* mice and not the brain in general.

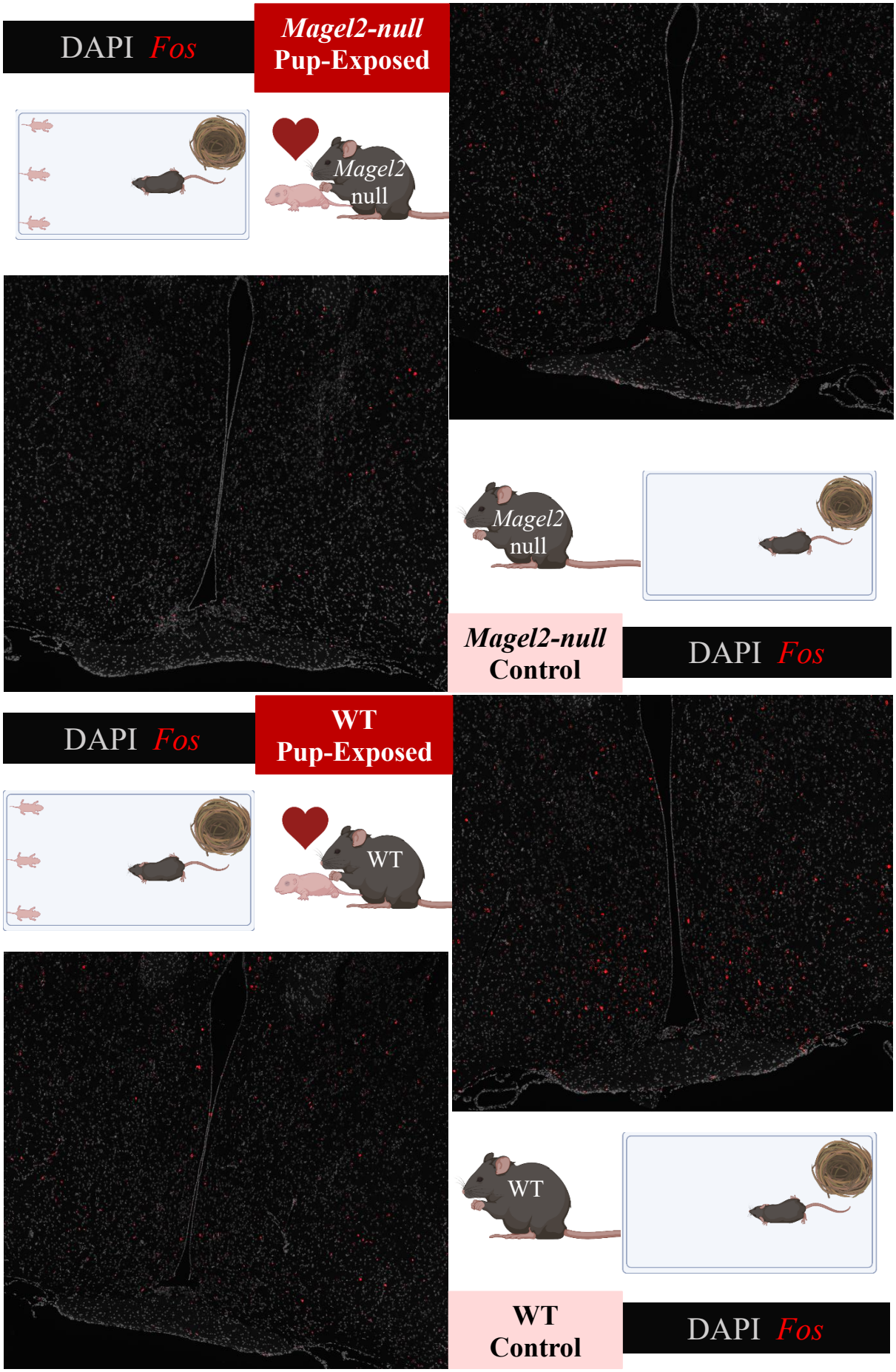

DAPI Fos
Magel2-null
Pup-Exposed
Magel2
null
Magel2
null
Magel2-null
Control
DAPI Fos
DAPI Fos
WT
Pup-Exposed
WT
WT
WT
Control
DAPI Fos

Figure 6.2. (Previous Page) Representative Preoptic Area *c-Fos* Images. *Magel2*-null mice (*Top*) and WT mice (*Bottom*) were either paired to produce litters and then used as the Pup-Exposed group (N=4 per genotype) or were left undisturbed to act as Controls (N=4 per genotype). Images present DAPI (Grey) stained nuclei alongside RNA molecules of *c-Fos* visualised in Red. Sections from Pup-Exposed animals are presented on the right and sections from Controls are presented on the left.

A reduction in molecules of *c-Fos* or in *c-Fos* positive cells in *Magel2*-null mice could be the consequence of structural abnormalities to the *Magel2*-null POA (i.e., a reduction in the number of POA cells globally). Here, I saw no significant genotype difference in the number of cells from the POA (Figure 6.3D; $F(1,14) = 0.067$, $p = 0.8$) with the *Magel2*-null POA comprising an average of 4425 cells per section and the WT POA comprising an average of 4516 cells per section. Nor did I observe any gross structural abnormalities with the POA or the brain sections as a whole (Figure 6.4) nor any differences in POA area ($F(1,14) = 0.37$, $p = 0.55$). This suggests that the reduction in *c-Fos* is due to internal cell activity or very specific cell loss and not a blanket reduction in cell number of the POA.

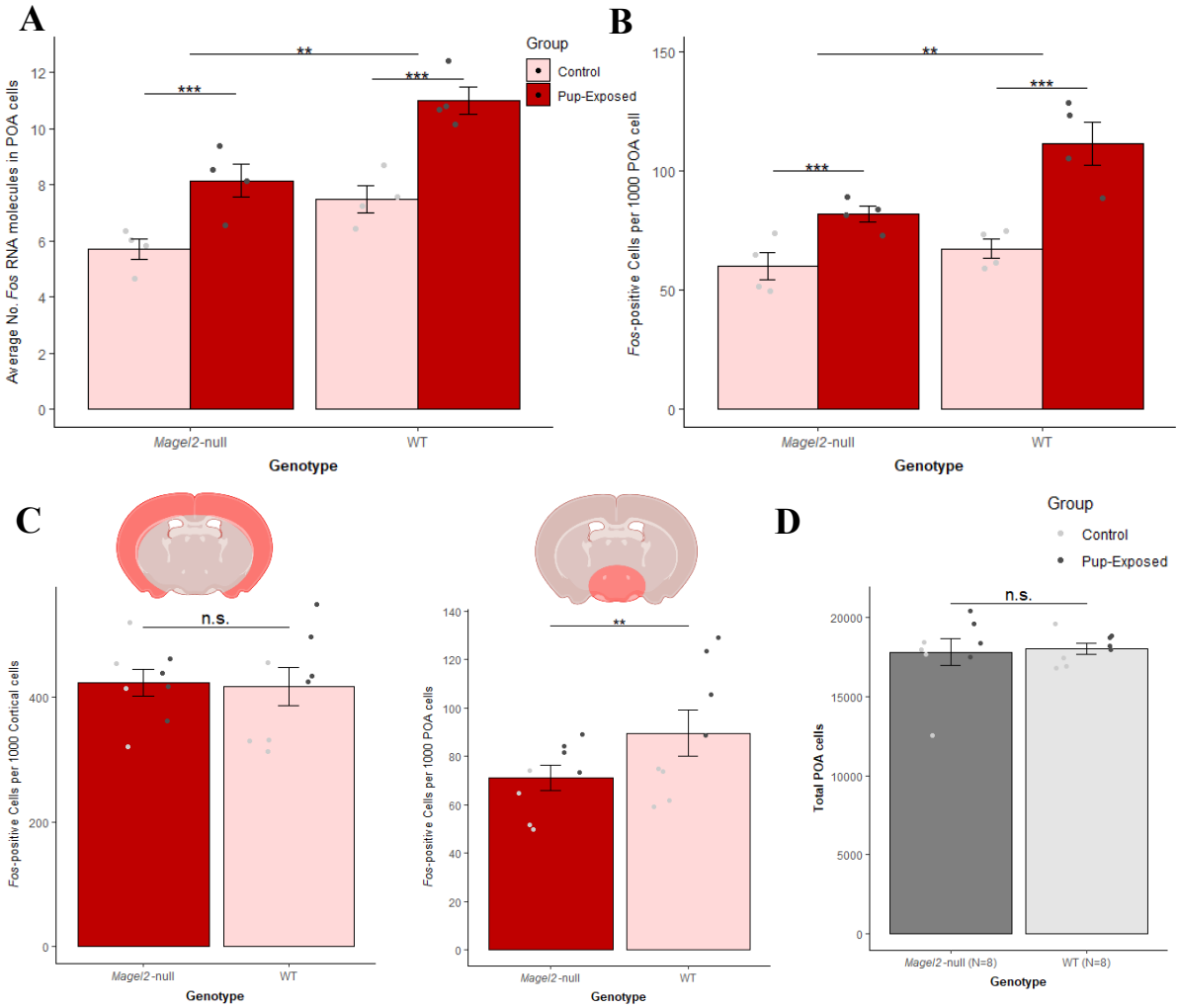


Figure 6.3. (Previous Page) *c-Fos* Expression in the POA of Pup-Exposed and WT mice. (A) Average number of *c-Fos* molecules present in POA cells of *Magel2*-null mice and WT mice either exposed to pups or controls. (B) Number of *c-Fos* positive (5+ molecules) cells per 1000 POA cells of *Magel2*-null mice and WT mice either exposed to pups or controls. (C) Number of *c-Fos* positive (5+ molecules) cells per 1000 Cortical cells (left) and per 1000 POA cells (right) in *Magel2*-null animals vs. WT animals (regardless of exposure). (D) Number of total POA cells per animal used in this study, split by *Magel2*-null and WT (regardless of condition).

6.3.2 Magel2-null mothers have reduced c-Fos activity in Gal-positive and Gal/Calcr-positive cells

The POA of *Magel2*-null mothers had fewer *c-Fos* positive cells upon being exposed to pups, but the POA is comprised of many cell types, and I wanted to know whether at least part of this *c-Fos* reduction would occur in cells positive for the parenting markers *Gal* and *Calcr* (Figure 6.4).

When looking at how many *c-Fos*-positive *Gal/Calcr*-positive cells per 1000 POA cells in the maternal brain (Figure 6.5A), I saw a significant main effect of Exposure ($F(1,12) = 86.6$, $p = 7.76 \times 10^{-7}$) with pup-exposed mice having 129% more *c-Fos*-positive *Gal/Calcr* cells and there was, again, a significant main effect of Genotype ($F(1,12) = 5.52$, $p = 0.037$, *Magel2*-null = 15.3% fewer). There was also no interaction effect, but I did see a more marked decrease between *Magel2*-null and WT in the pup-exposed condition (18.1% decrease in *Magel2*-null) as compared to the control *Magel2*-null and WTs (12% decrease).

Other types of galanin expressing cells have been linked with parenting behaviour and I was interested to know whether this *c-Fos* reduction would occur in galanin cells alone (Figure 6.5B). There was a significant main effect of Exposure ($F(1,12) = 20.022$, $p = 0.0008$) with pup-exposed mice having 56% more *c-Fos* positive *Gal* cells and there was, again, a significant main effect of Genotype ($F(1,12) = 9.405$, $p = 0.0098$) with *Magel2*-null females having a 23.2% reduction in *c-Fos* positive galanin cells of the POA. However, this was not found when looking at *Calcr*-positive cells (i.e., any cell expressing two or more RNA of Calcr (Figure 6.5C) which still showed a main effect of Exposure ($F(1,12) = 41.4$, $p = 3.2 \times 10^{-5}$) but unlike galanin, did not show a main effect of Genotype ($F(1,12) = 4.01$, $p = 0.07$). For all these cell types, when looking at average number of *c-Fos* molecules (Figure 6.5D), there were no effects of Genotype for *Gal/Calcr* ($F(1,12) = 0.93$, $p = 0.35$), *Gal* ($F(1,14) = 4.6$, $p = 0.051$) or *Calcr*

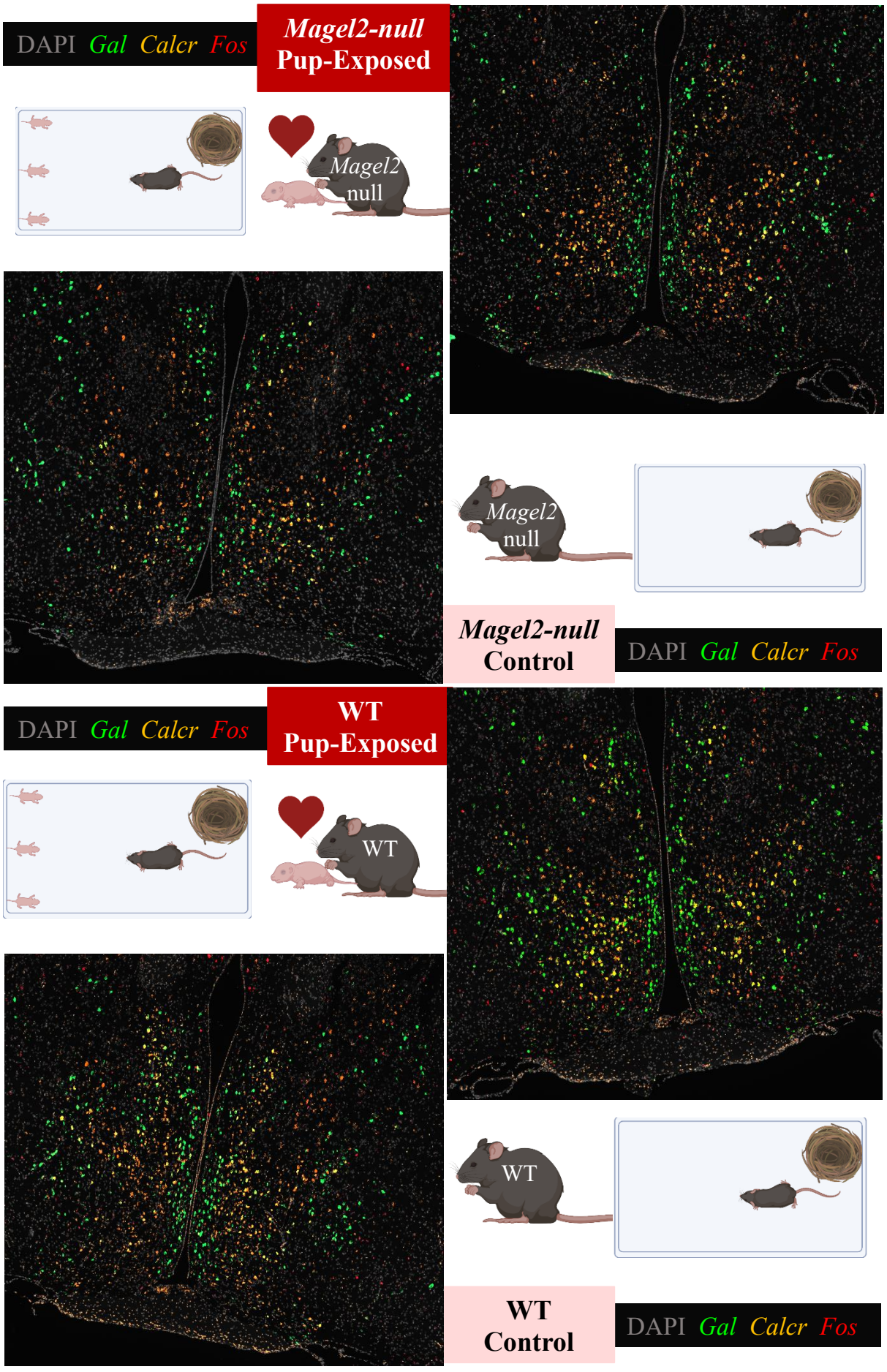

DAPI Gal Calcr Fos
Magel2-null
Pup-Exposed
Magel2
null
Magel2
null
Magel2-null
Control
DAPI Gal Calcr Fos
DAPI Gal Calcr Fos
WT
Pup-Exposed
WT
WT
WT
Control
DAPI Gal Calcr Fos

Figure 6.4. (Previous Page) Representative POA *Gal/Calcr c-Fos* Images. *Magel2*-null mice (*Top*) and WT mice (*Bottom*) were either paired to produce litters and then used as the Pup-Exposed group (N=4 per genotype) or were left undisturbed to act as Controls (N=4 per genotype). Images present DAPI (Grey) stained nuclei alongside RNA molecules of *Gal* (Green), *Calcr* (Orange) and *c-Fos* (Red). Sections from Pup-Exposed animals are presented on the right and sections from Controls are presented on the left.

($F(1,14) = 1.45$, $p = 0.25$) suggesting that the differences between WT and *Magel2*-null in *c-Fos* expression is not due to a reduction in activity levels of cells but rather a reduction in the number of *c-Fos* expressing cells. This was further reinforced by the high similarity of the H-scores and histograms of *c-Fos* expression of *Magel2*-null and WT mice (Figure 6.5E and Appendix Table A9) which suggests that while *Magel2*-null mothers have fewer *c-Fos* positive cells, they have a similar proportion of 'high', 'moderate' and 'low' expressing cells, suggesting that *c-Fos* expression in the *Magel2*-null cells is still manifesting normally.

6.3.3 Magel2-null mothers have a reduction in in Gal-positive and Gal/Calcr-positive cell numbers, explained primarily by a 20% reduction in galanin RNA molecules

Magel2-null mice have previously been associated with reductions in specific cell-types (specifically PVN oxytocin neurons – Schaller et al. (2010)). Here I wanted to know whether this reduction in *c-Fos* activity in the *Magel2*-null mothers would be reflected in a very specific loss of *Gal/Calcr* cells or *Gal* more generally and hence causing the reduction in *c-Fos* activity. In all the following comparisons, there was no main effect of Exposure or interaction effects, indicating that the pup-exposed parents and controls were not exhibiting any fundamental differences in *Gal/Calcr* positive cells or *Gal/Calcr* mRNA expression. When looking at the number of *Gal/Calcr* cells per 1000 POA cells (Figure 6.6A - the POA cell type showing the highest *Magel2* expression in Chapter 4) I found a main effect of Genotype ($F(1,12) = 6.878$, $p = 0.0223$) showing that *Magel2*-null females had 17.6% fewer *Gal/Calcr* positive cells in the POA. When looking at the impact on *Gal* positive cells and *Calcr* positive cells separately (Figure 6.6B), I found that there was a main effect of Genotype for *Gal* positive cells ($F(1,12) = 8.428$, $p = 0.0133$) with *Magel2*-null mice having 18.9% fewer *Gal* positive cells in their POA but there was no main effect for *Calcr* positive cells ($F(1,12) = 0.88$, $p = 0.368$, <10% reduction in *Magel2*-null). This was also replicated when looking purely at the number of RNA molecules of *Gal* in which there was a

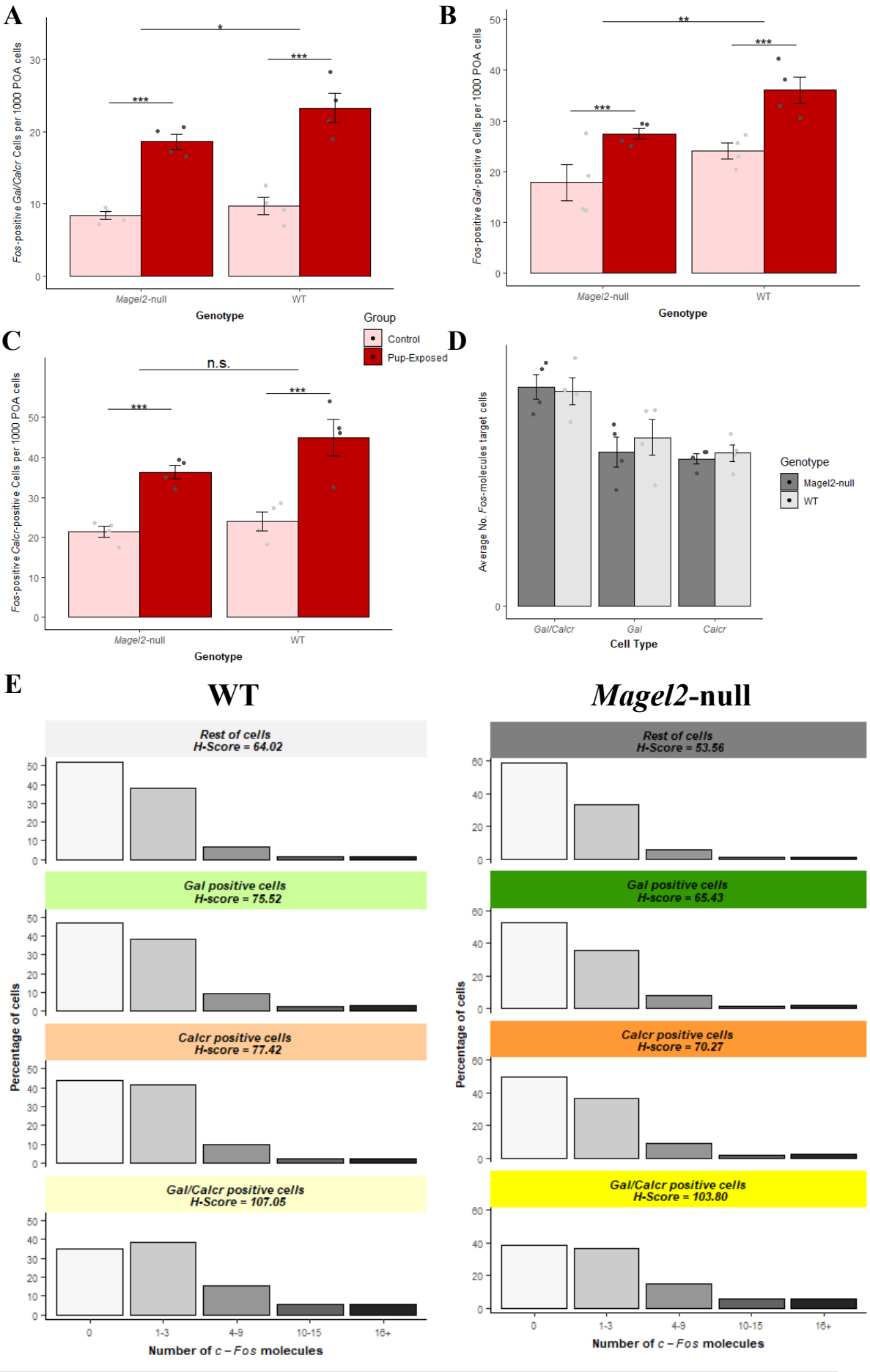

A
Fos-positive Gal/Calcr Cells per 1000 POA cells
Magel2-null
WT
Genotype
B
Fos-positive Gal-positive Cells per 1000 POA cells
Magel2-null
WT
Genotype
C
Group
Control
Pup-Exposed
n.s.
Fos-positive Calcr-positive Cells per 1000 POA cells
Magel2-null
WT
Genotype
D
Average No. Fos-molecules target cells
Gal/Calcr
Gal
Calcr
Cell Type
Genotype
Magel2-null
WT
E
WT
Magel2-null
Rest of cells
H-Score = 64.02
Gal positive cells
H-score = 75.52
Calcr positive cells
H-score = 77.42
Gal/Calcr positive cells
H-Score = 107.05
Rest of cells
H-Score = 53.56
Gal positive cells
H-score = 65.43
Calcr positive cells
H-score = 70.27
Gal/Calcr positive cells
H-Score = 103.80
Percentage of cells
0
1-3
4-9
10-15
16+
Number of c-Fos molecules

Figure 6.5. (Previous Page) *c-Fos* Expression in the POA of Pup-Exposed and WT mice. (A) Number of *c-Fos* positive (5+ molecules) cells also expressing *Gal* and *Calcr* (2+ molecules) per 1000 POA cells of *Magel2*-null mice and WT mice either exposed to pups or controls. (B) Number of *c-Fos* positive (5+ molecules) cells also expressing *Gal* (2+ molecules) per 1000 POA cells of *Magel2*-null mice and WT mice either exposed to pups or controls. (C) Number of *c-Fos* positive (5+ molecules) cells also expressing *Calcr* (2+ molecules) per 1000 POA cells of *Magel2*-null mice and WT mice either exposed to pups or controls. (D) Average number of *c-Fos* molecules in POA cell also expressing *Gal/Calcr*, *Gal* and *Calcr*, split by *Magel2*-null and WT (regardless of condition). (E) Histogram showing the percentage of cells with particular number of *c-Fos* molecules in *Gal/Calcr* positive cells, *Gal* positive cells, *Calcr* positive cells and the Rest of the POA cells. A larger percentage of *Gal/Calcr* cells expressed 4+ *Fos* RNA molecules explaining why this cell type has the highest H-Score. H-scores and histograms for all the cell type comparisons are remarkably similar when comparing *Magel2*-null (right) vs. WT mice (left).

main effect of Genotype ($F(1,12) = 5.708$, $p = 0.034$, 19.3% fewer in *Magel2*-null) and RNA molecules of *Calcr* in which there was no main effect ($F(1,12) = 1.08$, $p = 0.319$, 11.3% fewer in *Magel2*-null) (Figure 6.6C).

Finally, I examined whether controlling for galanin and/or *Gal/Calcr* positive cell number would remove the *c-Fos* differences I saw in these cell types. When assessing how many *c-Fos* positive *Gal/Calcr* cells each animal had per 1000 *Gal/Calcr* cells (Figure 6.6D), I still saw a main effect of Exposure ($F(1,12) = 89.5$, $p = 6.5x10\text{-}7$) which indicates that when normalizing the number of *Gal/Calcr* cells between pup-exposed and control animals, pup-exposed animals had more *c-Fos* positive cells. However, I no longer found a significant main effect of Genotype ($F(1,12) = 0.041$, $p = 0.844$) which was also the case when looking at *c-Fos* positive *Gal* cells per 1000 *Gal* cells identified (Figure 6.6D; Group - $F(1,12) = 34.3$, $p = 7.73x10\text{-}5$; Genotype – $F(1,12) = 0.99$, $p = 0.339$). *c-Fos* activity when normalizing for *Calcr* expressing cells (Figure 6.6) did not change the results maintaining a significant effect of Exposure was still significant ($F(1,12) = 43.7$, $p = 2.49x10^{-5}$) showing that *Calcr* cells in the POA were more active in pup-exposed mothers but still not showing a Genotype effect ($F(1,12) = 0.02$, $p = 0.88$).

6.4 Discussion

This chapter has shown that in the maternal cohort, the parenting deficit seen in *Magel2*-null mothers is associated with a reduction in *c-Fos* positive POA *Gal/Calcr* neurons that regulate parenting. This is predominantly explained by a reduction in POA galanin positive cells and considering the total number of cells in the POA is unchanged between genotypes, this is likely explained by a reduction in galanin expression in existing cells rather than a true

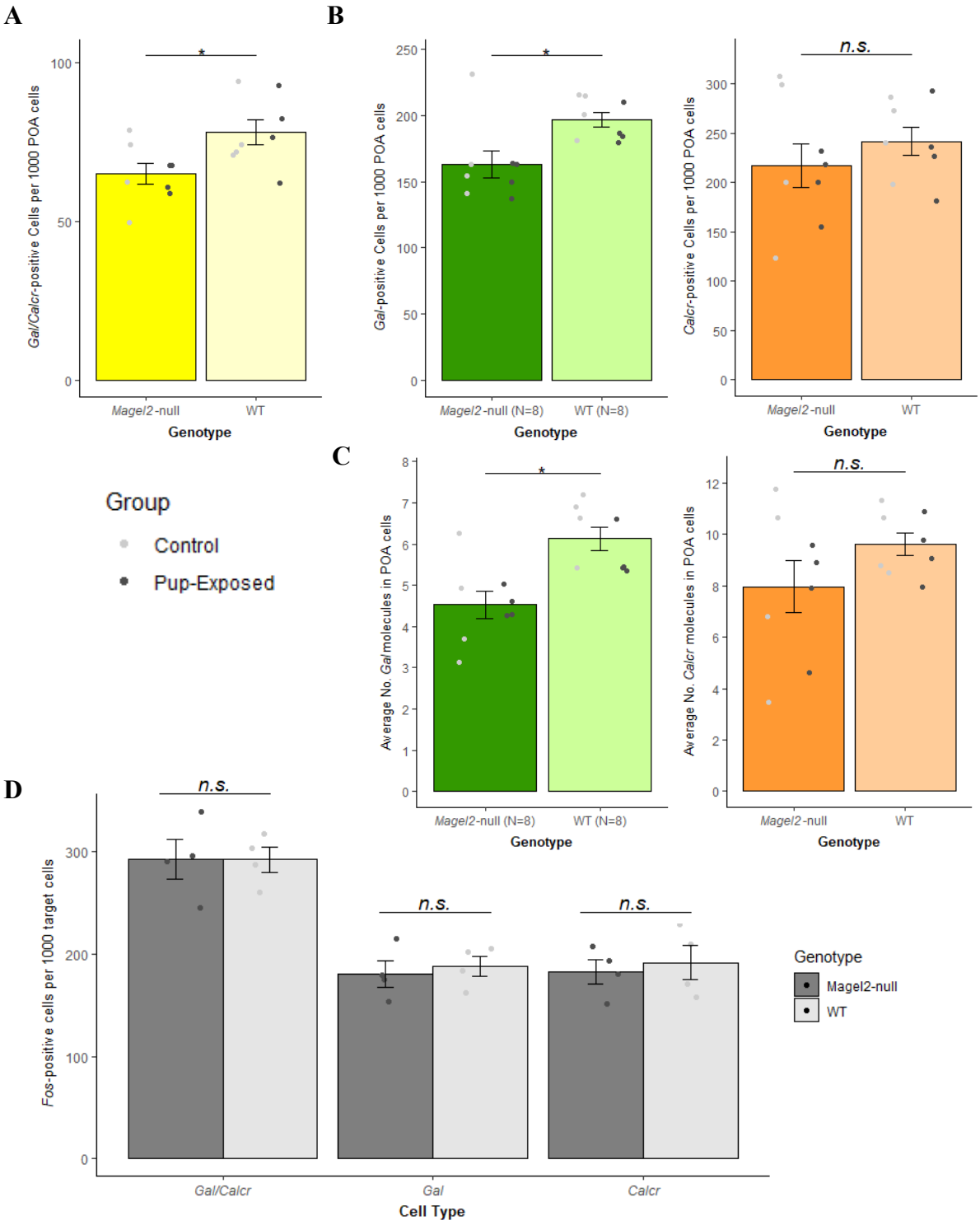


Figure 6.6. *Gal* and *Calcr* positive cells in the POA of Pup-Exposed and WT mice. (A) Number of *Gal*/*Calcr* positive (2+ molecules each) cells per 1000 POA cells of *Magel2*-null mice and WT (regardless of exposure). (B) Number of *Gal* (Left) and *Calcr* (Right) positive (2+ molecules) cells per 1000 POA cells of *Magel2*-null mice and WT mice (regardless of exposure). (C) Average number of *Gal* (Left) and *Calcr* (Right) molecules present in POA cells of Magel2-null animals vs. WT animals (regardless of exposure). (D Number of (*Gal/Calcr*, *Gal*, *Calcr*) cells also registering as *C-Fos* positive (5+ molecules) per 1000 of the respective cell (*Gal/Calcr*, *Gal*, *Calcr*). Significant differences seen in Figure 6.5A & B are no longer present here when normalising for *Gal* and *Gal/Calcr* cell number.

reduction in cell numbers i.e., cells that would normally be *Gal*-positive in WT mice are now no longer considered so in *Magel2*-null mothers as the cells do not express galanin. Mothers are the primary care givers in mammals and the primary model used when assessing parenting behaviour. *Magel2*-null mothers also showed the least dramatic insult to parenting in Chapter 5, hence having found molecular alterations in mothers would strongly suggest I would find equal if not more substantial deficits in fathers and virgin females.

A larger number of *c-Fos* positive cells in the POA of the pup-exposed mothers showed that activity in this area is high correlated with enacting parenting behaviour when exposed to pups (Numan and Numan, 1994, Mathieson et al., 2002, Wu et al., 2014, Moffitt et al., 2018). That the *Magel2*-null females had less *c-Fos* activity in the POA is highly suggestive that dysfunction in this area is leading to the reduction in parental behaviour seen in Chapter 5. This is further supported by the lack of a genotype difference in *c-Fos* positive cells of the cortex, suggesting that the *Magel2*-null *c-Fos* reduction is region specific and not brain-wide. The finding that this reduction in *c-Fos* also occurs in the very specific neuron type (*Gal/Calcr*), which was been intimately associated with parenting behaviour (Kohl and Dulac, 2018, Moffitt et al., 2018), strengthens the suggestion that parenting is affected in *Magel2*-null mice due to POA^{Gal} dysfunction.

Several ideas could explain this *c-Fos* reduction. One is that there is a structural insult to the POA of *Magel2*-null mice which reduces the number of cells and hence reduces the number of *c-Fos* positive cells indirectly. In response to this idea, I found there were no physical differences between the POA's of *Magel2*-null and WT mice. This included no structural abnormalities, no differences in area of POA, and no differences in cell number within the POA. This aligns with Mercer et al. (2009) who found that even though the hypothalamus has the highest expression of *Magel2,* detailed MRI data did not find any structural differences in the hypothalamus whereas there were regional size differences in other brain areas such as the amygdala.

Another idea would be that the POA of *Magel2*-null mice is structurally unaltered, but intricate neural mechanisms cause the cells to fire less hence a reduced *c-Fos* expression. This is unlikely to be the case because when looking at the average *c-Fos* expression in the *Gal/Calcr* cells as well as comparing the H-scores, indicated that these cells were expressing *c-Fos* at similar levels and the activity of these cells was not altered. One final idea would be that *Magel2*-null mice manifest insults to specific cell populations (i.e., the *Gal/Calcr* cells)

and despite no overall cell loss in the *Magel2*-null mice, I found they had significantly fewer *Gal*-positive and *Gal/Calcr* positive cells in the POA which, when adjusted for, completely explained the *c-Fos* difference between genotypes in these cells. Hence, the most parsimonious explanation for reduced *c-Fos* activity and the reason *Magel2*-null mice had less *c-Fos* activity in the POA^{Gal} is that *Magel2*-null mutants have fewer cells making up the POA^{Gal} population. Either this cellular loss is so specific that this is not detected when it comes to total POA cells or, the more likely scenario, is that the cell counts in the POA of *Magel2*-null mice are unchanged, but their expression rates of galanin are reduced, meaning fewer of the cells in the POA can be considered galanin positive, because they are not expressing *Gal* at the level normally seen in the WT. This would then suggest that deleting *Magel2* impairs the galanin neuropeptide system which may affect the performance of the POA^{Gal} Hub and hence parenting behaviour.

6.4.1 Caveats

Primary amongst the caveats is that this experiment has only been carried out using female mice/mothers despite having found a behavioural deficit in mothers, fathers and virgins. Finding the same galanin expression deficit in males/fathers would strengthen the work of this chapter substantially and potentially strengthen the idea that *Magel2*'s action is on the POA^{Gal} Hub.

This is also not an exhaustive account of the mechanisms by which *Magel2* could be affecting parenting behaviour and the POA. It is unlikely that the reduction in *c-Fos* expression in *Gal* positive cells alone explains the reduction in *c-Fos* activity more generally in the *Magel2*-null mice. This means there are other POA cell types which are also less active in *Magel2*-null mice upon exposure to pups but are not considered here without the appropriate probes to examine these. Furthermore, other mechanisms such as pituitary action (especially relevant given that the proteomic work of Chen et al. (2020) in the pituitary found that prolactin was one of the primary downregulated proteins in *Magel2*-null mice), dopamine, serotonin and oxytocin action are not considered here and could be having more substantial impacts on parenting behaviour in *Magel2*-null mice than the *Gal* expression in the POA. It could be the case that the reduction in *c-Fos* activity in parenting neurons of the POA is a consequence rather than a cause of some other mechanism, although the reduced galanin expression is less likely to be an indirect consequence.

6.4.2 Summary of Findings

- *Magel2*-null mothers had 20% fewer *c-Fos* positive cells in the POA when exposed to pups compared to WTs. This difference was not observed in the cortex
- *Magel2*-null mothers have 18.1% fewer *c-Fos* positive *Gal/Calcr* cells in the POA
- *Magel2*-null mothers have 19.3% less *Gal* expression in the POA leading to a reduced number of *Gal*-positive (18.9% less) and *Gal/Calcr*-positive cells (17.6% less) which naturally results in a reduced *c-Fos* response from these cells.

7 General Discussion

7.1 Overview

Some of the first phenotyping studies on imprinted gene mouse models found maternal care to be drastically affected (Lefebvre et al., 1998, Li et al., 1999). Since these studies in the late 1990's, maternal care has widely been seen as one of the phenotypic outcomes of disrupting imprinted genes and is consistently given as the primary evidence to support alternative theories for the origin of genomic imprinting (Curley et al., 2004). However, it has taken almost 20 years to identify another imprinted gene with parental care implications (Stohn et al., 2018) and in the meantime, doubts have been cast on the findings from the original two models (Denizot et al., 2016, Anunciado-Koza et al., 2022). Furthermore, recent years have seen a radical advance in the understanding of the neural circuitry co-ordinating parenting in mice, with it being revealed that a small population of neurons in the POA expressing galanin (Kohl et al., 2018, Wu et al., 2014) alongside other markers such as *Calcr,* (Moffitt et al., 2018) were sufficient and necessary to produce parenting behaviour in mice.

As it stands, parenting is still seen as a phenotypic consequence of genomic imprinting in mammals, yet this association still relies predominantly on experimental findings of only a handful of genes. One might expect that if imprinted genes were associated with parenting, they may act on the POA parenting circuitry, however, to date there has been no direct research exploring this area, and so the role of imprinted genes here is unknown. Based on this, the over-arching aim of this thesis was to assess whether imprinted genes would show enriched expression in the POA parenting neurons, and, if they did, whether one could use this data to identify new genes associated with parenting. Finally, I aimed to then confirm that these genes could show an impact on the POA parenting circuitry.

7.2 Main Findings

The main findings from each experimental chapter have been discussed in depth within the respective chapters, a brief summary of the three key findings from each chapter are displayed in Table 7.1. The following discussion will consider the implications of each chapter as well as the wider implication and future directions of this work as whole.

Table 7.1. Summary of key findings from each experimental chapter.

Chapter	Key Findings
Chapter 3: *Imprinted Gene Single-Cell RNAseq Enrichment*	• Imprinted gene expression was shown to be convergently enriched in specific hypothalamic neurons of the POA, ARC and SCN as well as the monoaminergic system of the brain. • Relating to parenting behaviour, imprinted genes showed enrichment in the prolactin, serotonin and dopamine systems, and more fundamentally, the galanin neurons of the POA which directly govern parenting behaviour. • 21 imprinted gene candidates were put forward as having a potential relationship with parenting based on their expression profile in POA^{Gal}, 3 of which have already shown parenting deficits in mouse models.
Chapter 4: *Magel2 Preoptic Area Expression*	• *Magel2* was shown to be highly expressed in the rostral hypothalamus and lateral septum and hotspots of *Magel2* expression in the rostral hypothalamus including the cells of the suprachiasmatic nucleus, paraventricular nucleus and the cells of the preoptic area. • Using probes in the PVN, *Magel2* was shown to be expressed highly in the *Avp* positive neurons and moderately in *Oxt* positive neurons • The strongest signal for *Magel2* was in the POA^{Gal} hub, specifically *Magel2* showed its highest expression and highest number of positive cells in the *Gal/Th* and *Gal/Calcr* cells of the POA.
Chapter 5: *Magel2-null Parenting Assessment*	• Mothers, fathers and virgin female mice null for *Magel2*, showed deficits in retrieval behaviour (fathers and virgins) and nest building (mothers and fathers) which could not be explained by motility or olfactory issues. • Mothers, fathers and virgin female mice null for *Magel2* also showed a reduction in the proportion of their time they spent engaged in pup-directed behaviour, which was mirrored by their lack of pup-preference during the Three Chambers test. • *Magel2*-null pups were not less preferentially retrieved and did not impact the retrieval behaviour of their parents.
Chapter 6: *Magel2-null Preoptic Area Phenotyping*	• *Magel2*-null mothers had 20% fewer *c-Fos* positive cells in the POA when exposed to pups compared to WTs. This difference was not observed in the cortex. • *Magel2*-null mothers had 18.1% fewer *c-Fos* positive *Gal/Calcr* cells in the POA • *Magel2*-null mothers had 19.3% less *Gal* expression in the POA leading to a reduced number of *Gal*-positive (18.9% less) and *Gal/Calcr*-positive cells (17.6% less) which naturally resulted in a reduced *c-Fos* response from these cells.

7.2.1 Implications for Bioinformatic Approach

Using publicly available genetic sequencing data, the work in Chapter 3 documented an accessible approach to behavioural genetics. It principally relied on using the gene expression data as a predictor of function, which in and of itself is not novel. However, a new aspect developed in this thesis was the use of a variety of sc-RNA seq datasets to pinpoint the expression of imprinted genes to cellular specificity. Identifying cells in which a gene is highly expressed should be a good hallmark for a cell type in which the gene is functionally active. Going beyond the expression of single genes, as I did in Chapter 3, and asking where the expression of a whole gene set is enriched, fundamentally asks whether a gene set like the imprinted genes show an enrichment pattern and hence some collective overarching functional focus. My approach provided a platform to identify where imprinted genes are expressed and the functional outcome they might have, which escapes the over-reliance on well-studied examples.

The rest of this thesis, Chapters 4-6, acted as a validation of this approach by analysing one candidate gene following a format explicitly laid out at the end of Chapter 3 - Step 1 (Chapter 4) – Validate candidate gene enrichment in target neuronal population, Step 2 (Chapter 5) – Confirm suspected behavioural deficit in mouse model, Step 3 (Chapter 6) – Investigate neuronal population phenotype and activity within mouse model that demonstrates behavioural deficit. Since my gene candidate (*Magel2)* showed both behavioural and the associated neuronal deficits, this thesis as a whole serves as an advert for the value of using expression data to identify functional consequences of genes that have been previously unaccounted. By following the three-step system above, one can validate and confirm the functional deficit as well as attempting to tie that deficit back to the cell types/circuitry which informed you of the function to begin with.

It is worth noting that whether a gene is ‘expressed’ in a cell is no guarantee that it will have any functional consequence, and indeed many genes appear to be transcribed in trace amounts without functional consequence (Mercer et al., 2012, Mortazavi et al., 2008, Hart et al., 2013) but considering that the majority of the genome in any individual cell is heavily methylated and expression of that gene is repressed (Yong et al., 2016), one would expect a gene to be expressed in more than trace amounts only when that gene product is providing some functional role. Without mRNAs from that gene transcribed in that cell, it would not be possible for that gene to influence the functioning of that cell to produce a different

functional state. As Chapters 4-6 have showed, *Magel2* expression here was heavily predictive of a functional consequence.

Though the bioinformatic approach laid out in this thesis has a lot of value, it also suffers from several caveats as detailed in Section 3.4.1. When it comes to imprinted genes, the primary issue is the lack of confirmation that the genes expressed in target cells are actually monoallelically expressed. It was not possible to confirm this with my current approach, although the hypothetical scenario in which genes which are monoallelically expressed in all previous experiments would be biallelic in just the enriched cell types identified in Chapter 3 is highly unlikely, it cannot be ruled out. It is important to remember that this approach by no means would be able to identify all populations that a gene set was active in, it only performs the equivalent of taking the highest peak of the collective enrichment signal, it would completely ignore genes with individualistic effects. Additionally, and very relevant for imprinted genes, any genes altering developmental processes that go on to shape adult tissues and cells would potentially not be detected when looking at enrichment of those genes in adult cells (which is no longer their site of action). Furthermore, cells and subtissue are interconnected networks of a variety of different cell types and gene expression in one cell type could have functional consequences for another, without the corresponding gene expression change in the affected cell (e.g., the role of astrocytes for neuronal synaptic plasticity (Allen and Barres, 2005, De Pittà et al., 2016). These various interconnected and time-point specific mechanisms are beyond the sensitivity of this methodology. However, in summary, I think the approach of surveying the expression of target genes to inform the experimenter of which behavioural or neurological phenotypes to assess is a more useful, targeted, cost effective and efficient approach than broad phenotyping.

7.2.2 Implications for Magel2

Magel2 is one of several Prader-Willi syndrome genes, although it also results in its own associated syndrome (Schaaf-Yang) when impacted in humans (Fountain and Schaaf, 2016). It is one of the heavily studied imprinted genes both for its clinical implications for feeding behaviour, growth and cognitive development (Tacer and Potts, 2017, Kamaludin et al., 2016, Fountain and Schaaf, 2016, Fountain and Schaaf, 2015, Bischof et al., 2007, Lee et al., 2000) but also as a risk-gene for autism (Fountain et al., 2017a, Schaaf et al., 2013). For the latter, recent animal work has focused on the oxytocin system in *Magel2* and offspring communication with mothers via USVs (Bosque Ortiz et al., 2022, Da Prato et al., 2022).

This thesis has suggested, for the first time, that disruption of *Magel2* is associated with parenting deficits. This is particularly interesting given that *Magel2* shares several hallmarks with *Peg3*, which has consistently been linked to parenting behaviours as the most studied imprinted gene in the context of mammalian parenting. These shared features include, growth deficits (Bischof et al., 2007, Curley et al., 2005, Li et al., 1999), impacts on the PVN oxytocin system (Li et al., 1999, Meziane et al., 2015, Schaller et al., 2010), USV deficits, (Bosque Ortiz et al., 2022, Da Prato et al., 2022, McNamara et al., 2018a), and deficits in suckling and lactation (Curley et al., 2004, Schaller et al., 2010). Interestingly, though disruption of both *Peg3* and *Magel2* show detrimental effects on litter survival rates (Schaller et al., 2010, Li et al., 1999), the effect has never been attributed to maternal performance in the *Magel2* model, most likely due to the survival rate not being as poor as the original *Peg3* model and the general clinical focus of the research around *Magel2*.

Naturally, one might ask, if *Magel2* has been shown to impact parenting, does this contribute anything to our understanding of the diseases *Magel2* is associated with. In short, I think very little. Prader-Willi and Schaaf Yang are diseases of development, with cognitive and behaviour difficulties that persist into later life (Fountain and Schaaf, 2016, Cassidy et al., 2000). Patients do not tend to reproduce; with research to suggest that males are infertile, and in line with this, there have been no reported successful pregnancies involving males who suffer from any of the human conditions associated with *Magel2* (Heksch et al., 2017, Gross-Tsur et al., 2012). In females, only four, successful pregnancies have been reported, though these were aided with medication. Of these pregnancies, two resulted in healthy, unaffected offspring, whilst two resulted in offspring inheriting the PWS deletion but developing Angelman's syndrome due to the maternal inheritance (Schulze et al., 2001, Åkefeldt et al., 1999). No comment was made on whether the mothers exhibited reduced interest in their infants or reduced maternal motivation. If *Magel2* did not affect reproductive capacity then a specific deficit of parenting potentially might manifest in later life, but I think the lack of this symptomology in Schaaf-Yang and Prader-Willi does not have any implications for the legitimacy of *Magel2* effect on parenting in mice.

Regarding social phenotypes, in mice, *Magel2*-null mice have been shown to demonstrate a reduced interest in social novelty but this did not manifest as any baseline social deficits (Fountain et al., 2017b, Mercer et al., 2009). In humans, as previously highlighted, there are links between the truncating mutations in *Magel2* and autism (Schaaf et al., 2013, Meziane et al., 2015, Fountain et al., 2017a). One might argue that reduced motivation to interact with

pups (i.e., a parenting deficit) could be mechanistically linked with these social deficits since parenting is a social behaviour for both parent and offspring. However, since *Magel2*-null mice do not show a general reduction in social motivation (Mercer et al., 2009) and the pups used in my assessments were not novel social cues, I think the parenting deficit may be more of a specific rather than a general loss of social interest. The work of this thesis more supports the idea that *Magel2* is performing one action in offspring (early life social engagement with mother) and another action in the parents (motivation to provide the care) in the form of co-adaption (Wolf and Hager, 2006) similar to the manifestations of *Peg3* (Curley et al., 2004).

The primary implication that this thesis has for research into *Magel2* (other than providing a novel behavioural phenotype) is providing further support and another example for what likely is *Magel2*'s major gene action in the brain, endosomal regulation and neuropeptide trafficking. Chen et al. (2020) found that the *Magel2* KO mouse model caused secretory granule abundance and altered neuropeptide production, which was mirrored in two PWS patient cell models. Specifically, *Magel2* regulates a WASH complex which prevents aberrant lysosomal degradation of SG neuropeptides, and hence deletion of *Magel2* impairs proper endosomal trafficking and recycling, resulting in a reduction of SGs and reduction in neuropeptide production. This would result in reduced neuropeptide release in the brain. The proteomic evidence from this group on *Magel2*-null murine hypothalamus showed that galanin (Alongside other neuropeptides such as: *Oxt, Avp, Sst* & *Trh*) is one of the major down-regulated proteins (Chen et al., 2020). Here, I was measuring galanin RNA levels which is not the same as the protein levels, but it is not inconceivable that the improper management of GAL protein would have feedback effects on *Gal* RNA levels in neurons, reflecting a broader functional deficit in the galanin neuropeptide system. Whether the phenotype observed in this thesis is the product of purely galanin dysfunction or some wider insult to several neuropeptide systems would need to be confirmed (especially given the downregulation of PRL protein in the murine pituitary - Chen et al. (2020)), but regardless, this thesis suggests that *Magel2* disruption is sufficient to alter parenting behaviour at least in part, due to aberrant galanin neuron activity.

7.2.3 Implications for Parenting Behaviour

If the enrichment of imprinted genes in the POA^{Gal} is substantiated by future research this will have implications for parenting behaviour more widely. Galanin has been recognized as a marker for parenting neurons for almost a decade (Wu et al., 2014) but whether the galanin

gene/protein actually plays a role in parenting behaviour or is just an indirect marker gene is poorly understood. If *Magel2*'s modulation of galanin levels in the POA is further substantiated and shown to causally impact parenting, along with the observed similar reduction in galanin protein in *Magel2*-null mice (Chen et al., 2020) this would then suggest that a partial loss of galanin expression would be enough to impact parenting behaviour. Additionally, it would suggest that galanin is performing a functional role for regulating these behaviours rather than just acting as a marker.

Furthermore, this thesis highlights the value of using groups other than mothers for assessing parenting behaviour. Fathers and virgin females rely on the same POA circuitry to produce parenting behaviour yet rely on different mechanisms to activate this circuitry and produce parental care (Rogers and Bales, 2019). Testing mothers, fathers and virgins expands the power of parenting behaviour analyses substantially. Notably, if deficits were observed in all groups, this may suggest one core or multiple overlapping mechanisms, while deficits in one group may suggest something specific e.g., to mothers (pregnancy hormones/placenta - Bridges (2016)) or to fathers (vomeronasal circuits - Liu et al. (2013)). The consistent insult to parenting I saw in Chapter 5 further strengthens the POA^{Gal} hub as a mechanism for the parenting behaviours that only assessing behaviours in the mother would not provide.

7.2.4 Implications for Genomic Imprinting

As highlighted previously, in recent years the initial findings surrounding the relationship between imprinted genes and parenting behaviour have been challenged. More recent models of *Peg3* and *Mest*, the two original maternal-care-associated imprinted genes, have failed to produce maternal care deficits in sophisticated models (Denizot et al., 2016, Anunciado-Koza et al., 2022) that lack foreign reporter material, such as LacZ, that may have had unintended neuronal, and hence behavioural, effects (Peck et al., 2021). As a consequence, these two genes cannot be considered as definitive examples of this relationship. Recent years have seen three new IGs characterized with parenting deficits (Stohn et al., 2018, Keshavarz and Tautz, 2021) but, the question remains whether it was correct to say that imprinted genes are associated with parenting.

In my opinion, the primary implication of this thesis has been the reinforcement of the relationship between imprinted genes and parenting, which was achieved in a number of ways. Firstly, through the demonstration that the POA^{Gal} is one of only a few cell-type-specific hotspots of imprinted genes. This cell type has been shown to specifically regulate

parenting behaviour (Kohl and Dulac, 2018, Kohl et al., 2018) and so, if in future research the other imprinted genes that are expressed here are also shown demonstrate functional consequence, there is a high likelihood they will be associated with parenting also. Secondly, I have identified a new imprinted gene that impacts parenting behaviour in mice, *Magel2*. This thesis has therefore added one more gene to a list that may now hopefully expand given the identification of imprinted gene targets generated in Chapter 3 of this thesis. Finally, I also found additional sites of imprinted gene enrichment with relevance for parenting such as dopamine, serotonin and prolactin. Overall, I have shown that imprinted genes express in the exact manner one would hope if parenting was an important function they regulate.

On a separate note, one of the more interesting insights from this thesis is the suggestion of an imprinted gene mechanism to regulate parenting behaviour and other hypothalamic behaviours mediated by the neuroendocrine system. Neuroendocrine cells were the principle enrichment I saw for imprinted genes in Chapter 3 where a substantial number of imprinted genes in the brain showed exaggerated expression with these neurons which are marked out by markers of neuropeptides and endosomal trafficking (e.g., *Baiap3* - Zhang et al. (2017), Ximerakis et al. (2019)). This general observation of IG enrichment in neuroendocrine cells aligns well with the suggestion that *Magel2* may be having its effect on parenting through trafficking and recycling of galanin in the POA. However, it should be noted that, at this stage, this is a purely hypothetical suggestion as I have not performed any experiments confirming this. This caveat notwithstanding, it is implied that this may be the underlying mechanism given the lack of structural changes and cell loss in the POA yet the model still shows a loss of *Gal* expression and a loss of galanin protein observed previously (Chen et al., 2020).

Imprinted genes are enriched in the adult brain suggesting that they should have functional consequence here (See Chapter 3). Though some IGs have been shown to alter synaptic processing (see Section 1.1.5) many do not have a mechanism for how they alter brain functioning post-natally. If imprinted genes were shown to regulate endosomal process more widely it would explain several of the enrichments identified in Chapter 3 (e.g., *Avp* cells, *Agrp* cells) including the galanin neurons of the POA and potentially their relationship with parenting behaviour. By regulating the amount of galanin RNA/protein, *Magel2* could effectively tone up or tone down the parental responsiveness of animals. This can be done without needing to influence the development of these neurons, just their neuroendocrine ability at a mature state. If multiple other genes can be shown to have similar phenotypes

when it comes to neuropeptide regulation it would suggest this is a process by which these late-to-evolve genes could regulate brain function and shift behaviour into a particular direction. Whether this mechanism is generic or specific is an important thing to test. Imprinted genes would unlikely evolve to modulate all neuroendocrine aspects of the brain, the specificity of the enrichment in the brain suggests that not all neuroendocrine locales are enriched and genes like *Magel2* seem to regulate endosomal activity in specific pathways which, if not an anomaly, suggests that imprinted genes are targeting specific pathways rather than generically involved in neuroendocrine regulation.

The final implication this thesis has for imprinted genes is regarding their evolution. Section 1.1.2 summarized the major theories for the origin of genomic imprinting. The effect of imprinted genes on maternal care has been difficult to square with conflict theory although attempts have been made. With the recent failed attempts to replicate the parenting deficits of *Mest* and *Peg3* (Denizot et al., 2016, Anunciado-Koza et al., 2022), conflict theory might no longer need to account for parenting if it turns out that LacZ activity is causing the parenting issues in these models. By strengthening the relationship between imprinted genes and parenting behaviour, this thesis has added to the body of literature which suggests that conflict theory, although very powerful as a theory, may not be capable of accounting for everything. This is further highlighted with the finding that deleting *Magel2* impairs the parenting behaviour of virgin females and, particularly, fathers. The attempts to square parenting with conflict theory have relied on the differential desires that the parents may have for their daughters when it came to parenting, there has been no consideration of fathers/sons (Haig, 2014, Wilkins and Haig, 2003, Haig, 1999). I think that two explanations to square paternal behaviour with imprinting exist, the first is that genomic imprinting did evolve in relation to maternal care, and the enhancement of this care in mothers was enhanced in fathers and virgins as a biproduct. The second idea is that, although genomic imprinting has evolved in the realms of conflict, these genes have persisted post-natally, to exhibit monoallelic expression to enhance parenting to adapt to the new demands created from the evolution of Eutherian mammals which applies to mothers, fathers and virgin animals.

Genomic imprinting is not the same amongst all mammals. Eutherian mammals have several hundred confirmed imprinted genes while the next closest group of mammals, the marsupials have only been confirmed to have a handful, with many of the eutherian imprinted genes either not having evolved or being expressed biallelically (Renfree et al., 2009, Stringer et al., 2014, Suzuki et al., 2011a, Suzuki et al., 2011b, Rapkins et al., 2006, Edwards et al., 2007).

This suggests that genomic imprinting evolved quite radically following the Eutherian-Marsupial split (Figure 7.1A), an idea further reinforced given that imprinted genes have not been found in the only other groups of mammals, the monotremes (Killian et al., 2001, Killian et al., 2000). As highlighted both in the introduction and in Figure 7.1 a number of unique properties mark the evolution from the therian ancestor to the first eutherians. Notably, the placenta evolved from a small invasive structure to a heavily invasive one, which is established early and lasts for a much longer gestational length (Guernsey et al., 2017, Renfree et al., 2009). Marsupials are born intensely altricial but have the ability to climb to the pouch where they see out most of their early development. Eutherian mammals do not use a pouch, but the longer gestation means that some eutherians are born more precocial and able to move and locate the mother to feed while others are still born intensely altricial. Either way, without the pouch, all eutherian mammals have a longer period out of the womb needing care for live-born young. Specifically, eutherian parents have to display higher and more robust maternal motivation, mothers now have to coordinate their high levels of maternal care with the birth of live young with no time for a learning period (hence the reliance on pregnancy hormones) and the offspring themselves are more dependent and display more care-demanding behaviours (such as USV's). Figure 7.1B displays this difference between mice and a very similar marsupial – the opossum.

This evolutionary transition from egg laying to pouch rearing to extended pregnancy/live birth aligns with the evolutionary expansion of imprinted genes in the marsupial (5-10) and then eutherian genome (100's) (Renfree et al., 2009, John and Surani, 2000). This could be coincidental but given the importance of imprinted genes for the placenta (Monk, 2015, Coan et al., 2005, John, 2017), and given that this thesis has shown that imprinted genes are convergently expressed in the neural circuitry for maintaining a heightened attention and motivation to care for offspring, then it is not infeasible that the evolution of IGs and Eutherian parenting could be linked. More explicitly, this would suggest that some imprinted genes could have evolved to be imprinted, and alter their expression, postnatally to fine-tune eutherians for the new challenges that come with their extended pregnancy and extended period of infant contact following birth, namely a more substantial placenta and higher parental motivation (Creeth et al., 2018). It is still plausible within this hypothesis that these genes originally evolved imprinted expression due to conflict reasons, but that these genes may not always be competing over post-natal resources like parenting but could have been co-opted into postnatal monoallelic expression to foster evolutionary change in the eutherian

lineage. However, hypotheses like this are still very speculative and much future work would be necessary (see below) to cement this relationship between imprinted genes and parenting behaviour before ideas about the co-evolution of the two could be substantiated.

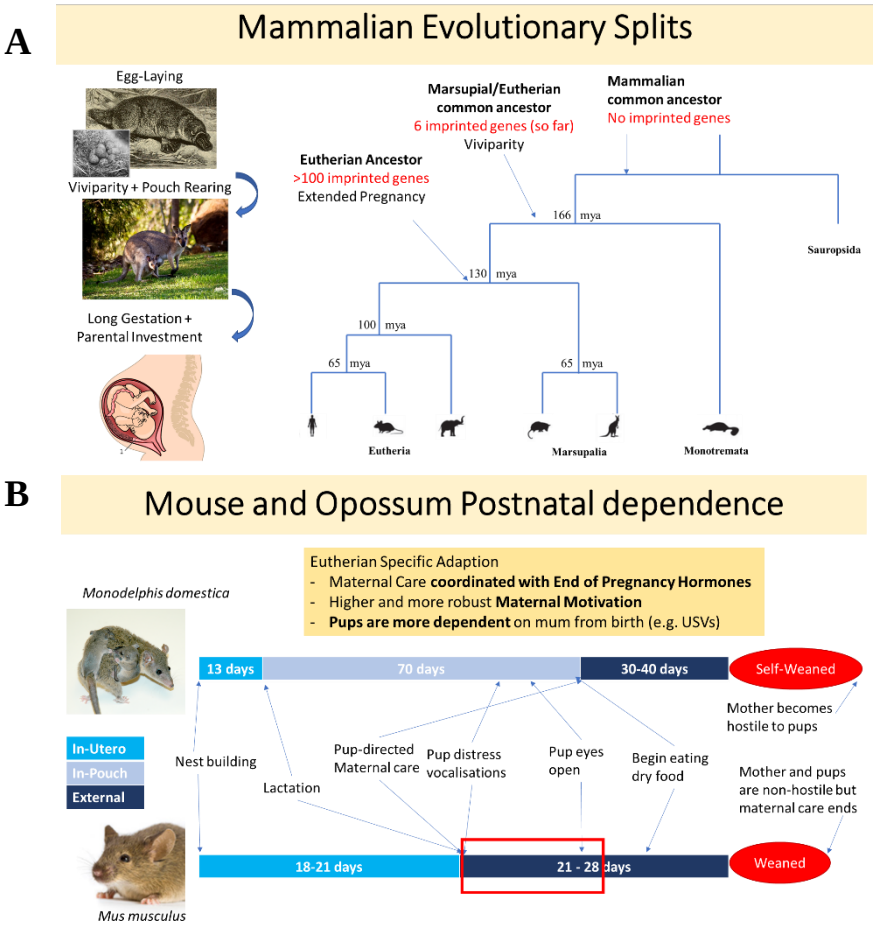


Figure 7.1 (Previous Page). Parental care and Imprinted genes share an evolutionary expansion during the Eutherian specialization. (A) Evolutionary expansion of imprinted genes mapped in evolutionary time with the evolutionary separation of monotremes and then marsupials from the Eutherian lineage, representing a transition from egg-laying to viviparity to longer gestation. (B) A comparison of developmental trajectory and offspring dependence in *Mus musculus* and *Monodelphis domestica* (Rousmaniere et al., 2010, Keyte and Smith, 2009). Mouse offspring are born at a later developmental stage than the opossum but, without the pouch, mice offspring have to transition from being blind and totally dependent on parents to a developed stage while exposed to the environment while the opossum hits these milestones within the security of the pouch and only emerges at the point where they can eat dry food. This increased dependence on mum and offspring appears to result in higher tolerance, with neither side resulting to hostility at the end of the parental care period which happens in opossum care. The red box highlights the critical time period in which Eutherian mammals have an enhanced dependency on their parents and hence the need for higher innate parental concern.

7.3 Future Directions

The principle future direction that has arisen from this thesis is the validation of more imprinted genes that displayed enriched expression in the parenting circuitry. *Magel2* was one of several promising candidates (incl. *Asb4, Bag3, Ndn, Nap1l5, Usp29*) that were highly expressed in galanin expressing neurons of the preoptic area. *Magel2* has now been demonstrated to have an association with parenting behaviour and POA^{Gal}. An exploration of whether these other genes show similar behavioural and molecular issues as the *Magel2*-null, as predicted by their expression profile, is essential further research. Firstly, this will further validate the method of using expression data to find novel imprinted gene function but will also serve to further support the ideas suggested in section 7.2.4, specifically, that imprinted genes are converging on activity of *Gal*-expressing POA neurons to modulate parenting behaviour.

In addition, most of the IGs that have previously been associated with parenting (*Mest, Peg3, Dio3, Calcr*) also showed enriched expression in the *Gal*-expressing POA in Chapter 3. However, the predominant hypothesis for how these genes impact this behaviour is via the oxytocin system of the PVN. It may be the case that *Magel2* was impacting on parenting via the oxytocin system (Schaller et al., 2010), and this is something to assess in the future, by exploring *c-Fos* response in these areas of the brain akin to the work described in Chapter 6. Regardless, this thesis has suggested the POA is important for imprinted genes, and that *Magel2* is functionally important in the POA. However, as of yet, POA dysfunction has not been assessed for models of *Peg3*. *Mest*, *Dio3* and *Peg13*. This is most likely an over-sight because it has been noted that *Peg3* (Li et al., 1999), *Peg13* (Davies et al., 2004) and *Dio3* (Escámez et al., 1999) are expressed at least as strongly in the POA as the PVN. Consequently, an investigation into the POA^{Gal} functionality of models of these mice would quickly confirm the validity of POA^{Gal} functionality being a target for imprinted genes affecting parenting behaviour.

Finally, with regards to *Magel2*-null mice (and future models of the other genes), there are a few directions worth taking. So far, all the mouse models used to assess imprinted genes in parenting have been global KO models, which as discussed in Sections 1.2.4 and 1.3, implicate multiple mechanisms that can impact parenting behaviour indirectly such as motility, olfaction, pituitary function, placental function and offspring behaviour. To truly cement the idea that dysfunction in galanin-expressing neurons in the POA is causing enough

of a disruption to create a parenting deficit, a *Gal*-specific KO line would greatly support this idea. However, not finding parenting deficits in this line would not necessarily go against the findings of this thesis, as it may be that the *Gal* disruption in combination with other mechanisms could build up to a parenting deficit. Additionally, though pup retrieval and nest-building are common practice for assessing parenting behaviour, these tests require an animal to perform in an artificial environment and experiment set-up. A potential avenue for future research that I feel would be beneficial would be to explore if genes that impact parenting had genuine deficits in the context of naturalistic parenting, i.e., how the parent engages with the highly demanding pups and how well do they get the nest ready a few days before and after birth. This could be done with a 24-hour monitoring system like the PhenoTyper cages (Noldus, Nottingham/UK).

7.4 Concluding Remarks

This thesis has contributed to our understanding of genomic imprinting and parenting in several ways. It has demonstrated that some neuron subtypes in the brain showed enriched imprinted gene expression in both expected neurons, such as the feeding neurons of the arcuate nucleus and some more unexpected ones, namely the galanin neurons of the POA. I have used this expression data to identify an imprinted gene with a parenting behaviour deficit, that was previously unknown, *Magel2*. This makes *Magel2* the 6th imprinted gene to show maternal care deficits. I additionally found the first imprinted gene to have an impact on paternal care and the parenting behaviour of virgin females. This thesis has also gone one step further and has proposed a mechanism by which *Magel2* may be creating these behavioural deficits; through modulating the activity of the galanin neurons in the POA in which *Magel2* is highly expressed. Finally, this thesis has provided the first suggestion that galanin functionality in the POA is one of the hotspots of imprinted gene activity, due to the identification of many imprinted genes showing high expression in the galanin neurons of the POA. This goes some way to explain why parenting deficits are observed in imprinted gene models as well as suggesting that the evolutionary improvement of parenting behaviour (alongside the expanded placenta and increased in-utero growth) may factor into the expansion of genomic imprinting in the Eutherian mammal genome.

References

ABI HABIB, W., BRIOUDE, F., AZZI, S., ROSSIGNOL, S., LINGLART, A., SOBRIER, M.-L., GIABICANI, É., STEUNOU, V., HARBISON, M. D. & LE BOUC, Y. 2019. Transcriptional profiling at the DLK1/MEG3 domain explains clinical overlap between imprinting disorders. *Science advances,* 5**,** eaau9425.

ABRAMOWITZ, L. K. & BARTOLOMEI, M. S. 2012. Genomic imprinting: recognition and marking of imprinted loci. *Current opinion in genetics & development,* 22**,** 72-78.

ABREU, A. P., DAUBER, A., MACEDO, D. B., NOEL, S. D., BRITO, V. N., GILL, J. C., CUKIER, P., THOMPSON, I. R., NAVARRO, V. M. & GAGLIARDI, P. C. 2013. Central precocious puberty caused by mutations in the imprinted gene MKRN3. *New England Journal of Medicine,* 368**,** 2467-2475.

ÅKEFELDT, A., TÖRNHAGE, C.-J. & GILLBERG, C. 1999. A woman with Prader–Willi syndrome gives birth to a healthy baby girl. *Developmental medicine and child neurology,* 41**,** 789-790.

AL ADHAMI, H., EVANO, B., LE DIGARCHER, A., GUEYDAN, C., DUBOIS, E., PARRINELLO, H., DANTEC, C., BOUSCHET, T., VARRAULT, A. & JOURNOT, L. 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 research,* 25**,** 353-367.

ALLEN, N. J. & BARRES, B. A. 2005. Signaling between glia and neurons: focus on synaptic plasticity. *Current opinion in neurobiology,* 15**,** 542-548.

ALSINA-LLANES, M., DE BRUN, V. & OLAZÁBAL, D. E. 2015. Development and expression of maternal behavior in naïve female C57BL/6 mice. *Developmental Psychobiology,* 57**,** 189-200.

ANDERGASSEN, D., DOTTER, C. P., WENZEL, D., SIGL, V., BAMMER, P. C., MUCKENHUBER, M., MAYER, D., KULINSKI, T. M., THEUSSL, H.-C. & PENNINGER, J. M. 2017. Mapping the mouse Allelome reveals tissue-specific regulation of allelic expression. *Elife,* 6**,** e25125.

ANDERGASSEN, D., MUCKENHUBER, M., BAMMER, P. C., KULINSKI, T. M., THEUSSL, H.-C., SHIMIZU, T., PENNINGER, J. M., PAULER, F. M. & HUDSON, Q. J. 2019. The Airn lncRNA does not require any DNA elements within its locus to silence distant imprinted genes. *PLoS Genetics,* 15**,** e1008268.

ANSELL, P. J., ZHOU, Y., SCHJEIDE, B.-M., KERNER, A., ZHAO, J., ZHANG, X. & KLIBANSKI, A. 2007. Regulation of growth hormone expression by Delta-like protein 1 (Dlk1). *Molecular and cellular endocrinology,* 271**,** 55-63.

ANUNCIADO-KOZA, R. P., STOHN, J. P., HERNANDEZ, A. & KOZA, R. A. 2022. Social and maternal behavior in mesoderm specific transcript (Mest)-deficient mice. *PloS one,* 17**,** e0271913.

ATES, T., ONCUL, M., DILSIZ, P., TOPCU, I. C., CIVAS, C. C., ALP, M. I., AKLAN, I., OZ, E. A., YAVUZ, Y. & YILMAZ, B. 2019. Inactivation of Magel2 suppresses oxytocin neurons through synaptic excitation-inhibition imbalance. *Neurobiology of disease,* 121**,** 58-64.

BABAK, T., DEVEALE, B., TSANG, E. K., ZHOU, Y., LI, X., SMITH, K. S., KUKURBA, K. R., ZHANG, R., LI, J. B. & VAN DER KOOY, D. 2015. Genetic conflict reflected in tissue-specific maps of genomic imprinting in human and mouse. *Nature genetics,* 47**,** 544-549.

BACH, K., PENSA, S., GRZELAK, M., HADFIELD, J., ADAMS, D. J., MARIONI, J. C. & KHALED, W. T. 2017. Differentiation dynamics of mammary epithelial cells revealed by single-cell RNA sequencing. *Nature communications,* 8**,** 1-11.
BAILEY, S. & ISOGAI, Y. 2022. Parenting as a model for behavioural switches. *Current Opinion in Neurobiology,* 73**,** 102543.
BARLOW, D., STÖGER, R., HERRMANN, B., SAITO, K. & SCHWEIFER, N. 1991. The mouse insulin-like growth factor type-2 receptor is imprinted and closely linked to the Tme locus. *Nature,* 349**,** 84-87.
BAROFSKY, A.-L., TAYLOR, J., TIZABI, Y., KUMAR, R. & JONES-QUARTEY, K. 1983. Specific neurotoxin lesions of median raphe serotonergic neurons disrupt maternal behavior in the lactating rat. *Endocrinology,* 113**,** 1884-1893.
BARON, M., VERES, A., WOLOCK, S. L., FAUST, A. L., GAUJOUX, R., VETERE, A., RYU, J. H., WAGNER, B. K., SHEN-ORR, S. S. & KLEIN, A. M. 2016. A single-cell transcriptomic map of the human and mouse pancreas reveals inter-and intra-cell population structure. *Cell systems,* 3**,** 346-360. e4.
BARTOLOMEI, M. S. & FERGUSON-SMITH, A. C. 2011. Mammalian genomic imprinting. *Cold Spring Harbor perspectives in biology,* 3**,** a002592.
BARTON, S. C., SURANI, M. & NORRIS, M. 1984. Role of paternal and maternal genomes in mouse development. *Nature,* 311**,** 374-376.
BEECHEY, C. 2000. Peg1/Mest locates distal to the currently defined imprinting region on mouse proximal chromosome 6 and identifies a new imprinting region affecting growth. *Cytogenetic and Genome Research,* 90**,** 309-314.
BENDESKY, A., KWON, Y.-M., LASSANCE, J.-M., LEWARCH, C. L., YAO, S., PETERSON, B. K., HE, M. X., DULAC, C. & HOEKSTRA, H. E. 2017. The genetic basis of parental care evolution in monogamous mice. *Nature,* 544**,** 434-439.
BIRD, A. 2002. DNA methylation patterns and epigenetic memory. *Genes & development,* 16**,** 6-21.
BISCHOF, J. M., STEWART, C. L. & WEVRICK, R. 2007. Inactivation of the mouse Magel2 gene results in growth abnormalities similar to Prader-Willi syndrome. *Human molecular genetics,* 16**,** 2713-2719.
BOILLOT, M. 2019. Hunger-activated AgRP neurons inhibit MPOA neurons controlling parenting. *Journal of Neuroscience,* 39**,** 6032-6034.
BONTHUIS, P. J., HUANG, W.-C., HÖRNDLI, C. N. S., FERRIS, E., CHENG, T. & GREGG, C. 2015. Noncanonical genomic imprinting effects in offspring. *Cell reports,* 12**,** 979-991.
BONTHUIS, P. J., STEINWAND, S., HÖRNDLI, C. N. S., EMERY, J., HUANG, W.-C., KRAVITZ, S., FERRIS, E. & GREGG, C. 2022. Noncanonical genomic imprinting in the monoamine system determines naturalistic foraging and brain-adrenal axis functions. *Cell reports,* 38**,** 110500.
BORIE, A. M., DROMARD, Y., DUFNER, D., POLLOZI, E., HUZARD, D., TÖMBÖLI, C., OLMA, A., MANNING, M., COLSON, P. & GUILLON, G. 2020. Control of social withdrawal of mice deficient for the autism gene Magel2 by restoration of vasopressin-oxytocin dialogue in septum. *bioRxiv***,** 800425.
BOSQUE ORTIZ, G. M., SANTANA, G. M. & DIETRICH, M. O. 2022. Deficiency of the paternally inherited gene Magel2 alters the development of separation-induced vocalization and maternal behavior in mice. *Genes, Brain and Behavior,* 21**,** e12776.
BOWERS, J. M., PEREZ-POUCHOULEN, M., EDWARDS, N. S. & MCCARTHY, M. M. 2013. Foxp2 mediates sex differences in ultrasonic vocalization by rat pups and directs order of maternal retrieval. *Journal of Neuroscience,* 33**,** 3276-3283.

BRICKLEY, S. G., ALLER, M. I., SANDU, C., VEALE, E. L., ALDER, F. G., SAMBI, H., MATHIE, A. & WISDEN, W. 2007. TASK-3 two-pore domain potassium channels enable sustained high-frequency firing in cerebellar granule neurons. *Journal of Neuroscience,* 27**,** 9329-9340.
BRIDGES, R. S. 1996. Biochemical basis of parental behavior in the rat. *Advances in the Study of Behavior.* Elsevier.
BRIDGES, R. S. 2015. Neuroendocrine regulation of maternal behavior. *Frontiers in neuroendocrinology,* 36**,** 178-196.
BRIDGES, R. S. 2016. Long-term alterations in neural and endocrine processes induced by motherhood in mammals. *Hormones and Behavior,* 77**,** 193-203.
BRIDGES, R. S. & FREEMARK, M. S. 1995. Human placental lactogen infusions into the medial preoptic area stimulate maternal behavior in steroid-primed, nulliparous female rats. *Hormones and behavior,* 29**,** 216-226.
BRIDGES, R. S., NUMAN, M., RONSHEIM, P. M., MANN, P. E. & LUPINI, C. E. 1990. Central prolactin infusions stimulate maternal behavior in steroid-treated, nulliparous female rats. *Proceedings of the National Academy of Sciences,* 87**,** 8003-8007.
BRIDGES, R. S., ROBERTSON, M. C., SHIU, R. P., FRIESEN, H. G., STUER, A. M. & MANN, P. E. 1996. Endocrine communication between conceptus and mother: placental lactogen stimulation of maternal behavior. *Neuroendocrinology,* 64**,** 57-64.
BROWN, R. S., AOKI, M., LADYMAN, S. R., PHILLIPPS, H. R., WYATT, A., BOEHM, U. & GRATTAN, D. R. 2017. Prolactin action in the medial preoptic area is necessary for postpartum maternal nursing behavior. *Proceedings of the National Academy of Sciences,* 114**,** 10779-10784.
BULLITT, E. 1990. Expression of c-fos-like protein as a marker for neuronal activity following noxious stimulation in the rat. *Journal of Comparative Neurology,* 296**,** 517-530.
BURLEY, N. T. & JOHNSON, K. 2002. The evolution of avian parental care. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences,* 357**,** 241-250.
BUTLER, M. G. 2009. Genomic imprinting disorders in humans: a mini-review. *Journal of assisted reproduction and genetics,* 26**,** 477-486.
CALDWELL, H. K. 2017. Oxytocin and vasopressin: powerful regulators of social behavior. *The Neuroscientist,* 23**,** 517-528.
CAMPBELL, J. N., MACOSKO, E. Z., FENSELAU, H., PERS, T. H., LYUBETSKAYA, A., TENEN, D., GOLDMAN, M., VERSTEGEN, A. M., RESCH, J. M. & MCCARROLL, S. A. 2017. A molecular census of arcuate hypothalamus and median eminence cell types. *Nature neuroscience,* 20**,** 484-496.
CARLSON, N. R. & THOMAS, G. J. 1968. Maternal behavior of mice with limbic lesions. *Journal of Comparative and Physiological Psychology,* 66**,** 731.
CASE, T. J. 1978. Endothermy and parental care in the terrestrial vertebrates. *The American Naturalist,* 112**,** 861-874.
CASSIDY, F. C. & CHARALAMBOUS, M. 2018. Genomic imprinting, growth and maternal–fetal interactions. *Journal of Experimental Biology,* 221**,** jeb164517.
CASSIDY, S. B., DYKENS, E. & WILLIAMS, C. A. 2000. Prader-Willi and Angelman syndromes: Sister imprinted disorders. *American journal of medical genetics,* 97**,** 136-146.
CATTANACH, B. M. & KIRK, M. 1985. Differential activity of maternally and paternally derived chromosome regions in mice. *Nature,* 315**,** 496-498.
CAVAILLÉ, J., BUITING, K., KIEFMANN, M., LALANDE, M., BRANNAN, C. I., HORSTHEMKE, B., BACHELLERIE, J.-P., BROSIUS, J. & HÜTTENHOFER, A.

2000. Identification of brain-specific and imprinted small nucleolar RNA genes exhibiting an unusual genomic organization. *Proceedings of the National Academy of Sciences,* 97**,** 14311-14316.
CHAMBERLAIN, S. J., JOHNSTONE, K. A., DUBOSE, A. J., SIMON, T. A., BARTOLOMEI, M. S., RESNICK, J. L. & BRANNAN, C. I. 2004. Evidence for genetic modifiers of postnatal lethality in PWS-IC deletion mice. *Human molecular genetics,* 13**,** 2971-2977.
CHAMPAGNE, F. A., CURLEY, J. P., SWANEY, W. T., HASEN, N. & KEVERNE, E. B. 2009. Paternal influence on female behavior: the role of Peg3 in exploration, olfaction, and neuroendocrine regulation of maternal behavior of female mice. *Behavioral neuroscience,* 123**,** 469.
CHAO, W. & D'AMORE, P. A. 2008. IGF2: epigenetic regulation and role in development and disease. *Cytokine & growth factor reviews,* 19**,** 111-120.
CHARALAMBOUS, M., COWLEY, M., GEOGHEGAN, F., SMITH, F. M., RADFORD, E. J., MARLOW, B. P., GRAHAM, C. F., HURST, L. D. & WARD, A. 2010. Maternally-inherited Grb10 reduces placental size and efficiency. *Developmental biology,* 337**,** 1-8.
CHARALAMBOUS, M., DA ROCHA, S. T., RADFORD, E. J., MEDINA-GOMEZ, G., CURRAN, S., PINNOCK, S. B., FERRÓN, S. R., VIDAL-PUIG, A. & FERGUSON-SMITH, A. C. 2014. DLK1/PREF1 regulates nutrient metabolism and protects from steatosis. *Proceedings of the National Academy of Sciences,* 111**,** 16088-16093.
CHARALAMBOUS, M., FERRON, S. R., DA ROCHA, S. T., MURRAY, A. J., ROWLAND, T., ITO, M., SCHUSTER-GOSSLER, K., HERNANDEZ, A. & FERGUSON-SMITH, A. C. 2012. Imprinted gene dosage is critical for the transition to independent life. *Cell metabolism,* 15**,** 209-221.
CHARALAMBOUS, M., SMITH, F. M., BENNETT, W. R., CREW, T. E., MACKENZIE, F. & WARD, A. 2003. Disruption of the imprinted Grb10 gene leads to disproportionate overgrowth by an Igf2-independent mechanism. *Proceedings of the National Academy of Sciences,* 100**,** 8292-8297.
CHARVET, C. J. & STRIEDTER, G. F. 2011. Developmental modes and developmental mechanisms can channel brain evolution. *Frontiers in Neuroanatomy,* 5**,** 4.
CHEN, H., VICTOR, A. K., KLEIN, J., TACER, K. F., TAI, D. J., DE ESCH, C., NUTTLE, A., TEMIROV, J., BURNETT, L. C. & ROSENBAUM, M. 2020. Loss of MAGEL2 in Prader-Willi syndrome leads to decreased secretory granule and neuropeptide production. *JCI insight,* 5.
CHEN, R., WU, X., JIANG, L. & ZHANG, Y. 2017. Single-cell RNA-seq reveals hypothalamic cell diversity. *Cell reports,* 18**,** 3227-3241.
CHEUNG, L. Y., GEORGE, A. S., MCGEE, S. R., DALY, A. Z., BRINKMEIER, M. L., ELLSWORTH, B. S. & CAMPER, S. A. 2018. Single-cell RNA sequencing reveals novel markers of male pituitary stem cells and hormone-producing cell types. *Endocrinology,* 159**,** 3910-3924.
CLEATON, M. A., DENT, C. L., HOWARD, M., CORISH, J. A., GUTTERIDGE, I., SOVIO, U., GACCIOLI, F., TAKAHASHI, N., BAUER, S. R. & CHARNOCK-JONES, D. S. 2016. Fetus-derived DLK1 is required for maternal metabolic adaptations to pregnancy and is associated with fetal growth restriction. *Nature genetics,* 48**,** 1473-1480.
CLEATON, M. A., EDWARDS, C. A. & FERGUSON-SMITH, A. C. 2014. Phenotypic outcomes of imprinted gene models in mice: elucidation of pre-and postnatal functions of imprinted genes. *Annual review of genomics and human genetics,* 15**,** 93-126.

COAN, P., BURTON, G. & FERGUSON-SMITH, A. 2005. Imprinted genes in the placenta–a review. *Placenta,* 26**,** S10-S20.
COCKBURN, A. 2006. Prevalence of different modes of parental care in birds. *Proceedings of the Royal Society B: Biological Sciences,* 273**,** 1375-1383.
CONSIGLIO, A. R. & BRIDGES, R. S. 2009. Circulating prolactin, MPOA prolactin receptor expression and maternal aggression in lactating rats. *Behavioural brain research,* 197**,** 97-102.
COSI, C., MILLAR, M., BELTRAN, M., SHERRY, L. & GATTI-MCARTHUR, S. 2021. Quantitative analysis of RNAscope staining for c-fos expression in mouse brain tissue as a measure of Neuronal Activation. *MethodsX,* 8**,** 101348.
COULSON, R. L., YASUI, D. H., DUNAWAY, K. W., LAUFER, B. I., VOGEL CIERNIA, A., ZHU, Y., MORDAUNT, C. E., TOTAH, T. S. & LASALLE, J. M. 2018. Snord116-dependent diurnal rhythm of DNA methylation in mouse cortex. *Nature communications,* 9**,** 1-11.
COWLEY, M., GARFIELD, A. S., MADON-SIMON, M., CHARALAMBOUS, M., CLARKSON, R. W., SMALLEY, M. J., KENDRICK, H., ISLES, A. R., PARRY, A. J. & CARNEY, S. 2014. Developmental programming mediated by complementary roles of imprinted Grb10 in mother and pup. *PLoS biology,* 12**,** e1001799.
COWLEY, M. & OAKEY, R. J. 2010. Retrotransposition and genomic imprinting. *Briefings in functional genomics,* 9**,** 340-346.
CREETH, H., MCNAMARA, G., ISLES, A. & JOHN, R. 2019. Imprinted genes influencing the quality of maternal care. *Frontiers in neuroendocrinology,* 53**,** 100732.
CREETH, H. D., MCNAMARA, G. I., TUNSTER, S. J., BOQUE-SASTRE, R., ALLEN, B., SUMPTION, L., EDDY, J. B., ISLES, A. R. & JOHN, R. M. 2018. Maternal care boosted by paternal imprinting in mammals. *PLoS biology,* 16**,** e2006599.
CURLEY, J., PINNOCK, S., DICKSON, S., THRESHER, R., MIYOSHI, N., SURANI, M. & KEVERNE, E. 2005. Increased body fat in mice with a targeted mutation of the paternally expressed imprinted gene Peg3. *The FASEB Journal,* 19**,** 1302-1304.
CURLEY, J. P., BARTON, S., SURANI, A. & KEVERNE, E. B. 2004. Coadaptation in mother and infant regulated by a paternally expressed imprinted gene. *Proceedings of the Royal Society of London. Series B: Biological Sciences,* 271**,** 1303-1309.
CURLEY, J. P. & MASHOODH, R. 2010. Parent-of-origin and trans-generational germline influences on behavioral development: The interacting roles of mothers, fathers, and grandparents. *Developmental psychobiology,* 52**,** 312-330.
DA PRATO, L. C., ZAYAN, U., ABDALLAH, D., POINT, V., SCHALLER, F., PALLESI-POCACHARD, E., MONTHEIL, A., CANAAN, S., GAIARSA, J.-L. & MUSCATELLI, F. 2022. Early life oxytocin treatment improves thermo-sensory reactivity and maternal behavior in neonates lacking the autism-associated gene Magel2. *Neuropsychopharmacology***,** 1-12.
DA ROCHA, S. T., TEVENDALE, M., KNOWLES, E., TAKADA, S., WATKINS, M. & FERGUSON-SMITH, A. C. 2007. Restricted co-expression of Dlk1 and the reciprocally imprinted non-coding RNA, Gtl2: implications for cis-acting control. *Developmental biology,* 306**,** 810-823.
DAVIES, W., LYNN, P. M., RELKOVIC, D. & WILKINSON, L. S. 2008. Imprinted genes and neuroendocrine function. *Frontiers in neuroendocrinology,* 29**,** 413-427.
DAVIES, W., SMITH, R. J., KELSEY, G. & WILKINSON, L. S. 2004. Expression patterns of the novel imprinted genes Nap1l5 and Peg13 and their non-imprinted host genes in the adult mouse brain. *Gene expression patterns,* 4**,** 741-747.
DE MICHELI, A. J., LAURILLIARD, E. J., HEINKE, C. L., RAVICHANDRAN, H., FRACZEK, P., SOUEID-BAUMGARTEN, S., DE VLAMINCK, I., ELEMENTO,

O. & COSGROVE, B. D. 2020. Single-cell analysis of the muscle stem cell hierarchy identifies heterotypic communication signals involved in skeletal muscle regeneration. *Cell reports,* 30**,** 3583-3595. e5.
DE PITTÀ, M., BRUNEL, N. & VOLTERRA, A. 2016. Astrocytes: Orchestrating synaptic plasticity? *Neuroscience,* 323**,** 43-61.
DEACON, R. M. 2006. Assessing nest building in mice. *Nature protocols,* 1**,** 1117-1119.
DECHIARA, T. M., ROBERTSON, E. J. & EFSTRATIADIS, A. 1991. Parental imprinting of the mouse insulin-like growth factor II gene. *Cell,* 64**,** 849-859.
DENG, Q., RAMSKÖLD, D., REINIUS, B. & SANDBERG, R. 2014. Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. *Science,* 343**,** 193-196.
DENIZOT, A.-L., BESSON, V., CORRERA, R. M., MAZZOLA, A., LOPES, I., COURBARD, J.-R., MARAZZI, G. & SASSOON, D. A. 2016. A novel mutant allele of Pw1/Peg3 does not affect maternal behavior or nursing behavior. *PLoS genetics,* 12**,** e1006053.
DENT, C., HUMBY, T., LEWIS, K., PLAGGE, A., FISCHER-COLBRIE, R., WILKINS, J., WILKINSON, L. & ISLES, A. R. 2016. Impulsive choices in mice lacking imprinted Nesp55. *Genes, Brain and Behavior,* 15**,** 693-701.
DENT, C. L. & ISLES, A. R. 2014. Brain-expressed imprinted genes and adult behaviour: the example of Nesp and Grb10. *Mammalian genome,* 25**,** 87-93.
DEVEALE, B., VAN DER KOOY, D. & BABAK, T. 2012. Critical evaluation of imprinted gene expression by RNA–Seq: a new perspective. *PLoS genetics,* 8**,** e1002600.
DONOVAN, M. H. & TECOTT, L. H. 2013. Serotonin and the regulation of mammalian energy balance. *Frontiers in neuroscience,* 7**,** 36.
DOODY, J. S., BURGHARDT, G. M. & DINETS, V. 2013. Breaking the social–non-social dichotomy: a role for reptiles in vertebrate social behavior research? *Ethology,* 119**,** 95-103.
DULAC, C., O'CONNELL, L. A. & WU, Z. 2014. Neural control of maternal and paternal behaviors. *Science,* 345**,** 765-770.
EDWARDS, C. A., RENS, W., CLARKE, O., MUNGALL, A. J., HORE, T., GRAVES, J. A. M., DUNHAM, I., FERGUSON-SMITH, A. C. & FERGUSON-SMITH, M. A. 2007. The evolution of imprinting: chromosomal mapping of orthologues of mammalian imprinted domains in monotreme and marsupial mammals. *BMC evolutionary biology,* 7**,** 1-12.
EGGERMANN, T., DAVIES, J. H., TAUBER, M., VAN DEN AKKER, E., HOKKEN-KOELEGA, A., JOHANSSON, G. & NETCHINE, I. 2021. Growth restriction and genomic imprinting-overlapping phenotypes support the concept of an imprinting network. *Genes,* 12**,** 585.
EHLEN, J. C., JONES, K. A., PINCKNEY, L., GRAY, C. L., BURETTE, S., WEINBERG, R. J., EVANS, J. A., BRAGER, A. J., ZYLKA, M. J. & PAUL, K. N. 2015. Maternal Ube3a loss disrupts sleep homeostasis but leaves circadian rhythmicity largely intact. *Journal of Neuroscience,* 35**,** 13587-13598.
ELWOOD, R. W. 1985. Inhibition of infanticide and onset of paternal care in male mice (Mus musculus). *Journal of Comparative Psychology,* 99**,** 457.
ESCÁMEZ, M. A. J., GUADAÑO-FERRAZ, A., CUADRADO, A. & BERNAL, J. 1999. Type 3 iodothyronine deiodinase is selectively expressed in areas related to sexual differentiation in the newborn rat brain. *Endocrinology,* 140**,** 5443-5446.
ESQUILIANO, D. R., GUO, W., LIANG, L., DIKKES, P. & LOPEZ, M. F. 2009. Placental glycogen stores are increased in mice with H19 null mutations but not in those with insulin or IGF type 1 receptor mutations. *Placenta,* 30**,** 693-699.

FAUST, K. M., CAROUSO-PECK, S., ELSON, M. R. & GOLDSTEIN, M. H. 2020. The origins of social knowledge in altricial species. *Annual review of developmental psychology,* 2**,** 225.
FEIL, R. & BERGER, F. 2007. Convergent evolution of genomic imprinting in plants and mammals. *Trends in Genetics,* 23**,** 192-199.
FERGUSON-SMITH, A. C. 2011. Genomic imprinting: the emergence of an epigenetic paradigm. *Nature Reviews Genetics,* 12**,** 565-575.
FERREY, A. E., SANTASCOY, N., MCCRORY, E. J., THOMPSON-BOOTH, C., MAYES, L. C. & RUTHERFORD, H. J. 2016. Motivated attention and reward in parenting. *Parenting,* 16**,** 284-301.
FERRIS, C. F., KULKARNI, P., SULLIVAN, J. M., HARDER, J. A., MESSENGER, T. L. & FEBO, M. 2005. Pup suckling is more rewarding than cocaine: evidence from functional magnetic resonance imaging and three-dimensional computational analysis. *Journal of Neuroscience,* 25**,** 149-156.
FERRÓN, S., RADFORD, E., DOMINGO-MUELAS, A., KLEINE, I., RAMME, A., GRAY, D., SANDOVICI, I., CONSTANCIA, M., WARD, A. & MENHENIOTT, T. 2015. Differential genomic imprinting regulates paracrine and autocrine roles of IGF2 in mouse adult neurogenesis. *Nature communications,* 6**,** 1-12.
FERRÓN, S. R., CHARALAMBOUS, M., RADFORD, E., MCEWEN, K., WILDNER, H., HIND, E., MORANTE-REDOLAT, J. M., LABORDA, J., GUILLEMOT, F. & BAUER, S. R. 2011. Postnatal loss of Dlk1 imprinting in stem cells and niche astrocytes regulates neurogenesis. *Nature,* 475**,** 381-385.
FOUNTAIN, M. D., ATEN, E., CHO, M. T., JUUSOLA, J., WALKIEWICZ, M. A., RAY, J. W., XIA, F., YANG, Y., GRAHAM, B. H. & BACINO, C. A. 2017a. The phenotypic spectrum of Schaaf-Yang syndrome: 18 new affected individuals from 14 families. *Genetics in Medicine,* 19**,** 45-52.
FOUNTAIN, M. D. & SCHAAF, C. P. 2015. MAGEL2 and Oxytocin—Implications in Prader-Willi Syndrome and Beyond. *Biological Psychiatry,* 78**,** 78-80.
FOUNTAIN, M. D. & SCHAAF, C. P. 2016. Prader-Willi syndrome and Schaaf-Yang syndrome: neurodevelopmental diseases intersecting at the MAGEL2 gene. *Diseases,* 4**,** 2.
FOUNTAIN, M. D., TAO, H., CHEN, C. A., YIN, J. & SCHAAF, C. P. 2017b. Magel2 knockout mice manifest altered social phenotypes and a deficit in preference for social novelty. *Genes, Brain and Behavior,* 16**,** 592-600.
FRESARD, L., LEROUX, S., SERVIN, B., GOURICHON, D., DEHAIS, P., CRISTOBAL, M. S., MARSAUD, N., VIGNOLES, F., BED'HOM, B. & COVILLE, J.-L. 2014. Transcriptome-wide investigation of genomic imprinting in chicken. *Nucleic acids research,* 42**,** 3768-3782.
FREY, W. D. & KIM, J. 2014. APeg3: regulation of Peg3 through an evolutionarily conserved ncRNA. *Gene,* 540**,** 251-257.
FREY, W. D., SHARMA, K., CAIN, T. L., NISHIMORI, K., TERUYAMA, R. & KIM, J. 2018. Oxytocin receptor is regulated by Peg3. *Plos one,* 13**,** e0202476.
FROST, J. M. & MOORE, G. E. 2010. The importance of imprinting in the human placenta. *PLoS genetics,* 6**,** e1001015.
GABORY, A., RIPOCHE, M.-A., LE DIGARCHER, A., WATRIN, F., ZIYYAT, A., FORNÉ, T., JAMMES, H., AINSCOUGH, J. F., SURANI, M. A. & JOURNOT, L. 2009. H19 acts as a trans regulator of the imprinted gene network controlling growth in mice. *Development,* 136**,** 3413-3421.
GARFIELD, A. S., COWLEY, M., SMITH, F. M., MOORWOOD, K., STEWART-COX, J. E., GILROY, K., BAKER, S., XIA, J., DALLEY, J. W. & HURST, L. D. 2011.

Distinct physiological and behavioural functions for parental alleles of imprinted Grb10. *Nature,* 469**,** 534-538.
GOTTER, A. L., SANTARELLI, V. P., DORAN, S. M., TANNENBAUM, P. L., KRAUS, R. L., ROSAHL, T. W., MEZIANE, H., MONTIAL, M., REISS, D. R. & WESSNER, K. 2011. TASK-3 as a potential antidepressant target. *Brain research,* 1416**,** 69-79.
GRATTAN, D. & KOKAY, I. 2008. Prolactin: a pleiotropic neuroendocrine hormone. *Journal of neuroendocrinology,* 20**,** 752-763.
GRATTAN, D. R. 2015. 60 years of neuroendocrinology: the hypothalamo-prolactin axis. *Journal of Endocrinology,* 226**,** T101-T122.
GRATTAN, D. R., STEYN, F. J., KOKAY, I. C., ANDERSON, G. M. & BUNN, S. J. 2008. Pregnancy-induced adaptation in the neuroendocrine control of prolactin secretion. *Journal of neuroendocrinology,* 20**,** 497-507.
GREGG, C., ZHANG, J., WEISSBOURD, B., LUO, S., SCHROTH, G. P., HAIG, D. & DULAC, C. 2010. High-resolution analysis of parent-of-origin allelic expression in the mouse brain. *science,* 329**,** 643-648.
GROSS-TSUR, V., HIRSCH, H. J., BENARROCH, F. & ELDAR-GEVA, T. 2012. The FSH-inhibin axis in prader-willi syndrome: heterogeneity of gonadal dysfunction. *Reproductive Biology and Endocrinology,* 10**,** 1-7.
GROSS, M. R. 2005. The evolution of parental care. *The Quarterly review of biology,* 80**,** 37-45.
GUBERNICK, D. J. 2013. *Parental care in mammals,* Springer Science & Business Media.
GUERNSEY, M. W., CHUONG, E. B., CORNELIS, G., RENFREE, M. B. & BAKER, J. C. 2017. Molecular conservation of marsupial and eutherian placentation and lactation. *Elife,* 6**,** e27450.
GUYON, A., TARDY, M. P., ROVERE, C., NAHON, J.-L., BARHANIN, J. & LESAGE, F. 2009. Glucose inhibition persists in hypothalamic neurons lacking tandem-pore K+ channels. *Journal of Neuroscience,* 29**,** 2528-2533.
HAIG, D. 1999. Multiple paternity and genomic imprinting. *Genetics,* 151**,** 1229-1231.
HAIG, D. 2000. The kinship theory of genomic imprinting. *Annual review of ecology and systematics***,** 9-32.
HAIG, D. 2014. Coadaptation and conflict, misconception and muddle, in the evolution of genomic imprinting. *Heredity,* 113**,** 96-103.
HAIG, D. 2015. Maternal–fetal conflict, genomic imprinting and mammalian vulnerabilities to cancer. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 370**,** 20140178.
HAMILTON, W. D. 1963. The evolution of altruistic behavior. *The American Naturalist,* 97**,** 354-356.
HAMILTON, W. D. 1964. The genetical evolution of social behaviour. II. *Journal of theoretical biology,* 7**,** 17-52.
HAN, X., WANG, R., ZHOU, Y., FEI, L., SUN, H., LAI, S., SAADATPOUR, A., ZHOU, Z., CHEN, H. & YE, F. 2018. Mapping the mouse cell atlas by microwell-seq. *Cell,* 172**,** 1091-1107. e17.
HART, T., KOMORI, H. K., LAMERE, S., PODSHIVALOVA, K. & SALOMON, D. R. 2013. Finding the active genes in deep RNA-seq gene expression studies. *BMC genomics,* 14**,** 1-7.
HEKSCH, R., KAMBOJ, M., ANGLIN, K. & OBRYNBA, K. 2017. Review of Prader-Willi syndrome: the endocrine approach. *Translational pediatrics,* 6**,** 274.
HENCKEL, A. & ARNAUD, P. 2010. Genome-wide identification of new imprinted genes. *Briefings in functional genomics,* 9**,** 304-314.

HERNANDEZ, A., FIERING, S., MARTINEZ, E., GALTON, V. A. & ST. GERMAIN, D. 2002. The gene locus encoding iodothyronine deiodinase type 3 (Dio3) is imprinted in the fetus and expresses antisense transcripts. *Endocrinology,* 143**,** 4483-4486.
HERNANDEZ, A. & STOHN, J. P. 2018. The type 3 deiodinase: epigenetic control of brain thyroid hormone action and neurological function. *International Journal of Molecular Sciences,* 19**,** 1804.
HO, Y., HU, P., PEEL, M. T., CHEN, S., CAMARA, P. G., EPSTEIN, D. J., WU, H. & LIEBHABER, S. A. 2020. Single-cell transcriptomic analysis of adult mouse pituitary reveals sexual dimorphism and physiologic demand-induced cellular plasticity. *Protein & cell,* 11**,** 565-583.
HOOK, P. W., MCCLYMONT, S. A., CANNON, G. H., LAW, W. D., MORTON, A. J., GOFF, L. A. & MCCALLION, A. S. 2018. Single-cell RNA-Seq of mouse dopaminergic neurons informs candidate gene selection for sporadic Parkinson disease. *The American Journal of Human Genetics,* 102**,** 427-446.
HORE, T. A., RAPKINS, R. W. & GRAVES, J. A. M. 2007. Construction and evolution of imprinted loci in mammals. *TRENDS in Genetics,* 23**,** 440-448.
HOSHIYA, H., MEGURO, M., KASHIWAGI, A., OKITA, C. & OSHIMURA, M. 2003. Calcr, a brain-specific imprinted mouse calcitonin receptor gene in the imprinted cluster of the proximal region of chromosome 6. *Journal of human genetics,* 48**,** 208-211.
HUANG, K. W., OCHANDARENA, N. E., PHILSON, A. C., HYUN, M., BIRNBAUM, J. E., CICCONET, M. & SABATINI, B. L. 2019. Molecular and anatomical organization of the dorsal raphe nucleus. *Elife,* 8.
HUDSON, Q., KULINSKI, T., HUETTER, S. & BARLOW, D. 2010. Genomic imprinting mechanisms in embryonic and extraembryonic mouse tissues. *Heredity,* 105**,** 45-56.
HUERTA-OCAMPO, I., SLACK, R., BEECHEY, C., SKINNER, J., PETERS, J. & CHRISTIAN, H. Overexpression of the imprinted gene Neuronatin represses normal pituitary differentiation. Endocrine Abstracts, 2004. Bioscientifica.
HURST, L. D. & MCVEAN, G. T. 1997. Growth effects of uniparental disomies and the conflict theory of genomic imprinting. *Trends in Genetics,* 13**,** 436-443.
HUTTER, B., HELMS, V. & PAULSEN, M. 2006. Tandem repeats in the CpG islands of imprinted genes. *Genomics,* 88**,** 323-332.
INESON, J., STAYNER, C., HAZLETT, J., SLOBBE, L., ROBSON, E., LEGGE, M. & ECCLES, M. R. 2012. Somatic reactivation of expression of the silent maternal Mest allele and acquisition of normal reproductive behaviour in a colony of Peg1/Mest mutant mice. *Journal of Reproduction and Development***,** 1204160456-1204160456.
ISLES, A. R., MCNAMARA, G. I. & JOHN, R. M. 2019. Genomic imprinting and neurobehavioral programming by adverse early life environments: evidence from studying Cdkn1c. *Current opinion in behavioral sciences,* 25**,** 31-35.
IVANOVA, E. & KELSEY, G. 2011. Imprinted genes and hypothalamic function. *Journal of molecular endocrinology,* 47**,** R67-R74.
IWASA, Y. 1998. 7 The Conflict Theory of Genomic Imprinting: How Much Can Be Explained? *Current topics in developmental biology,* 40**,** 255-293.
JACOBS, F. M., VAN DER LINDEN, A. J., WANG, Y., VON OERTHEL, L., SUL, H. S., BURBACH, J. P. H. & SMIDT, M. P. 2009. Identification of Dlk1, Ptpru and Klhl1 as novel Nurr1 target genes in meso-diencephalic dopamine neurons.
JIANG, Y.-H., PAN, Y., ZHU, L., LANDA, L., YOO, J., SPENCER, C., LORENZO, I., BRILLIANT, M., NOEBELS, J. & BEAUDET, A. L. 2010. Altered ultrasonic vocalization and impaired learning and memory in Angelman syndrome mouse model with a large maternal deletion from Ube3a to Gabrb3. *PloS one,* 5**,** e12278.

JOHN, R. M. 2017. Imprinted genes and the regulation of placental endocrine function: Pregnancy and beyond. *Placenta,* 56**,** 86-90.
JOHN, R. M. & SURANI, M. A. 2000. Genomic imprinting, mammalian evolution, and the mystery of egg-laying mammals. *Cell,* 101**,** 585-588.
JONES, K. A., HAN, J. E., DEBRUYNE, J. P. & PHILPOT, B. D. 2016. Persistent neuronal Ube3a expression in the suprachiasmatic nucleus of Angelman syndrome model mice. *Scientific reports,* 6**,** 1-13.
JOSEPH, B., WALLÉN-MACKENZIE, Å., BENOIT, G., MURATA, T., JOODMARDI, E., OKRET, S. & PERLMANN, T. 2003. p57Kip2 cooperates with Nurr1 in developing dopamine cells. *Proceedings of the National Academy of Sciences,* 100**,** 15619-15624.
KAMALUDIN, A. A., SMOLARCHUK, C., BISCHOF, J. M., EGGERT, R., GREER, J. J., REN, J., LEE, J. J., YOKOTA, T., BERRY, F. B. & WEVRICK, R. 2016. Muscle dysfunction caused by loss of Magel2 in a mouse model of Prader-Willi and Schaaf-Yang syndromes. *Human molecular genetics,* 25**,** 3798-3809.
KANEKO-ISHINO, T., KUROIWA, Y., MIYOSHI, N., KOHDA, T., SUZUKI, R., YOKOYAMA, M., VIVILLE, S., BARTON, S. C., ISHINO, F. & SURANI, M. A. 1995. Peg1/Mest imprinted gene on chromosome 6 identified by cDNA subtraction hybridization. *Nature genetics,* 11**,** 52-59.
KELSEY, G. & FEIL, R. 2013. New insights into establishment and maintenance of DNA methylation imprints in mammals. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 368**,** 20110336.
KENKEL, W. M., PERKEYBILE, A. M. & CARTER, C. S. 2017. The neurobiological causes and effects of alloparenting. *Developmental neurobiology,* 77**,** 214-232.
KESHAVARZ, M. & TAUTZ, D. 2021. The imprinted lncRNA Peg13 regulates sexual preference and the sex-specific brain transcriptome in mice. *Proceedings of the National Academy of Sciences,* 118.
KESSLER, M. S., BOSCH, O. J., BUNCK, M., LANDGRAF, R. & NEUMANN, I. D. 2011. Maternal care differs in mice bred for high vs. low trait anxiety: impact of brain vasopressin and cross-fostering. *Social neuroscience,* 6**,** 156-168.
KEVERNE, E. 2014. Significance of epigenetics for understanding brain development, brain evolution and behaviour. *Neuroscience,* 264**,** 207-217.
KEVERNE, E. B., FUNDELE, R., NARASIMHA, M., BARTON, S. C. & SURANI, M. A. 1996a. Genomic imprinting and the differential roles of parental genomes in brain development. *Developmental Brain Research,* 92**,** 91-100.
KEVERNE, E. B., MARTEL, F. L. & NEVISON, C. M. 1996b. Primate brain evolution: genetic and functional considerations. *Proceedings of the Royal Society of London. Series B: Biological Sciences,* 263**,** 689-696.
KEYTE, A. L. & SMITH, K. K. 2009. Opossum (Monodelphis domestica).
KILLIAN, J. K., BYRD, J. C., JIRTLE, J. V., MUNDAY, B. L., STOSKOPF, M. K., MACDONALD, R. G. & JIRTLE, R. L. 2000. M6P/IGF2R imprinting evolution in mammals. *Molecular cell,* 5**,** 707-716.
KILLIAN, J. K., NOLAN, C. M., STEWART, N., MUNDAY, B. L., ANDERSEN, N. A., NICOL, S. & JIRTLE, R. L. 2001. Monotreme IGF2 expression and ancestral origin of genomic imprinting. *Journal of Experimental Zoology,* 291**,** 205-212.
KOHL, J., AUTRY, A. E. & DULAC, C. 2017. The neurobiology of parenting: A neural circuit perspective. *Bioessays,* 39**,** 1-11.
KOHL, J., BABAYAN, B. M., RUBINSTEIN, N. D., AUTRY, A. E., MARIN-RODRIGUEZ, B., KAPOOR, V., MIYAMISHI, K., ZWEIFEL, L. S., LUO, L. & UCHIDA, N. 2018. Functional circuit architecture underlying parental behaviour. *Nature,* 556**,** 326-331.

KOHL, J. & DULAC, C. 2018. Neural control of parental behaviors. *Current opinion in neurobiology,* 49**,** 116-122.
KOLB, B. & NONNEMAN, A. J. 1974. Frontolimbic lesions and social behavior in the rat. *Physiology & Behavior,* 13**,** 637-643.
KÖLLIKER, M., SMISETH, P. T. & ROYLE, N. J. 2014. Evolution of parental care. *The Princeton Guide to Evolution,* 663-670.
KOVÁCS, K. 2008. Measurement of immediate-early gene activation-c-fos and beyond. *Journal of neuroendocrinology,* 20**,** 665-672.
KOZA, R. A., NIKONOVA, L., HOGAN, J., RIM, J.-S., MENDOZA, T., FAULK, C., SKAF, J. & KOZAK, L. P. 2006. Changes in gene expression foreshadow diet-induced obesity in genetically identical mice. *PLoS genetics,* 2**,** e81.
KOZLOV, S. V., BOGENPOHL, J. W., HOWELL, M. P., WEVRICK, R., PANDA, S., HOGENESCH, J. B., MUGLIA, L. J., VAN GELDER, R. N., HERZOG, E. D. & STEWART, C. L. 2007. The imprinted gene Magel2 regulates normal circadian output. *Nature genetics,* 39**,** 1266-1272.
KURODA, K. O., TACHIKAWA, K., YOSHIDA, S., TSUNEOKA, Y. & NUMAN, M. 2011. Neuromolecular basis of parental behavior in laboratory mice and rats: with special emphasis on technical issues of using mouse genetics. *Progress in Neuro-Psychopharmacology and Biological Psychiatry,* 35**,** 1205-1231.
KUROIWA, Y., KANEKO-ISHINO, T., KAGITANI, F., KOHDA, T., LI, L.-L., TADA, M., SUZUKI, R., YOKOYAMA, M., SHIROISHI, T. & WAKANA, S. 1996. Peg3 imprinted gene on proximal chromosome 7 encodes for a zinc finger protein. *Nature genetics,* 12**,** 186-190.
LA MANNO, G., GYLLBORG, D., CODELUPPI, S., NISHIMURA, K., SALTO, C., ZEISEL, A., BORM, L. E., STOTT, S. R., TOLEDO, E. M. & VILLAESCUSA, J. C. 2016. Molecular diversity of midbrain development in mouse, human, and stem cells. *Cell,* 167**,** 566-580. e19.
LARSEN, C. M. & GRATTAN, D. 2012. Prolactin, neurogenesis, and maternal behaviors. *Brain, behavior, and immunity,* 26**,** 201-209.
LASSI, G., BALL, S. T., MAGGI, S., COLONNA, G., NIEUS, T., CERO, C., BARTOLOMUCCI, A., PETERS, J. & TUCCI, V. 2012. Loss of Gnas imprinting differentially affects REM/NREM sleep and cognition in mice. *PLoS Genet,* 8**,** e1002706.
LASSI, G., PRIANO, L., MAGGI, S., GARCIA-GARCIA, C., BALZANI, E., EL-ASSAWY, N., PAGANI, M., TINARELLI, F., GIARDINO, D. & MAURO, A. 2016. Deletion of the Snord116/SNORD116 alters sleep in mice and patients with Prader-Willi syndrome. *Sleep,* 39**,** 637-644.
LAUKOTER, S., PAULER, F. M., BEATTIE, R., AMBERG, N., HANSEN, A. H., STREICHER, C., PENZ, T., BOCK, C. & HIPPENMEYER, S. 2020. Cell-type specificity of genomic imprinting in cerebral cortex. *Neuron,* 107**,** 1160-1179. e9.
LAWTON, B. R., SEVIGNY, L., OBERGFELL, C., REZNICK, D., O'NEILL, R. J. & O'NEILL, M. J. 2005. Allelic expression of IGF2 in live-bearing, matrotrophic fishes. *Development genes and evolution,* 215**,** 207-212.
LEE, A., CLANCY, S. & FLEMING, A. S. 1999. Mother rats bar-press for pups: effects of lesions of the mpoa and limbic sites on maternal behavior and operant responding for pup-reinforcement. *Behavioural brain research,* 100**,** 15-31.
LEE, S., KOZLOV, S., HERNANDEZ, L., CHAMBERLAIN, S. J., BRANNAN, C. I., STEWART, C. L. & WEVRICK, R. 2000. Expression and imprinting of MAGEL2 suggest a role in Prader–Willi syndrome and the homologous murine imprinting phenotype. *Human Molecular Genetics,* 9**,** 1813-1819.

LEFEBVRE, L., VIVILLE, S., BARTON, S. C., ISHINO, F., KEVERNE, E. B. & SURANI, M. A. 1998. Abnormal maternal behaviour and growth retardation associated with loss of the imprinted gene Mest. *Nature genetics,* 20**,** 163-169.

LEIN, E. S., HAWRYLYCZ, M. J., AO, N., AYRES, M., BENSINGER, A., BERNARD, A., BOE, A. F., BOGUSKI, M. S., BROCKWAY, K. S. & BYRNES, E. J. 2007. Genome-wide atlas of gene expression in the adult mouse brain. *Nature,* 445**,** 168-176.

LERCH-HANER, J. K., FRIERSON, D., CRAWFORD, L. K., BECK, S. G. & DENERIS, E. S. 2008. Serotonergic transcriptional programming determines maternal behavior and offspring survival. *Nature neuroscience,* 11**,** 1001-1003.

LI, J.-Y., CHAI, B.-X., ZHANG, W., WANG, H. & MULHOLLAND, M. W. 2010. Expression of ankyrin repeat and suppressor of cytokine signaling box protein 4 (Asb-4) in proopiomelanocortin neurons of the arcuate nucleus of mice produces a hyperphagic, lean phenotype. *Endocrinology,* 151**,** 134-142.

LI, L., KEVERNE, E., APARICIO, S., ISHINO, F., BARTON, S. & SURANI, M. 1999. Regulation of maternal behavior and offspring growth by paternally expressed Peg3. *Science,* 284**,** 330-334.

LIN, S.-P., YOUNGSON, N., TAKADA, S., SEITZ, H., REIK, W., PAULSEN, M., CAVAILLE, J. & FERGUSON-SMITH, A. C. 2003. Asymmetric regulation of imprinting on the maternal and paternal chromosomes at the Dlk1-Gtl2 imprinted cluster on mouse chromosome 12. *Nature genetics,* 35**,** 97-102.

LIU, H.-X., LOPATINA, O., HIGASHIDA, C., FUJIMOTO, H., AKTHER, S., INZHUTOVA, A., LIANG, M., ZHONG, J., TSUJI, T. & YOSHIHARA, T. 2013. Displays of paternal mouse pup retrieval following communicative interaction with maternal mates. *Nature communications,* 4**,** 1-8.

LORENC, A., LINNENBRINK, M., MONTERO, I., SCHILHABEL, M. B. & TAUTZ, D. 2014. Genetic differentiation of hypothalamus parentally biased transcripts in populations of the house mouse implicate the Prader–Willi syndrome imprinted region as a possible source of behavioral divergence. *Molecular biology and evolution,* 31**,** 3240-3249.

LUI, J. C., FINKIELSTAIN, G. P., BARNES, K. M. & BARON, J. 2008. An imprinted gene network that controls mammalian somatic growth is down-regulated during postnatal growth deceleration in multiple organs. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology,* 295**,** R189-R196.

LUKAS, D. & HUCHARD, E. 2014. The evolution of infanticide by males in mammalian societies. *Science,* 346**,** 841-844.

LUO, Y.-W., XU, Y., CAO, W.-Y., ZHONG, X.-L., DUAN, J., WANG, X.-Q., HU, Z.-L., LI, F., ZHANG, J.-Y. & ZHOU, M. 2015. Insulin-like growth factor 2 mitigates depressive behavior in a rat model of chronic stress. *Neuropharmacology,* 89**,** 318-324.

MACDONALD, W. A. 2012. Epigenetic mechanisms of genomic imprinting: common themes in the regulation of imprinted regions in mammals, plants, and insects. *Genetics research international,* 2012.

MACKAY, D. J. & TEMPLE, I. K. 2017. Human imprinting disorders: principles, practice, problems and progress. *European journal of medical genetics,* 60**,** 618-626.

MADON-SIMON, M., COWLEY, M., GARFIELD, A. S., MOORWOOD, K., BAUER, S. R. & WARD, A. 2014. Antagonistic roles in fetal development and adult physiology for the oppositely imprinted Grb10 and Dlk1genes. *BMC biology,* 12**,** 1-23.

MARTÍN-SÁNCHEZ, A., VALERA-MARÍN, G., HERNÁNDEZ-MARTÍNEZ, A., LANUZA, E., MARTÍNEZ-GARCÍA, F. & AGUSTÍN-PAVÓN, C. 2015. Wired for

motherhood: induction of maternal care but not maternal aggression in virgin female CD1 mice. *Frontiers in behavioral neuroscience,* 9**,** 197.
MARTINET, C., MONNIER, P., LOUAULT, Y., BENARD, M., GABORY, A. & DANDOLO, L. 2016. H19 controls reactivation of the imprinted gene network during muscle regeneration. *Development,* 143**,** 962-971.
MATHIESON, W., WILKINSON, M., BROWN, R. E., BOND, T. L., TAYLOR, S. W. & NEUMANN, P. E. 2002. FOS and FOSB expression in the medial preoptic nucleus pars compacta of maternally active C57BL/6J and DBA/2J mice. *Brain research,* 952**,** 170-175.
MATSUSHITA, N., MUROI, Y., KINOSHITA, K.-I. & ISHII, T. 2015. Comparison of c-Fos expression in brain regions involved in maternal behavior of virgin and lactating female mice. *Neuroscience letters,* 590**,** 166-171.
MATSUURA, K. 2020. Genomic imprinting and evolution of insect societies. *Population Ecology,* 62**,** 38-52.
MATTSON, B. & MORRELL, J. 2005. Preference for cocaine-versus pup-associated cues differentially activates neurons expressing either Fos or cocaine-and amphetamine-regulated transcript in lactating, maternal rodents. *Neuroscience,* 135**,** 315-328.
MCGRATH, J. & SOLTER, D. 1984. Completion of mouse embryogenesis requires both the maternal and paternal genomes. *Cell,* 37**,** 179-183.
MCNAMARA, G., CREETH, H., HARRISON, D., TANSEY, K., ANDREWS, R., ISLES, A. & JOHN, R. 2018a. Loss of offspring Peg3 reduces neonatal ultrasonic vocalizations and increases maternal anxiety in wild-type mothers. *Human molecular genetics,* 27**,** 440-450.
MCNAMARA, G. I., DAVIS, B. A., BROWNE, M., HUMBY, T., DALLEY, J. W., XIA, J., JOHN, R. M. & ISLES, A. R. 2018b. Dopaminergic and behavioural changes in a loss-of-imprinting model of Cdkn1c. *Genes, Brain and Behavior,* 17**,** 149-157.
MERCER, R. E., KWOLEK, E. M., BISCHOF, J. M., VAN EEDE, M., HENKELMAN, R. M. & WEVRICK, R. 2009. Regionally reduced brain volume, altered serotonin neurochemistry, and abnormal behavior in mice null for the circadian rhythm output gene Magel2. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics,* 150**,** 1085-1099.
MERCER, R. E. & WEVRICK, R. 2009. Loss of magel2, a candidate gene for features of Prader-Willi syndrome, impairs reproductive function in mice. *PLoS One,* 4**,** e4291.
MERCER, T. R., GERHARDT, D. J., DINGER, M. E., CRAWFORD, J., TRAPNELL, C., JEDDELOH, J. A., MATTICK, J. S. & RINN, J. L. 2012. Targeted RNA sequencing reveals the deep complexity of the human transcriptome. *Nature biotechnology,* 30**,** 99-104.
MESSERSCHMIDT, D. 2012. Should I stay or should I go: protection and maintenance of DNA methylation at imprinted genes. *Epigenetics,* 7**,** 969-975.
MEZIANE, H., SCHALLER, F., BAUER, S., VILLARD, C., MATARAZZO, V., RIET, F., GUILLON, G., LAFITTE, D., DESARMENIEN, M. G. & TAUBER, M. 2015. An early postnatal oxytocin treatment prevents social and learning deficits in adult mice deficient for Magel2, a gene involved in Prader-Willi syndrome and autism. *Biological psychiatry,* 78**,** 85-94.
MILLER, J. L., GOLDSTONE, A. P., COUCH, J. A., SHUSTER, J., HE, G., DRISCOLL, D. J., LIU, Y. & SCHMALFUSS, I. M. 2008. Pituitary abnormalities in Prader–Willi syndrome and early onset morbid obesity. *American Journal of Medical Genetics Part A,* 146**,** 570-577.

MILLERSHIP, S. J., VAN DE PETTE, M. & WITHERS, D. J. 2019. Genomic imprinting and its effects on postnatal growth and adult metabolism. *Cellular and Molecular Life Sciences,* 76**,** 4009-4021.
MOFFITT, J. R., BAMBAH-MUKKU, D., EICHHORN, S. W., VAUGHN, E., SHEKHAR, K., PEREZ, J. D., RUBINSTEIN, N. D., HAO, J., REGEV, A. & DULAC, C. 2018. Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. *Science,* 362**,** eaau5324.
MONK, D. 2015. Genomic imprinting in the human placenta. *American journal of obstetrics and gynecology,* 213**,** S152-S162.
MONK, D., MACKAY, D. J., EGGERMANN, T., MAHER, E. R. & RICCIO, A. 2019. Genomic imprinting disorders: lessons on how genome, epigenome and environment interact. *Nature Reviews Genetics,* 20**,** 235-248.
MONTALBÁN-LORO, R., LASSI, G., LOZANO-UREÑA, A., PEREZ-VILLALBA, A., JIMÉNEZ-VILLALBA, E., CHARALAMBOUS, M., VALLORTIGARA, G., HORNER, A. E., SAKSIDA, L. M. & BUSSEY, T. J. 2021. Dlk1 dosage regulates hippocampal neurogenesis and cognition. *Proceedings of the National Academy of Sciences,* 118**,** e2015505118.
MOORE, T. & HAIG, D. 1991. Genomic imprinting in mammalian development: a parental tug-of-war. *Trends in genetics,* 7**,** 45-49.
MORTAZAVI, A., WILLIAMS, B. A., MCCUE, K., SCHAEFFER, L. & WOLD, B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature methods,* 5**,** 621-628.
NAMBA, T., TANIGUCHI, M., MURATA, Y., TONG, J., WANG, Y., OKUTANI, F., YAMAGUCHI, M. & KABA, H. 2016. Activation of arginine vasopressin receptor 1a facilitates the induction of long-term potentiation in the accessory olfactory bulb of male mice. *Neuroscience Letters,* 634**,** 107-113.
NEGI, S. K. & GUDA, C. 2017. Global gene expression profiling of healthy human brain and its application in studying neurological disorders. *Scientific reports,* 7**,** 1-12.
NUMAN, M. 1974. Medial preoptic area and maternal behavior in the female rat. *Journal of comparative and physiological psychology,* 87**,** 746.
NUMAN, M. 1988. Neural basis of maternal behavior in the rat. *Psychoneuroendocrinology,* 13**,** 47-62.
NUMAN, M. 2017. Parental behavior.
NUMAN, M. & CORODIMAS, K. P. 1985. The effects of paraventricular hypothalamic lesions on maternal behavior in rats. *Physiology & behavior,* 35**,** 417-425.
NUMAN, M. & INSEL, T. R. 2003. Paternal behavior. *The Neurobiology of Parental Behavior***,** 246-267.
NUMAN, M. & NUMAN, M. J. 1994. Expression of Fos-like immunoreactivity in the preoptic area of maternally behaving virgin and postpartum rats. *Behavioral neuroscience,* 108**,** 379.
NUMAN, M., ROSENBLATT, J. S. & KOMISARUK, B. R. 1977. Medial preoptic area and onset of maternal behavior in the rat. *Journal of comparative and physiological psychology,* 91**,** 146.
NUMAN, M. & SMITH, H. G. 1984. Maternal behavior in rats: evidence for the involvement of preoptic projections to the ventral tegmental area. *Behavioral neuroscience,* 98**,** 712.
NUMAN, M. & STOLZENBERG, D. S. 2009. Medial preoptic area interactions with dopamine neural systems in the control of the onset and maintenance of maternal behavior in rats. *Frontiers in neuroendocrinology,* 30**,** 46-64.

O'BRIEN, E. K. & WOLF, J. B. 2017. The coadaptation theory for genomic imprinting. *Evolution letters,* 1**,** 49-59.
O'HIGGINS, M., ROBERTS, I. S. J., GLOVER, V. & TAYLOR, A. 2013. Mother-child bonding at 1 year; associations with symptoms of postnatal depression and bonding in the first few weeks. *Archives of women's mental health,* 16**,** 381-389.
OKAE, H., HIURA, H., NISHIDA, Y., FUNAYAMA, R., TANAKA, S., CHIBA, H., YAEGASHI, N., NAKAYAMA, K., SASAKI, H. & ARIMA, T. 2012. Re-investigation and RNA sequencing-based identification of genes with placenta-specific imprinted expression. *Human molecular genetics,* 21**,** 548-558.
OLAZÁBAL, D., ABERCROMBIE, E., ROSENBLATT, J. & MORRELL, J. 2004. The content of dopamine, serotonin, and their metabolites in the neural circuit that mediates maternal behavior in juvenile and adult rats. *Brain research bulletin,* 63**,** 259-268.
ONO, R., NAKAMURA, K., INOUE, K., NARUSE, M., USAMI, T., WAKISAKA-SAITO, N., HINO, T., SUZUKI-MIGISHIMA, R., OGONUKI, N. & MIKI, H. 2006. Deletion of Peg10, an imprinted gene acquired from a retrotransposon, causes early embryonic lethality. *Nature genetics,* 38**,** 101-106.
OTTO, S. P. & GERSTEIN, A. C. 2008. The evolution of haploidy and diploidy. *Current Biology,* 18**,** R1121-R1124.
PANDA, S., ANTOCH, M. P., MILLER, B. H., SU, A. I., SCHOOK, A. B., STRAUME, M., SCHULTZ, P. G., KAY, S. A., TAKAHASHI, J. S. & HOGENESCH, J. B. 2002. Coordinated transcription of key pathways in the mouse by the circadian clock. *Cell,* 109**,** 307-320.
PATTEN, M. M., COWLEY, M., OAKEY, R. J. & FEIL, R. 2016. Regulatory links between imprinted genes: evolutionary predictions and consequences. *Proceedings of the Royal Society B: Biological Sciences,* 283**,** 20152760.
PAULO, E., WU, D., WANG, Y., ZHANG, Y., WU, Y., SWANEY, D. L., SOUCHERAY, M., JIMENEZ-MORALES, D., CHAWLA, A. & KROGAN, N. J. 2018. Sympathetic inputs regulate adaptive thermogenesis in brown adipose tissue through cAMP-Salt inducible kinase axis. *Scientific reports,* 8**,** 1-14.
PECK, L. J., PATEL, R., DIAZ, P., WINTLE, Y. M., DICKENSON, A. H., TODD, A. J., CALVO, M. & BENNETT, D. L. 2021. Studying independent Kcna6 knock-out mice reveals toxicity of exogenous LacZ to central nociceptor terminals and differential effects of Kv1. 6 on acute and neuropathic pain sensation. *Journal of Neuroscience,* 41**,** 9141-9162.
PEREZ, J. D., RUBINSTEIN, N. D. & DULAC, C. 2016. New perspectives on genomic imprinting, an essential and multifaceted mode of epigenetic control in the developing and adult brain. *Annual review of neuroscience,* 39**,** 347-384.
PEREZ, J. D., RUBINSTEIN, N. D., FERNANDEZ, D. E., SANTORO, S. W., NEEDLEMAN, L. A., HO-SHING, O., CHOI, J. J., ZIRLINGER, M., CHEN, S.-K. & LIU, J. S. 2015. Quantitative and functional interrogation of parent-of-origin allelic expression biases in the brain. *Elife,* 4**,** e07860.
PETERS, J. 2014. The role of genomic imprinting in biology and disease: an expanding view. *Nature Reviews Genetics,* 15**,** 517-530.
PICARD, C., SILVY, M., GERARD, C., BUFFAT, C., LAVAQUE, E., FIGARELLA-BRANGER, D., DUFOUR, H., GABERT, J., BECKERS, A. & BRUE, T. 2007. Gsα overexpression and loss of Gsα imprinting in human somatotroph adenomas: association with tumor size and response to pharmacologic treatment. *International journal of cancer,* 121**,** 1245-1252.

PIEDRAHITA, J. A. 2011. The role of imprinted genes in fetal growth abnormalities. *Birth Defects Research Part A: Clinical and Molecular Teratology,* 91**,** 682-692.
PIRES, N. D. & GROSSNIKLAUS, U. 2014. Different yet similar: evolution of imprinting in flowering plants and mammals. *F1000prime reports,* 6.
PLAGGE, A., GORDON, E., DEAN, W., BOIANI, R., CINTI, S., PETERS, J. & KELSEY, G. 2004. The imprinted signaling protein XLαs is required for postnatal adaptation to feeding. *Nature genetics,* 36**,** 818-826.
PLAGGE, A., ISLES, A. R., GORDON, E., HUMBY, T., DEAN, W., GRITSCH, S., FISCHER-COLBRIE, R., WILKINSON, L. S. & KELSEY, G. 2005. Imprinted Nesp55 influences behavioral reactivity to novel environments. *Molecular and cellular biology,* 25**,** 3019-3026.
PLASSCHAERT, R. N. & BARTOLOMEI, M. S. 2014. Genomic imprinting in development, growth, behavior and stem cells. *Development,* 141**,** 1805-1813.
PORTFORS, C. V. & PERKEL, D. J. 2014. The role of ultrasonic vocalizations in mouse communication. *Current opinion in neurobiology,* 28**,** 115-120.
PULIX, M. & PLAGGE, A. 2020. Imprinted Genes and Hypothalamic Function. *Developmental Neuroendocrinology.* Springer.
RAPKINS, R. W., HORE, T., SMITHWICK, M., AGER, E., PASK, A. J., RENFREE, M. B., KOHN, M., HAMEISTER, H., NICHOLLS, R. D. & DEAKIN, J. E. 2006. Recent assembly of an imprinted domain from non-imprinted components. *PLoS genetics,* 2**,** e182.
RAU, A. R. & HENTGES, S. T. 2017. The relevance of AgRP neuron-derived GABA inputs to POMC neurons differs for spontaneous and evoked release. *Journal of Neuroscience,* 37**,** 7362-7372.
REICHOVA, A., SCHALLER, F., BUKATOVA, S., BACOVA, Z., MUSCATELLI, F. & BAKOS, J. 2021. The impact of oxytocin on neurite outgrowth and synaptic proteins in Magel2-deficient mice. *Developmental neurobiology,* 81**,** 366-388.
REIK, W., COLLICK, A., NORRIS, M. L., BARTON, S. C. & SURANI, M. A. 1987. Genomic imprinting determines methylation of parental alleles in transgenic mice. *Nature,* 328**,** 248-251.
REIK, W., DEAN, W. & WALTER, J. 2001. Epigenetic reprogramming in mammalian development. *Science,* 293**,** 1089-1093.
RENFREE, M. B., HORE, T. A., SHAW, G., MARSHALL GRAVES, J. A. & PASK, A. J. 2009. Evolution of genomic imprinting: insights from marsupials and monotremes. *Annual review of genomics and human genetics,* 10**,** 241-262.
RIBEIRO, A. C., MUSATOV, S., SHTEYLER, A., SIMANDUYEV, S., ARRIETA-CRUZ, I., OGAWA, S. & PFAFF, D. W. 2012. siRNA silencing of estrogen receptor-α expression specifically in medial preoptic area neurons abolishes maternal care in female mice. *Proceedings of the National Academy of Sciences,* 109**,** 16324-16329.
RIEDMAN, M. L. 1982. The evolution of alloparental care and adoption in mammals and birds. *The Quarterly review of biology,* 57**,** 405-435.
RIENECKER, K. D., CHAVASSE, A. T., MOORWOOD, K., WARD, A. & ISLES, A. R. 2018. Detailed analysis of paternal knockout Grb10 mice suggests effects on social stability, rather than social dominance. *bioRxiv***,** 493692.
RILLING, J. K. & YOUNG, L. J. 2014. The biology of mammalian parenting and its effect on offspring social development. *Science,* 345**,** 771-776.
ROGERS, F. D. & BALES, K. L. 2019. Mothers, fathers, and others: neural substrates of parental care. *Trends in neurosciences,* 42**,** 552-562.
ROMANOV, R. A., ZEISEL, A., BAKKER, J., GIRACH, F., HELLYSAZ, A., TOMER, R., ALPAR, A., MULDER, J., CLOTMAN, F. & KEIMPEMA, E. 2017. Molecular

interrogation of hypothalamic organization reveals distinct dopamine neuronal subtypes. *Nature neuroscience,* 20**,** 176-188.
ROSENBLATT, J. S. & CEUS, K. 1998. Estrogen implants in the medial preoptic area stimulate maternal behavior in male rats. *Hormones and behavior,* 33**,** 23-30.
ROUSMANIERE, H., SILVERMAN, R., WHITE, R. A., SASAKI, M. M., WILSON, S. D., MORRISON, J. T. & CRUZ, Y. P. 2010. Husbandry of Monodelphis domestica in the study of mammalian embryogenesis. *Lab animal,* 39**,** 219-226.
RYMER, T. L. & PILLAY, N. 2018. An integrated understanding of paternal care in mammals: Lessons from the rodents. *Journal of Zoology,* 306**,** 69-76.
SAPER, C. B. & LOWELL, B. B. 2014. The hypothalamus. *Current Biology,* 24**,** R1111-R1116.
SCAGLIOTTI, V., COSTA FERNANDES ESSE, R., WILLIS, T. L., HOWARD, M., CARRUS, I., LODGE, E., ANDONIADOU, C. L. & CHARALAMBOUS, M. 2021. Dynamic Expression of Imprinted Genes in the Developing and Postnatal Pituitary Gland. *Genes,* 12**,** 509.
SCHAAF, C. P., GONZALEZ-GARAY, M. L., XIA, F., POTOCKI, L., GRIPP, K. W., ZHANG, B., PETERS, B. A., MCELWAIN, M. A., DRMANAC, R. & BEAUDET, A. L. 2013. Truncating mutations of MAGEL2 cause Prader-Willi phenotypes and autism. *Nature genetics,* 45**,** 1405-1408.
SCHALLER, F., WATRIN, F., STURNY, R., MASSACRIER, A., SZEPETOWSKI, P. & MUSCATELLI, F. 2010. A single postnatal injection of oxytocin rescues the lethal feeding behaviour in mouse newborns deficient for the imprinted Magel2 gene. *Human Molecular Genetics,* 19**,** 4895-4905.
SCHAUM, N., KARKANIAS, J., NEFF, N. F., MAY, A. P., QUAKE, S. R., WYSS-CORAY, T., DARMANIS, S., BATSON, J., BOTVINNIK, O. & CHEN, M. B. 2018. Single-cell transcriptomic characterization of 20 organs and tissues from individual mice creates a tabula muris. *BioRxiv***,** 237446.
SCHEIBER, I. B., WEIß, B. M., KINGMA, S. A. & KOMDEUR, J. 2017. The importance of the altricial–precocial spectrum for social complexity in mammals and birds–a review. *Frontiers in zoology,* 14**,** 1-20.
SCHULZ, R., MENHENIOTT, T. R., WOODFINE, K., WOOD, A. J., CHOI, J. D. & OAKEY, R. J. 2006. Chromosome-wide identification of novel imprinted genes using microarrays and uniparental disomies. *Nucleic acids research,* 34**,** e88-e88.
SCHULZE, A., MOGENSEN, H., HAMBORG-PETERSEN, B., GRAEM, N., ØSTERGAARD, J. & BRØNDUM-NIELSEN, K. 2001. Fertility in Prader-Willi syndrome: a case report with Angelman syndrome in the offspring. *Acta paediatrica,* 90**,** 455-459.
SCOTT, R. J. & SPIELMAN, M. 2006. Genomic imprinting in plants and mammals: how life history constrains convergence. *Cytogenetic and genome research,* 113**,** 53-67.
SEISENBERGER, S., ANDREWS, S., KRUEGER, F., ARAND, J., WALTER, J., SANTOS, F., POPP, C., THIENPONT, B., DEAN, W. & REIK, W. 2012. The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. *Molecular cell,* 48**,** 849-862.
SFERRUZZI-PERRI, A., VAUGHAN, O., COAN, P., SUCIU, M., DARBYSHIRE, R., CONSTANCIA, M., BURTON, G. & FOWDEN, A. 2011. Placental-specific Igf2 deficiency alters developmental adaptations to undernutrition in mice. *Endocrinology,* 152**,** 3202-3212.
SHINOZUKA, K., YANO-NASHIMOTO, S., YOSHIHARA, C., TOKITA, K., KURACHI, T., MATSUI, R., WATANABE, D., INOUE, K.-I., TAKADA, M. & MORIYA-ITO, K. 2022. A calcitonin receptor-expressing subregion of the medial preoptic area is

involved in alloparental tolerance in common marmosets. *Communications biology,* 5**,** 1-18.
SIBLEY, C., COAN, P., FERGUSON-SMITH, A., DEAN, W., HUGHES, J., SMITH, P., REIK, W., BURTON, G., FOWDEN, A. & CONSTANCIA, M. 2004. Placental-specific insulin-like growth factor 2 (Igf2) regulates the diffusional exchange characteristics of the mouse placenta. *Proceedings of the National Academy of Sciences,* 101**,** 8204-8208.
SIBLEY, C. P., TURNER, M. A., CETIN, I., AYUK, P., BOYD, C., D'SOUZA, S. W., GLAZIER, J. D., GREENWOOD, S. L., JANSSON, T. & POWELL, T. 2005. Placental phenotypes of intrauterine growth. *Pediatric research,* 58**,** 827-832.
SMALLWOOD, S. A. & KELSEY, G. 2012. De novo DNA methylation: a germ cell perspective. *Trends in Genetics,* 28**,** 33-42.
SMILEY, K. O., BROWN, R. S. & GRATTAN, D. R. 2022. Prolactin action is necessary for parental behavior in male mice. *Journal of Neuroscience,* 42**,** 8308-8327.
SMITH, F., GARFIELD, A. & WARD, A. 2006. Regulation of growth and metabolism by imprinted genes. *Cytogenetic and genome research,* 113**,** 279-291.
SMITH, S. E., ZHOU, Y.-D., ZHANG, G., JIN, Z., STOPPEL, D. C. & ANDERSON, M. P. 2011. Increased gene dosage of Ube3a results in autism traits and decreased glutamate synaptic transmission in mice. *Science translational medicine,* 3**,** 103ra97-103ra97.
ST-PIERRE, J., HIVERT, M.-F., PERRON, P., POIRIER, P., GUAY, S.-P., BRISSON, D. & BOUCHARD, L. 2012. IGF2 DNA methylation is a modulator of newborn's fetal growth and development. *Epigenetics,* 7**,** 1125-1132.
ST JAMES-ROBERTS, I., CONROY, S. & WILSHER, K. 1998. Links between maternal care and persistent infant crying in the early months. *Child: care, health and development,* 24**,** 353-376.
STACK, E. C., BALAKRISHNAN, R., NUMAN, M. J. & NUMAN, M. 2002. A functional neuroanatomical investigation of the role of the medial preoptic area in neural circuits regulating maternal behavior. *Behavioural brain research,* 131**,** 17-36.
STACK, E. C. & NUMAN, M. 2000. The temporal course of expression of c-Fos and Fos B within the medial preoptic area and other brain regions of postpartum female rats during prolonged mother–young interactions. *Behavioral neuroscience,* 114**,** 609.
STAGKOURAKIS, S., SMILEY, K. O., WILLIAMS, P., KAKADELLIS, S., ZIEGLER, K., BAKKER, J., BROWN, R. S., HARKANY, T., GRATTAN, D. R. & BROBERGER, C. 2020. A neuro-hormonal circuit for paternal behavior controlled by a hypothalamic network oscillation. *Cell,* 182**,** 960-975. e15.
STEINHOFF, C., PAULSEN, M., KIELBASA, S., WALTER, J. & VINGRON, M. 2009. Expression profile and transcription factor binding site exploration of imprinted genes in human and mouse. *BMC genomics,* 10**,** 1-15.
STELZER, Y., SAGI, I., YANUKA, O., EIGES, R. & BENVENISTY, N. 2014. The noncoding RNA IPW regulates the imprinted DLK1-DIO3 locus in an induced pluripotent stem cell model of Prader-Willi syndrome. *Nature genetics,* 46**,** 551-557.
STEUERNAGEL, L., LAM, B. Y., KLEMM, P., DOWSETT, G. K., BAUDER, C. A., TADROSS, J. A., HITSCHFELD, T. S., DEL RIO MARTIN, A., CHEN, W. & DE SOLIS, A. J. 2022. HypoMap—a unified single-cell gene expression atlas of the murine hypothalamus. *Nature Metabolism,* 4**,** 1402-1419.
STEVENSON, E. L. & CALDWELL, H. K. 2012. The vasopressin 1b receptor and the neural regulation of social behavior. *Hormones and Behavior,* 61**,** 277-282.
STOHN, J. P., MARTINEZ, M. E., ZAFER, M., LÓPEZ-ESPÍNDOLA, D., KEYES, L. M. & HERNANDEZ, A. 2018. Increased aggression and lack of maternal behavior in

Dio3-deficient mice are associated with abnormalities in oxytocin and vasopressin systems. *Genes, Brain and Behavior,* 17**,** 23-35.
STOLZENBERG, D. S. & CHAMPAGNE, F. A. 2016. Hormonal and non-hormonal bases of maternal behavior: The role of experience and epigenetic mechanisms. *Hormones and behavior,* 77**,** 204-210.
STOLZENBERG, D. S. & RISSMAN, E. F. 2011. Oestrogen-independent, experience-induced maternal behaviour in female mice. *Journal of neuroendocrinology,* 23**,** 345-354.
STRINGER, J., PASK, A., SHAW, G. & RENFREE, M. B. 2014. Post-natal imprinting: evidence from marsupials. *Heredity,* 113**,** 145-155.
SUBRAMANIAN, A., TAMAYO, P., MOOTHA, V. K., MUKHERJEE, S., EBERT, B. L., GILLETTE, M. A., PAULOVICH, A., POMEROY, S. L., GOLUB, T. R. & LANDER, E. S. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences,* 102**,** 15545-15550.
SURANI, M., BARTON, S. C. & NORRIS, M. 1984. Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. *Nature,* 308**,** 548-550.
SURMACZ, B., NOISA, P., RISNER-JANICZEK, J. R., HUI, K., UNGLESS, M., CUI, W. & LI, M. 2012. DLK1 promotes neurogenesis of human and mouse pluripotent stem cell-derived neural progenitors via modulating Notch and BMP signalling. *Stem Cell Reviews and Reports,* 8**,** 459-471.
SUZUKI, S., SHAW, G., KANEKO-ISHINO, T., ISHINO, F. & RENFREE, M. B. 2011a. Characterisation of marsupial PHLDA2 reveals eutherian specific acquisition of imprinting. *BMC evolutionary biology,* 11**,** 1-9.
SUZUKI, S., SHAW, G., KANEKO-ISHINO, T., ISHINO, F. & RENFREE, M. B. 2011b. The evolution of mammalian genomic imprinting was accompanied by the acquisition of novel CpG islands. *Genome Biology and Evolution,* 3**,** 1276-1283.
SWANEY, W. T., CURLEY, J. P., CHAMPAGNE, F. A. & KEVERNE, E. B. 2008. The paternally expressed gene Peg3 regulates sexual experience-dependent preferences for estrous odors. *Behavioral neuroscience,* 122**,** 963.
SYMMONS, O., CHANG, M., MELLIS, I. A., KALISH, J. M., PARK, J., SUSZTÁK, K., BARTOLOMEI, M. S. & RAJ, A. 2019. Allele-specific RNA imaging shows that allelic imbalances can arise in tissues through transcriptional bursting. *PLoS genetics,* 15**,** e1007874.
TACER, K. F. & POTTS, P. R. 2017. Cellular and disease functions of the Prader–Willi syndrome gene MAGEL2. *Biochemical Journal,* 474**,** 2177-2190.
TACHIKAWA, K. S., YOSHIHARA, Y. & KURODA, K. O. 2013. Behavioral transition from attack to parenting in male mice: a crucial role of the vomeronasal system. *Journal of Neuroscience,* 33**,** 5120-5126.
TEAM, R. 2015. RStudio: integrated development for R. *RStudio, Inc., Boston, MA URL http://www. rstudio. com,* 42**,** 14.
TEAM, R. C. 2013. R: A language and environment for statistical computing. Vienna, Austria.
TENNESE, A. A. & WEVRICK, R. 2011. Impaired hypothalamic regulation of endocrine function and delayed counterregulatory response to hypoglycemia in Magel2-null mice. *Endocrinology,* 152**,** 967-978.
TIMSHEL, P. N., THOMPSON, J. J. & PERS, T. H. 2020. Genetic mapping of etiologic brain cell types for obesity. *Elife,* 9.
TRIVERS, R. & BURT, A. 1999. Kinship and genomic imprinting. *Genomic imprinting.* Springer.

TSAI, C.-E., LIN, S.-P., ITO, M., TAKAGI, N., TAKADA, S. & FERGUSON-SMITH, A. C. 2002. Genomic imprinting contributes to thyroid hormone metabolism in the mouse embryo. *Current Biology,* 12**,** 1221-1226.

TUCCI, V. 2016. Genomic imprinting: a new epigenetic perspective of sleep regulation. *PLoS genetics,* 12**,** e1006004.

TUCCI, V., ISLES, A. R., KELSEY, G., FERGUSON-SMITH, A. C., BARTOLOMEI, M. S., BENVENISTY, N., BOURC'HIS, D., CHARALAMBOUS, M., DULAC, C. & FEIL, R. 2019. Genomic imprinting and physiological processes in mammals. *Cell,* 176**,** 952-965.

TUNSTER, S. J., BOQUÉ-SASTRE, R., MCNAMARA, G. I., HUNTER, S. M., CREETH, H. D. & JOHN, R. M. 2018. Peg3 deficiency results in sexually dimorphic losses and gains in the normal repertoire of placental hormones. *Frontiers in cell and developmental biology,* 6**,** 123.

TUNSTER, S. J., CREETH, H. & JOHN, R. M. 2016a. The imprinted Phlda2 gene modulates a major endocrine compartment of the placenta to regulate placental demands for maternal resources. *Developmental biology,* 409**,** 251-260.

TUNSTER, S. J., JENSEN, A. B. & JOHN, R. M. 2013. Imprinted genes in mouse placental development and the regulation of fetal energy stores. *Reproduction,* 145**,** R117-R137.

TUNSTER, S. J., MCNAMARA, G., CREETH, H. & JOHN, R. M. 2016b. Increased dosage of the imprinted Ascl2 gene restrains two key endocrine lineages of the mouse Placenta. *Developmental biology,* 418**,** 55-65.

TUNSTER, S. J., VAN DE PETTE, M. & JOHN, R. M. 2011. Fetal overgrowth in the Cdkn1c mouse model of Beckwith-Wiedemann syndrome. *Disease models & mechanisms,* 4**,** 814-821.

ÚBEDA, F. & GARDNER, A. 2015. Mother and offspring in conflict: why not? *PLoS Biology,* 13**,** e1002084.

VAGENA, E., CRNETA, J., ENGSTRÖM, P., HE, L., YULYANINGSIH, E., KORPEL, N. L., CHEANG, R. T., BACHOR, T. P., HUANG, A. & MICHEL, G. 2022. ASB4 modulates central melanocortinergic neurons and calcitonin signaling to control satiety and glucose homeostasis. *Science Signaling,* 15**,** eabj8204.

VAN DE PETTE, M., ABBAS, A., FEYTOUT, A., MCNAMARA, G., BRUNO, L., TO, W. K., DIMOND, A., SARDINI, A., WEBSTER, Z. & MCGINTY, J. 2017. Visualizing changes in Cdkn1c expression links early-life adversity to imprint mis-regulation in adults. *Cell reports,* 18**,** 1090-1099.

VAN DE PETTE, M., TUNSTER, S. J., MCNAMARA, G. I., SHELKOVNIKOVA, T., MILLERSHIP, S., BENSON, L., PEIRSON, S., CHRISTIAN, M., VIDAL-PUIG, A. & JOHN, R. M. 2016. Cdkn1c boosts the development of brown adipose tissue in a murine model of silver russell syndrome. *PLoS genetics,* 12**,** e1005916.

VARRAULT, A., GUEYDAN, C., DELALBRE, A., BELLMANN, A., HOUSSAMI, S., AKNIN, C., SEVERAC, D., CHOTARD, L., KAHLI, M. & LE DIGARCHER, A. 2006. Zac1 regulates an imprinted gene network critically involved in the control of embryonic growth. *Developmental cell,* 11**,** 711-722.

VIJAYAKUMAR, A., YAKAR, S. & LEROITH, D. 2011. The intricate role of growth hormone in metabolism. *Frontiers in endocrinology,* 2**,** 32.

VOM SAAL, F. S. & HOWARD, L. S. 1982. The regulation of infanticide and parental behavior: implications for reproductive success in male mice. *Science,* 215**,** 1270-1272.

WALLACE, M. L., BURETTE, A. C., WEINBERG, R. J. & PHILPOT, B. D. 2012. Maternal loss of Ube3a produces an excitatory/inhibitory imbalance through neuron type-specific synaptic defects. *Neuron,* 74**,** 793-800.
WAN, L. B. & BARTOLOMEI, M. S. 2008. Regulation of imprinting in clusters: noncoding RNAs versus insulators. *Advances in genetics,* 61**,** 207-223.
WANG, X., SOLOWAY, P. D. & CLARK, A. G. 2011. A survey for novel imprinted genes in the mouse placenta by mRNA-seq. *Genetics,* 189**,** 109-122.
WANG, X., SUN, Q., MCGRATH, S. D., MARDIS, E. R., SOLOWAY, P. D. & CLARK, A. G. 2008. Transcriptome-wide identification of novel imprinted genes in neonatal mouse brain. *PloS one,* 3**,** e3839.
WAXMAN, D. J. & FRANK, S. J. 2000. Growth hormone action. *Principles of Molecular Regulation.* Springer.
WEINBERG-SHUKRON, A., BEN-YAIR, R., TAKAHASHI, N., DUNJIĆ, M., SHTRIKMAN, A., EDWARDS, C. A., FERGUSON-SMITH, A. C. & STELZER, Y. 2022. Balanced gene dosage control rather than parental origin underpins genomic imprinting. *Nature communications,* 13**,** 1-12.
WEINSTEIN, L. S., XIE, T., QASEM, A., WANG, J. & CHEN, M. 2010. The role of GNAS and other imprinted genes in the development of obesity. *International journal of obesity,* 34**,** 6-17.
WEN, S. A., MA, D., ZHAO, M., XIE, L., WU, Q., GOU, L., ZHU, C., FAN, Y., WANG, H. & YAN, J. 2020. Spatiotemporal single-cell analysis of gene expression in the mouse suprachiasmatic nucleus. *Nature neuroscience,* 23**,** 456-467.
WHIPPLE, A. J., BRETON-PROVENCHER, V., JACOBS, H. N., CHITTA, U. K., SUR, M. & SHARP, P. A. 2020. Imprinted maternally expressed microRNAs antagonize paternally driven gene programs in neurons. *Molecular cell,* 78**,** 85-95. e8.
WILKINS, J. F. 2014. Genomic imprinting of Grb10: coadaptation or conflict? *PLoS biology,* 12**,** e1001800.
WILKINS, J. F. & HAIG, D. 2003. Inbreeding, maternal care and genomic imprinting. *Journal of Theoretical Biology,* 221**,** 559-564.
WILKINS, J. F., ÚBEDA, F. & VAN CLEVE, J. 2016. The evolving landscape of imprinted genes in humans and mice: conflict among alleles, genes, tissues, and kin. *Bioessays,* 38**,** 482-489.
WILLIAMS, G. C. 1966. Natural selection, the costs of reproduction, and a refinement of Lack's principle. *The American Naturalist,* 100**,** 687-690.
WILLIAMSON, C. M., TURNER, M. D., BALL, S. T., NOTTINGHAM, W. T., GLENISTER, P., FRAY, M., TYMOWSKA-LALANNE, Z., PLAGGE, A., POWLES-GLOVER, N. & KELSEY, G. 2006. Identification of an imprinting control region affecting the expression of all transcripts in the Gnas cluster. *Nature genetics,* 38**,** 350-355.
WIRTSHAFTER, H. S. & WILSON, M. A. 2021. Lateral septum as a nexus for mood, motivation, and movement. *Neuroscience & Biobehavioral Reviews,* 126**,** 544-559.
WOLF, J. B. 2013. Evolution of genomic imprinting as a coordinator of coadapted gene expression. *Proceedings of the National Academy of Sciences,* 110**,** 5085-5090.
WOLF, J. B., COWLEY, M. & WARD, A. 2015. Coadaptation between mother and offspring: why not? *PLoS biology,* 13**,** e1002085.
WOLF, J. B. & HAGER, R. 2006. A maternal–offspring coadaptation theory for the evolution of genomic imprinting. *PLoS biology,* 4**,** e380.
WU, Z., AUTRY, A. E., BERGAN, J. F., WATABE-UCHIDA, M. & DULAC, C. G. 2014. Galanin neurons in the medial preoptic area govern parental behaviour. *Nature,* 509**,** 325-330.

WUTZ, A., SMRZKA, O. W., SCHWEIFER, N., SCHELLANDER, K., WAGNER, E. F. & BARLOW, D. P. 1997. Imprinted expression of the Igf2r gene depends on an intronic CpG island. *Nature,* 389**,** 745-749.

WUTZ, A., THEUSSL, H., DAUSMAN, J., JAENISCH, R., BARLOW, D. & WAGNER, E. 2001. Non-imprinted Igf2r expression decreases growth and rescues the Tme mutation in mice. *Development,* 128**,** 1881-1887.

XIMERAKIS, M., LIPNICK, S. L., INNES, B. T., SIMMONS, S. K., ADICONIS, X., DIONNE, D., MAYWEATHER, B. A., NGUYEN, L., NIZIOLEK, Z. & OZEK, C. 2019. Single-cell transcriptomic profiling of the aging mouse brain. *Nature neuroscience,* 22**,** 1696-1708.

YEVTODIYENKO, A. & SCHMIDT, J. V. 2006. Dlk1 expression marks developing endothelium and sites of branching morphogenesis in the mouse embryo and placenta. *Developmental dynamics: an official publication of the American Association of Anatomists,* 235**,** 1115-1123.

YOKOI, F., DANG, M. T., LI, J. & LI, Y. 2006. Myoclonus, motor deficits, alterations in emotional responses and monoamine metabolism in ε-sarcoglycan deficient mice. *Journal of biochemistry,* 140**,** 141-146.

YONG, W.-S., HSU, F.-M. & CHEN, P.-Y. 2016. Profiling genome-wide DNA methylation. *Epigenetics & chromatin,* 9**,** 1-16.

YOSHIHARA, C., NUMAN, M. & KURODA, K. O. 2017. Oxytocin and parental behaviors. *Behavioral Pharmacology of Neuropeptides: Oxytocin***,** 119-153.

YOSHIHARA, C., TOKITA, K., MARUYAMA, T., KANEKO, M., TSUNEOKA, Y., FUKUMITSU, K., MIYAZAWA, E., SHINOZUKA, K., HUANG, A. J. & NISHIMORI, K. 2021. Calcitonin receptor signaling in the medial preoptic area enables risk-taking maternal care. *Cell Reports,* 35**,** 109204.

YU, S., YU, D., LEE, E., ECKHAUS, M., LEE, R., CORRIA, Z., ACCILI, D., WESTPHAL, H. & WEINSTEIN, L. S. 1998. Variable and tissue-specific hormone resistance in heterotrimeric Gs protein α-subunit (Gsα) knockout mice is due to tissue-specific imprinting of the Gsα gene. *Proceedings of the National Academy of Sciences,* 95**,** 8715-8720.

ZEISEL, A., HOCHGERNER, H., LÖNNERBERG, P., JOHNSSON, A., MEMIC, F., VAN DER ZWAN, J., HÄRING, M., BRAUN, E., BORM, L. E. & LA MANNO, G. 2018. Molecular architecture of the mouse nervous system. *Cell,* 174**,** 999-1014. e22.

ZHANG, X., JIANG, S., MITOK, K. A., LI, L., ATTIE, A. D. & MARTIN, T. F. 2017. BAIAP3, a C2 domain–containing Munc13 protein, controls the fate of dense-core vesicles in neuroendocrine cells. *Journal of Cell Biology,* 216**,** 2151-2166.

ZHONG, J., LIANG, M., AKTHER, S., HIGASHIDA, C., TSUJI, T. & HIGASHIDA, H. 2014. c-Fos expression in the paternal mouse brain induced by communicative interaction with maternal mates. *Molecular brain,* 7**,** 1-11.

A Appendices

Appendix A1 –Dataset specific information and workflow specifics (Chapter 3 and Appendix A2)

1) Multi-Organ level

Mouse Cell Atlas (adult and mouse – 43 tissues) (Han et al., 2018)

The Mouse Cell Atlas (MCA) involved the collection and sequencing of 50+ tissues in an attempt to distinguish genetic signatures of tissues across a mouse's life, of which 43 passed the author's quality control measures. These tissues included: 21 adult tissues (from 6-10 weeks old C57BL/6J mice, including the mammary gland sequenced at four separate life stages), 11 fetal tissues (E14.5), six neonatal tissues and 4 stem cell derived populations. An Expression matrix ('MCA_Figure2-batch-removed.txt.tar.gz') of 60,000 cells of high quality (~1500 cells from the 43 tissues) was downloaded from Figshare (https://figshare.com/articles/MCA_DGE_Data/5435866) alongside the cell metadata ('MCA_Figure2_Cell.info.xlsx'). Raw data is available from GEO repository, GSE108097. Data were filtered for the 18 unique adult tissues (including the mammary gland from virgin females only). Data were then scaled to 100,000 reads and log normalised. Tissue cell identities were acquired from the 'Tissue' annotations. Data were run through the core workflow once for all adult tissues.

Tabula Muris (adult mouse – 20 tissues) (Schaum et al., 2018)

The *Tabula Muris* Consortium released the *Tabula Muris* (TM) soon after the MCA. The *Tabula muris* comprises 20 adult mouse organs which partially overlap with the MCA data; however, TM sequenced these tissues to a much greater depth and using two sequencing methods. Droplet sequencing allowed the surveying of a large number of cells at relatively low coverage and was completed for 13 organs. Fluorescent activated cell sorting (FACs) allowed for the characterisation of fewer cells from the organ but at much higher sensitivity and depth and was completed for all 20 organs. This was completed using 3 female mice and 4 male mice, all 10-15 weeks old and on a C57BL/6JN background. The FACs dataset was

selected for analysis as it contained the largest number of tissues and the only one to include the brain (broken down as myeloid brain cells (neurons, glia, endothelial etc) and non-myeloid brain cells (microglia and macrophages)). Clustering for cells in this dataset allowed cell identification at the Tissue and cell identity level. All data were downloaded through Figshare as Robjects (https://figshare.com/articles/Robject_files_for_tissues_processed_by_Seurat/5821263/1). Raw data and processed data were accessible from these files and cell metadata were acquired as annotation csv's from the respective sequencing technology Figshare page (FACS - https://figshare.com/articles/Single-cell_RNA- seq_data_from_Smart-seq2_sequencing_of_FACS_sorted_cells_v2_/5829687). Raw data is also available from GEO repository, GSE109774. Raw data were used to identify genes expressed in 20 or more cells and processed data were then used unaltered for the analysis. Tissue cell identities were acquired from the 'tissue' annotations and run through the core workflow once for all adult tissues.

2) Whole Brain level

Mouse Brain Atlas (Zeisel et al., 2018)

{Zeisel, 2018 #48@@author-year} created the Mouse Brain Atlas (MBA) in 2018 – a comprehensive sequencing of the adolescent mouse nervous system. This involved sequencing cells from the entire central and the peripheral nervous system, including the enteric nervous system. Cells were clustered into class (e.g., neuron, oligodendrocyte), region of origin (e.g., hypothalamus, pons) and specific cell type by neurotransmitter (e.g., serotonergic, dopaminergic) and most likely nuclei of origin. Data were downloaded from mousebrain.org (http://mousebrain.org/downloads.html) and the Level 5 data were downloaded as a one cell per column loom file which also included the cell metadata as column attributes ("l5.all.loom"). Data were then log normalised and scaled to 5,000 reads. Cell lineage labels were acquired from 'Class' annotations, cell tissue labels were acquired from 'tissue' annotations and cell identity labels were acquired from 'ClusterName' annotations. The core workflow as run three times, once for cell lineage with all cells, before restricting the analysis to neurons only and running it once for nervous system regions and once for the specific neuron identities across the nervous system. Marker genes to identify neuron subpopulations were acquired from combined enrichment analysis with trinaization scores carried out by the original authors located from their online resource (http://mousebrain.org/). Two additional marker genes were included – *Esr1* and *Prlr* – for

Deinh5 neurons based on our own analyses. Raw data is also available from SRA repository, SRP135960

Single-Cell Transcriptome profiling of the aging mouse brain (Ximerakis et al., 2019)

To complement the analysis of the mouse brain atlas, we also analysed Ximerakis et al. (2019)'s single cell profile of the young and old mouse brains. Although the original aim of the study was to identify aging-related genes/pathways, the data could also be used to investigate differential expression between the different cell types within the brain, grouped by cell lineage. 16 C57BL/6 mice (8 at 2/3 months and 8 at 21/22 months) had whole brains dissected and hindbrain removed before tissue was dissociated into single cells. For this analysis we chose to only include the 2–3-month-old mouse data which was the most comparable to the age of mice from other studies used by us. Expression data and meta data were acquired from Gene Expression Omnibus through accession no. GSE129788, scaled to 10,000 reads and log normalised. Data were run through the workflow once comparing all major CNS (minus hindbrain) cell types.

3) Brain Nuclei level

3a) Whole Hypothalamus

Chen et al. (2017)

Chen et al. (2017) sequenced adult (8-10 weeks) whole hypothalamus of female B6D2F1 mice (C57B6 female × DBA2 male). Processed/normalised data were downloaded as an R object from Gene Expression Omnibus through accession no. GSE87544 alongside count data and a metadata file with the cell identities. The original 14,437 cells were filtered to 3,319 high quality cells by the authors (containing > 2000 unique genes expressed) and this filtered number was maintained for this analysis. Cell identities included 11 non-neuronal and 34 neuronal identities identified as Snap25/Syt1-high (and split into GABA and Glut populations alongside a Hista cluster). Cells without a cluster (labelled as 'zothers') were discarded from the analysis leaving 2275 neurons. Marker genes supplied for these neuronal identities were identified by the original authors. Cell identities for all cells were taken from 'SVM_clusterID' annotations from the metadata file. The core workflow as then run twice, once for all cells and once for just the neurons.

Romanov et al. (2017)

To complement the previous study, the data from Romanov et al. (2017) were also analysed. The authors sequenced cells from "a central column" of the mouse hypothalamus which spanned from the posterior POA to the ARC (rostro-caudal), PVN to the VLH (dorsal-lateral). Mice were on C57BL6/N background, of both sexes and aged 14-28 days. The expression matrix and meta data was acquired from a single file available from Gene Expression Omnibus through accession no. GSE74672. The expression matrix was prefiltered by the authors to 2,882 high quality cells (>1,500 molecules excluding rRNA) and was scaled to 10,000 reads and log2 normalised. Cell identities included 6 non-neuronal identities and 62 neuronal subtypes (clustered from 898 neurons) identified by expression of key neuropeptides (e.g., *Avp*, *Oxt*, DA). The 'level1 class' annotations were used for non-neuronal populations and the 'level2 class (neurons only)" annotations were used for the 62 neuronal identities. The core workflow was then run twice, once for all cells and once for just the neurons.

3b) Specific Hypothalamic Nuclei

Several datasets focusing on individual regions of the hypothalamus were in existence prior to this investigation and each was investigated individually for imprinted enrichment.

Preoptic Area (POA) (Moffitt et al., 2018)

The Preoptic area is of particular interest to parenting behaviour and feeding behaviour containing neural populations that contribute to a number of innate hypothalamic behaviour outputs. Moffitt et al. (2018) conducted a molecular, spatial, and functional characterisation of the murine POA and as part of this investigation they profiled 31,299 cells from male and female adult mouse POA. Alongside 21 major cell classes, 18,553 neurons were clustered into 66 distinct neuronal populations (of which 56 originated from the POA). Barcodes, Features and Matrix files were acquired from Gene Expression Omnibus through accession no. GSE113576 and used to create a Seurat object for the analysis. Data were scaled to 10,000 UMI and log10-normalised. Cluster identities were acquired from Supplementary Table 1. accompanying the original publication (https://science.sciencemag.org/content/362/6416/eaau5324). Neuronal cluster identities were taken from 'Neuronal cluster (determined from clustering of inhibitory or excitatory neurons)' annotations. Data were run through the workflow thrice, once with all neurons and twice with only POA neurons (removing the 9 suspected extra-POA groups) with upregulated

genes capped at the standard 1FC cap which generated a large number of over-represented sub populations and so again at the more conservative 2FC cap.

Arcuate Nucleus (ARC) (Campbell et al., 2017)

Campbell et al. (2017) sequenced the arcuate-median eminence complex of adult C57BL/6J mice under different feeding conditions. They identified 65 distinct cellular identities of which 34 were unique neuronal identities defined by key neuropeptides and novel markers (and 24 originated from the arcuate nucleus). Raw and processed/filtered expression matrices alongside the cell metadata (including cluster annotations) were acquired from Gene Expression Omnibus through accession no. GSE93374. Raw data were scaled to 10,000 transcripts per cell (following Timshel et al. (2020) and log normalised. Cell labels were acquired from the "Subcluster" annotations from the metadata. Data were run through the workflow twice, once with all neurons and a second time with only ARC neurons (removing the 10 suspected non-ARC subpopulations identified by the original authors).

Suprachiasmatic Nucleus (SCN) (Wen et al., 2020)

In a spatio-temporal manner, Wen et al. (2020) used single-cell RNA sequencing to investigate the basic and circadian expression of various cell types in the SCN by micro dissecting the SCN and surrounding areas of 55 virgin 7-8 week old male C57BL/6J mice and sequencing them at 12 circadian time points (5 per timepoint). For the main experiment, conducted via Drop-seq, 8 major cell types (incl. neurons) were identified and within the neuronal cluster, 5 SCN neuronal subtypes were identified (ignoring neuronal subtypes likely from other neighbouring brain regions). Neuron identities were then reconfirmed with 10x Genomics, processing an additional 8,679 cells including 1,251 SCN neurons. These neurons were classified into the same 5 neuronal clusters as drop-seq and found a high convergence between the groups. Since the 10x data was more comprehensive for neurons, this data were chosen to look for neural specific enrichment only. The 10X Seurat files was acquired from Gene Expression Omnibus from accession no. GSE132608 and a Seurat object was created. Cell metadata were acquired through personal communication with the authors. SCN neuron identities were acquired from the 'cluster' annotation. The data were scaled by 5000 reads per cell, log-normalised and run through the workflow once.

3c) Monoaminergic Nuclei

In addition to the specific hypothalamic nuclei, I decided to investigate imprinted gene expression in other brain areas of interest, namely the dopamine neurons of the midbrain and the various neurons of the Dorsal Raphe Nucleus, both of which had datasets available and were regions of over-representation in the analysis of the Mouse Brain Atlas (Romanov et al., 2017).

Dorsal Raphe Nucleus (Huang et al., 2019)

The Dorsal Raphe Nucleus is a critical nucleus in the central nervous system (CNS) responsible for a significant proportion of neuromodulators, and most significantly, approximately a third of all serotonergic neurons. Huang et al. (2019) sequenced the dorsal raphe nucleus and surrounding areas of 8-10-week-old C57BL/6J mice. They discovered 17 major clusters (of which 5 were major neuron classes e.g., serotonergic, dopaminergic) and the neuron classes were found to have an additional 17 subclusters. Raw data is available at GEO depository, GSE134163. Data were downloaded as processed and annotated Seurat objects (already log normalised, scaled to 10,000 UMI) for each of the major cell types from Harvard Dataverse (https://doi.org/10.7910/DVN/ QB5CC8). The workflow was run three times, once using major cell cluster identities acquired from the 'Curated_cellTypeLabels' annotations and once using just neurons with the major cell clusters and a final time using just the neurons with the sub cluster identities acquired from the 'Curated_subtypeLabels.'

Brain-wide Dopamine Neurons (Hook et al., 2018)

The other key neuron type that suggested a presence of imprinted genes was the dopamine neuron population. Hook et al. (2018) sequenced 473 dopamine neurons in the C57BL/6J mouse brain at embryonic (E15.5) and early postnatal (P7). Midbrain and Forebrain dopamine neurons were microdissected from E15.5 mouse brains and the olfactory bulb, midbrain and forebrain (arcuate nucleus) dopamine neurons from P7 mice, all identified using FACS, were also dissected. A total of 13 cell populations were found (4 at E15.5 and 9 at P7) which included clustering of dopamine neurons from the arcuate nucleus, VTA, PAG and SN individually. The raw and normalised expression matrix, pre-filtered for the 396 dopamine neurons meeting the author's quality criteria was acquired from github (https://github.com/ pwh124/sc-da-parkinsons) along with accompanying cell metadata (cell cluster identities). Raw data is available at GEO repository, GSE108020. Data were run through the workflow once.

Developing Midbrain (La Manno et al., 2016)

As a final investigation into the presence of imprinted genes in the midbrain, the data from La Manno et al. (2016) were analysed. The authors sequenced single cells from the CD-1 mouse ventral midbrain at time points between E11.5 – E18.5 covering dopaminergic progenitor specification and differentiation. Over this time period, 26 distinct cellular identities were identified including a variety of unique neuronal identities (e.g., serotonergic, dopaminergic, GABAergic). Data were downloaded from Gene Expression Omnibus using accession no. GSE76381. The expression matrix was scaled to 10,000 reads and log normalised. Metadata was included as a row within the gene expression matrix and cell identities were prescribed using the 'cell-type' annotations. Data were run through the workflow once using all cell identities.

3d) Pituitary Gland

Anterior Pituitary Gland (Ho et al., 2020)

Ho et al. (2020) sequenced the anterior pituitary gland using both 10X genomics and Drop-seq platforms. The 10Xgenomics experiment involved 2,780 cells from 8-week-old CD1 males and females. The Drop-seq experiment sequenced cells from WT mice at 8 weeks but also 13week old and *mt*/*hGRF* transgenic mic, for consistency only the 8-week WT mice were analysed resulting in 4,663 cells. Both 10x and Drop-seq matrices were acquired from a public google drive as RDS datasets - https://drive.google.com/drive/folders/1cbwkQl1Hh70xsToNhjOdom_ru7XGA3l-. Cell metadata for both experiments were acquired from personal correspondence with the authors and cell identities were acquired from the 'cluster' and 'genotype' annotation. Raw data are available at GEO depository, GSE146619. Analysis pathways to filter, normalise and scale the data were acquired from https://github.com/wulabupenn/mPit for both datasets and data were processed until the point of pre-clustering following the author's workflow. Additionally, genes were filtered to those expressed in 20 or more pituitary cells. Drop-seq data was additionally filtered to select only cells from 8-week-old-WTs. The workflow was run once for each dataset comparing all cell types across both sexes.

Pituitary Gland (Cheung et al., 2018)

As an additional independent analysis of the pituitary gland was available through Cheung et al. (2018)'s sequencing of 6, 7-week-old male C57BL/6 pituitary glands. This involved

13,663 cells with many overlapping cell types with Ho et al., 2020. Files necessary to create a Seurat object were acquired from Gene Expression Omnibus through accession no. GSE120410 and manual cell clusters were acquired from personal correspondence with the author's. Data were scaled to 10,000 reads and log transformed. The workflow was run once, for all major cell types. Cells were run through the workflow once including all cell types.

4) Other Mouse Organ level

Mouse cell atlas (adult and developing mouse – 43 tissues) (Han et al., 2018)

For the analysis on all 43 tissues, the workflow was run twice, once for all 43 tissue identities ('Tissue' column) and once again for the 63 major cell identities ('Cluster ID column') across the 43 tissues. For the adult tissue only, there were a total of 20,744 cells assigned to 292 unique specific cell identities from the adult mouse, annotated from the 'annotations' column. The core workflow was run once for all the adults cells and then again for the tissue types on interest from the previous whole tissue analysis, namely the pancreas, bladder and mammary gland. Genes were filtered to 20 cells for each analysis.

Tabula Muris cell populations (adult mouse – 20 tissues) (Schaum et al., 2018)

Cell assignments were acquired from the 'free_annotation' and 'subtissue' column of the metadata, to provide cells with either a cell identity or at the least a sub tissue identity. The core workflow was run once for all the cells and then again for cells from only the pancreas to allow direct comparisons to other datasets. Genes were filtered to 20 cells for each analysis.

Pancreas (Baron et al., 2016)

The Pancreas was followed up independently after it was found as an enriched tissue in adults through both datasets. Baron et al. (2016) sequenced 2000 pancreatic cells from 5 mice and identified 13 cell clusters including the main islets (alpha, beta, etc.). Raw data and cell metadata were acquired together for the mouse data from two files from Gene Expression Omnibus through accession no. GSE84133. The expression matrix included cell cluster labels through the 'assigned_cluster' annotations. Data were normalised to 10,000 UMI and log transformed and run through the core workflow once for the major cell types.

Muscle cells (De Micheli et al., 2020)

Muscle tissue was followed up independently after found as an enriched tissue in adult in the TM dataset. De Micheli et al., (2020) sequenced cells from injured tibalis anterior muscles at 0-, 2-, 5- or 7-days post injury (notexin injection). The non FACs sorted sample was used for this analysis and the raw, processed and metadata were acquired from Gene Expression Omnibus through accession no. GSE143437. Metadata included cell identities as 'cell-annotation' annotation. Raw data were used to apply the 20-cell filter to the genes and the processed data were filtered and run through the workflow once for all major cell types.

Mammary Gland Epithelial Cells (Bach et al., 2017)

The mammary gland, although not identified as a tissue of interest from the cross-tissue analysis, has been raised as a potential site of imprinted gene enrichment, particularly in the context of marsupials and very recently restated as a site of imprinting in mammals generally. Bach et al. (2017) use single-cell RNA sequencing to characterize the mammary epithelial cells across four developmental stages of the mammary gland (nulliparous, mid gestation, lactation, and post involution). Processed data and metadata were acquired through personal correspondence with the authors (although raw data is available from Gene Expression Omnibus through accession no. GSE106273). The matrix was scaled using the supplied size factors and log2 normalised. Cell annotations were taken from the 'SuperCluster' annotations involving 8 unique cell clusters and the 'Condition' annotations for developmental stage and cells were run through the workflow twice, once for cell types and once for developmental stage.

Appendix A2 – Level 4 Analyses (Chapter 3)

Embryo vs. Neonatal vs. Adult Analysis

Imprinted gene are over-represented in foetal and neonatal tissues over adult at whole level analysis

The MCA allowed one further multi-organ comparison since the adult tissues were sequenced alongside a selection of embryonic (E14.5) and neonatal tissues, I carried analysed enrichment across tissue types and again across major cell subpopulations.

Appendix Table A2.1 shows the analysis of global cell populations encompassing all tissues sequenced in the MCA (incl. fetal, neonatal and extra-embryonic tissues). There is a distinct bias for overrepresentation in fetal (5/9) and neonatal (5/6) tissues over adult tissues (1/19); the pancreas being the only adult tissue to be considered over-represented for imprinted genes in this analysis. Both the placenta and embryonic mesenchyme were over-represented for imprinted genes.

Appendix Table A2.1 Imprinted gene enrichment in tissues derived from adult, foetal, embryonic and stem-cell derived cells in the MCA (Han et al., 2018). *Up Reg* – number of upregulated genes with $q \leq 0.05$ and Log2FC ≥ 1 (total number of genes in the dataset in brackets); *IG* – number of imprinted genes upregulated with $q \leq 0.05$ and Log2FC ≥ 1 (total number of IGs in the dataset in brackets); *ORA p – p* value from over representation analysis on groups with minimum 5% of total IGs; *ORA q* – Bonferroni corrected *p* value from ORA; *Mean FC IG* – mean fold change for upregulated imprinted genes; *Mean FC Rest* – mean fold change for all other upregulated genes; *GSEA p – p* value from Gene Set Enrichment Analysis for identity groups with 15+ IGs and Mean FC IG > Mean FC Rest; *GSEA q* – Bonferroni corrected *p* values from GSEA.

Tissue Identity (43)	Up Reg (22,810)	IG (107)	ORA p	ORA q	Mean FC IG	Mean FC Rest	GSEA p	GSEA q
Neonatal Muscle	1078	28	6.79E-14	**2.17E-12**	4.62	4.21	0.0084	**0.0168**
Neonatal Rib	1223	26	5.60E-11	**1.79E-09**	4.33	4.71	-	-
Fetal Stomach	470	15	5.46E-09	**1.75E-07**	5.77	5.33	0.2246	0.4492
Neonatal Calvaria	1415	24	2.97E-08	**9.50E-07**	5.07	5.56	-	-
Placenta	1094	19	7.28E-07	**2.33E-05**	9.36	92.73	-	-
Embryonic Mesenchyme	1013	18	1.07E-06	**3.43E-05**	3.66	4.72	-	-
Neonatal Heart	1035	18	1.46E-06	**4.66E-05**	3.23	9.30	-	-
Pancreas	3466	36	1.58E-06	**5.05E-05**	3.98	6.47	-	-
Fetal Intestine	622	13	7.04E-06	**0.0002**	3.55	4.47	-	-
Neonatal Skin	1199	17	4.18E-05	**0.0013**	3.45	5.04	-	-
Fetal Kidney	792	13	8.66E-05	**0.0028**	2.98	4.85	-	-

Secondly, these cells were then recategorized into 67 tissue-spanning cell types (Appendix Table A2.2). Over-representation was seen principally in the mesenchymal stem cell derived cells (osteoblast, chondrocyte, myocyte, muscle cell, cartilage) spread across neonatal and adult tissue. Endocrine cells (mainly derived from the pancreas) were also over-represented. However, of specific interest were the trophoblast progenitor cells isolated from the placenta, which was the only cell type across all sequenced tissues to show a significant GSEA (Appendix Figure A2.1).

Appendix Table A2.2 Imprinted gene enrichment in global cell types across adult, foetal, embryonic and stem-cell derived tissues in the MCA (Han et al., 2018). Column descriptions can be found in the legend of Appendix Table A2.1.

Identity	Up Reg (22,810)	IG (107)	ORA p	ORA q	Mean FC IG	Mean FC Rest	GSEA p
Myocyte	807	26	4.33E-15	**2.08E-13**	3.99	4.31	-
Cartilage cell	475	16	7.42E-10	**3.56E-08**	7.07	8.17	-
Stromal cell	2200	31	1.39E-08	**6.67E-07**	4.08	4.62	-
Chondrocyte	1414	24	2.93E-08	**1.41E-06**	5.38	7.93	-
Myoblast	854	17	4.69E-07	**2.25E-05**	3.34	4.76	-
Pancreatic acinar cell	496	12	3.74E-06	**0.0002**	3.79	10.66	-
Trophoblast progenitor cell	1512	21	6.22E-06	**0.0003**	14.73	11.00	**0.003**
Cardiac muscle cell	854	14	4.60E-05	**0.0022**	3.29	14.51	-
Cycling cell	458	10	5.98E-05	**0.0029**	3.17	4.82	-
Endothelial cell	1549	19	9.93E-05	**0.0048**	3.18	5.94	-
Muscle cell	999	14	0.0002	**0.0115**	2.87	4.14	-

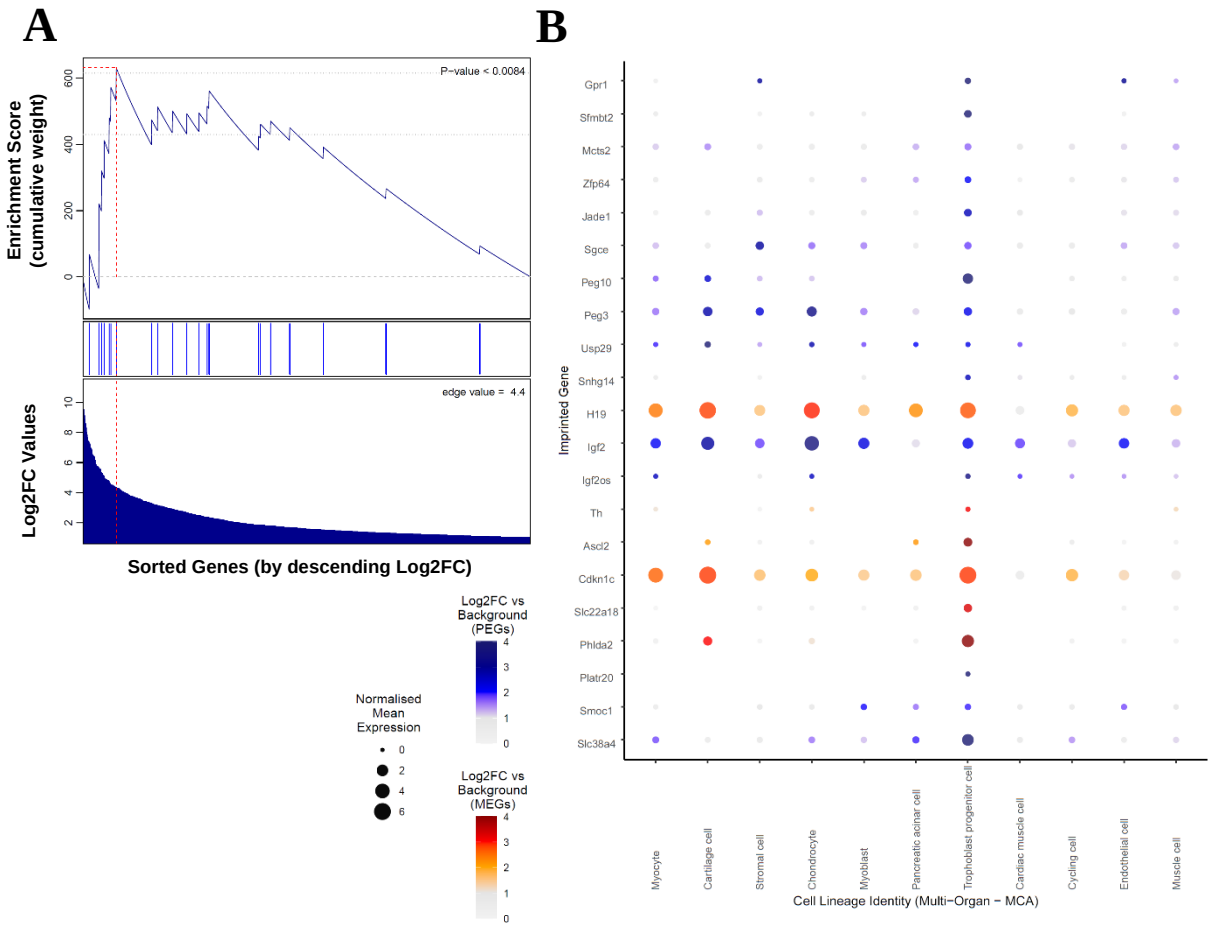
Adult tissues cell population analyses

Imprinted genes are over-represented in stromal and mesenchymal cell populations of the MCA & *TM*

I began by analysing the adult MCA and *Tabula Muris* at the cell-population specific level – to find the cell types driving the enrichment in tissues (other than the brain) e.g., bladder and muscle tissues). Pancreatic enrichment will be the focus of the following section.

Cells from the 18 adult tissues of the MCA were distinguished into 292 tissue specific cell types. Cell-types with an over-representation of imprinted genes from the analysis of the adult MCA cells are shown in Appendix Table A2.3. Stromal cells from various organs (incl. pancreas, bladder and mammary gland) were found to be the major over-represented cell type. MCA stromal cells were identified by Han et al. (2018) as “connective tissue cells that

support the function of parenchymal cells" and were identified based on expression of collagens, laminins, elastin and fibronectin. This includes cells such as fibroblasts, extracellular matrix and mesenchymal stem/stromal cells (Meirelles, Chagastelles & Nardi, 2006, Schafer, 2012; Valkenburg, Groot & Pienta, 2018).



Appendix Figure A2.1. Gene Set Enrichment Analysis of Imprinted Genes for genes upregulated in Progenitor Trophoblast cells. (*A*) GSEA graph - In the GSEA analysis, genes are sorted by strength by which they mark this neuronal cluster (sorted by fold change values) indicated by the bar (middle). Fold change values are displayed along the bottom of the graph. The genes are arrayed left (strongest marker) to right and blue lines mark where imprinted genes fall on this array. The vertical axis indicates an accumulating weight, progressing from left to right and increasing or decreasing depending on whether the next gene is an imprinted gene or not. The *p*-value represents the probability of observing the maximum value of the score (red dashed line) if the imprinted genes are distributed randomly along the horizontal axis. (*B*) Dot plot of imprinted genes upregulated in the Trophoblast Progenitor Cells plotted across all imprinted gene over-represented cell types in the MCA. Size of points represented absolute mean expression; colour represented the size of the Log2FC value for that cell type (e.g., trophoblast progenitor cells) vs. all other cells. MEGs and PEGs possess distinct colour scales.

Appendix Table A2.3. Imprinted gene over-representation in MCA adult cell populations (Han et al., 2018). Column descriptions can be found in the legend of Appendix Table A2.1.

Identity	Up Reg (20,534)	IG (95)	ORA *p*	ORA *q*	Mean FC IG	Mean FC Rest	GSEA *p*	GSEA *q*
Stromal cell_Dpt high (Bladder)	1697	24	5.44E-07	**4.90E-05**	5.51	5.03	0.15	0.60
Endocrine cell (Pancreas)	666	14	2.25E-06	**0.0002**	12.06	15.68	-	-
Stromal cell_Smoc2 high (Pancreas)	1593	22	2.68E-06	**0.0002**	9.45	5.36	0.07	0.27
Stromal cell_Mfap4 high (Pancreas)	1164	18	5.57E-06	**0.0005**	6.08	7.56	-	-
Î²-cell (Pancreas)	1413	19	2.14E-05	**0.0019**	11.14	12.18	-	-
Stromal cell_Dcn high (Lung)	762	13	4.82E-05	**0.0043**	4.88	6.92	-	-
Stromal cell_Col3a1 high (Mammary-Gland)	802	13	8.10E-05	**0.0073**	9.09	5.42	-	-
Acinar cell(Pancreas)	260	7	0.0002	**0.0185**	26.39	47.94	-	-
Stromal cell_Pi16 high (Mammary-Gland)	655	11	0.0002	**0.0198**	6.83	8.03	-	-
Dividing cell (Pancreas)	1054	14	0.0003	**0.0305**	17.2	15.21	-	-
Mesenchymal stromal cell (Bladder)	1056	14	0.0003	**0.0311**	4.89	5.69	-	-
Myelinating oligodendrocyte (Brain)	2566	24	0.0005	**0.0452**	6.94	12.54	-	-

Cells from the 20 tissues of the TM were distilled into 89 tissue-spanning cell types. Significant cells from the analysis of the *Tabula Muris* adult cells are shown in Appendix Table A2.4. Over-representation was predominantly seen in pancreatic and muscle-based cell types but cell types from the brain, bladder, mammary gland and adipose were also found over-represented from this large-scale comparison. Importantly, stromal cells, originating from the mammary gland and lung, and various mesenchymal stem cells were sites of enrichment, overlapping with the results from the MCA

Appendix Table A2.4. Imprinted gene over-representation in *Tabula Muris* adult cell populations (Schaum et al., 2018). Column descriptions can be found in the legend of Appendix Table A2.1.

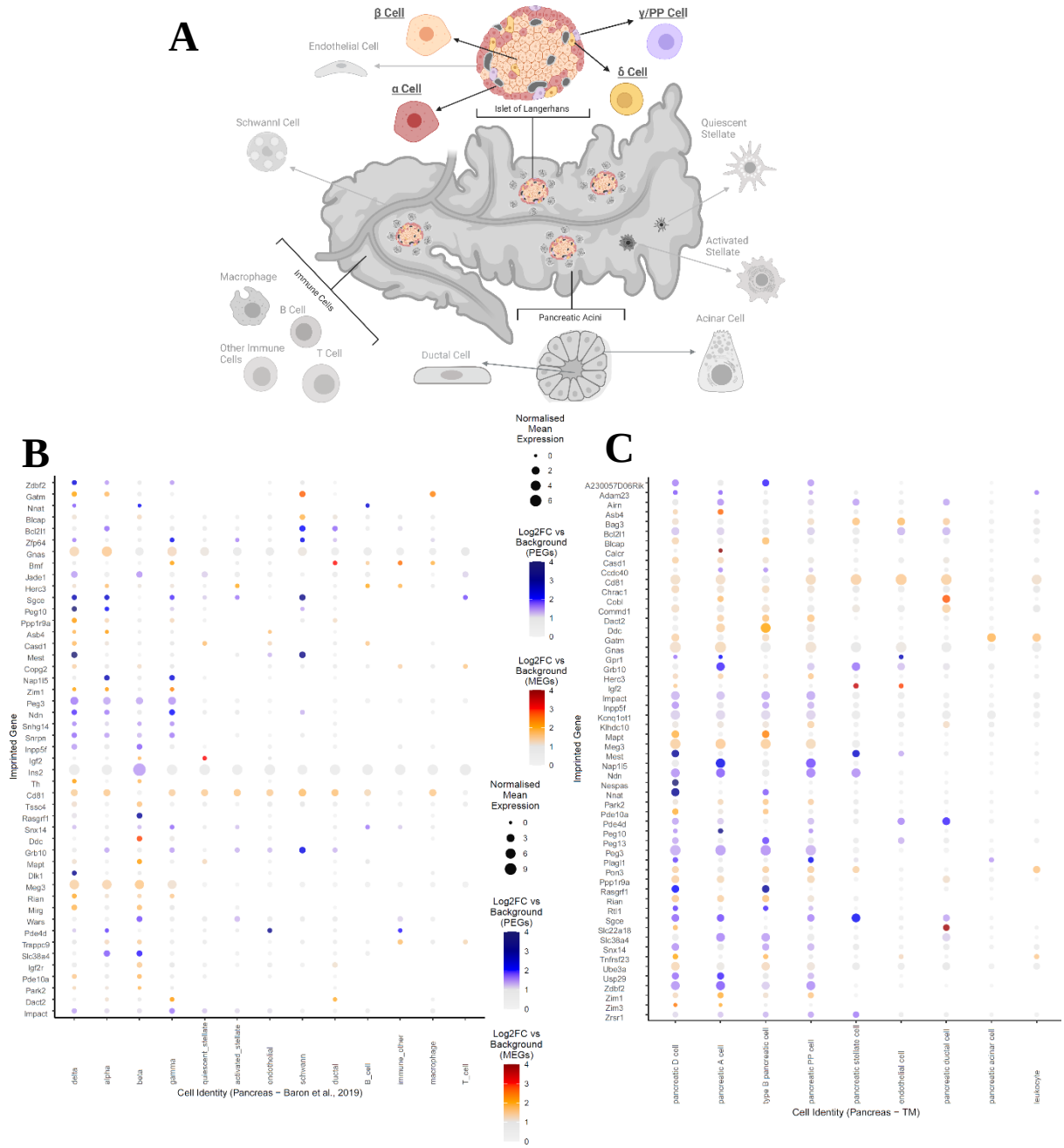
Identity	Up Reg (20,839)	IG (107)	ORA *p*	ORA *q*	Mean FC IG	Mean FC Rest	GSEA *p*	GSEA *q*
Skeletal muscle satellite stem cell (Diaphragm)	462	24	1.01E-17	**5.65E-16**	9.22	6.94	0.32	1
Skeletal muscle satellite cell (Limb Muscle)	833	23	3.05E-11	**1.71E-09**	16.53	6.91	0.06	0.77
Mesenchymal stem cell (Diaphragm & Limb Muscle)	1289	24	2.78E-08	**1.56E-06**	4.69	4.25	0.16	1

Pancreatic stellate cell	1953	28	3.89E-07	**2.18E-05**	4.75	6.03	-	-
Mesenchymal cell (Trachae)	2500	32	5.80E-07	**3.25E-05**	4.71	4.91	-	-
Stromal cell (Mammary Gland & Lung)	1528	24	6.48E-07	**3.63E-05**	3.43	4.20	-	-
Mesenchymal stem cell of adipose	2396	30	2.35E-06	**0.0001**	5.04	4.77	0.41	1
Pancreatic D cell	4713	45	5.54E-06	**0.0003**	11.57	7.89	0.09	1
Mesenchymal cell (Bladder)	2820	30	5.86E-05	**0.0033**	4.62	5.79	-	-
Neuron (Brain Non-Myeloid)	5492	47	6.22E-05	**0.0035**	15.58	16.95	-	-
Endocardial cell (Heart)	912	15	6.46E-05	**0.0036**	11.37	5.68	0.0146	0.2044
Pancreatic B cell	5024	44	7.28E-05	**0.0041**	14.61	8.89	0.0151	0.2114
Oligodendrocyte precursor cell (Brain Non-Myeloid)	3969	36	0.0002	**0.0134**	6.69	12.41	-	-
Pancreatic PP cell	4385	38	0.0004	**0.0215**	8.64	6.64	0.1114	1
Oligodendrocyte (Brain Non-Myeloid)	1860	21	0.0004	**0.0252**	7.37	11.70	-	-
Pancreatic A cell	4179	35	0.0014	0.0776	11.43	7.17	0.0069	0.0966

Imprinted genes are over-represented in pancreatic endocrine cell types

Looking at the pancreas specifically, the MCA identified a variety of cells from the pancreas – Acinar, Ductal, Endothelial, Immune, Smooth muscle, and Stromal but did not distinguish all the islet cells, only the Beta cells and the rest as endocrine cells. From the whole cell analysis in Appendix Table A2.3, I saw that the endocrine pancreas and the beta cells were over-represented alongside the three stromal cell subpopulations. The stromal cells of the pancreas were the only over-represented cell type when analysing the pancreas alone (Appendix Table A3.20A) but imprinted genes had a 3.9x greater mean FC in the beta cells than the rest suggesting the genes upregulated here were dramatically so.

The *Tabula Muris* identified a similar selection of pancreatic cells and within the tissue-wide analysis (Appendix Table A2.4) I saw over-representation of imprinted genes in three of four types of endocrine islet cells in the pancreas (Beta, Delta and Polypeptide, but not Alpha) alongside the top hit, the stellate cells - a multifunctional cell in the endocrine and exocrine pancreas that express collagen and fibronectin and promotes fibrosis when activated (Zhou et al., 2019). When analysing the pancreas cells alone (Appendix A3.20B) over-representation was seen in the δ cells ($q = 0.004$) and the stellate cells ($q = 0.033$) only.



Appendix Figure A2.2. (*A*) Pancreatic cell types with imprinted gene over-representation in Baron et al. (2016) - Over-represented cell types are bold, underlined and not in greyscale. (*B&C*) Dot Plots showing imprinted gene expression across all cell types in Baron et al. (2016) (*B*) and *Tabula Muris* (*C*) pancreatic datasets – Imprinted genes with upregulated expression in at least one endocrine cell were plotted for each dataset independently across all cell types. See legend of Appendix Figure A2.1 for how to interpret the plot.

To provide another independent pancreatic cell analysis, I analysed the pancreatic data from Baron et al. (2016), detailed in Appendix Table A2.5. Again, converging with the TM, the endocrine cell populations of the pancreas were over-represented (although this time all 4 types), with the δ and α cells as the top hits, and analysing these pancreatic cells alone saw no

over-representation in the stellate cells. See Appendix Figure A2.2A for a summary graphic of the pancreatic analysis. Appendix Figure A2.2B & C presents dot plots of the imprinted

Appendix Table A2.5. Imprinted gene over-representation in Pancreatic adult cell populations (Baron et al., 2016). Column descriptions can be found in the legend of Appendix Table A2.1.

Identity	Up Reg (11,418)	IG (77)	ORA *p*	ORA *q*	Mean FC IG	Mean FC Rest	GSEA *p*	GSEA *q*
Delta cell	1271	23	6.17E-06	**5.56E-05**	4.27	4.07	0.3616	0.7232
Alpha cell	1091	18	0.0003	**0.0024**	2.46	2.93	-	-
Beta cell	2073	25	0.0017	**0.0157**	2.33	2.63	-	-
Gamma/PP cell	565	10	0.0045	**0.0404**	2.64	3.36	-	-
Quiescent Stellate cell	1119	11	0.1298	1	4.31	6.87	-	-
Activated Stellate cell	840	6	0.5037	1	6.85	7.58	-	-
Endothelial cell	2085	14	0.5538	1	5.70	22.82	-	-
Schwann cell	2001	11	0.8146	1	5.60	6.02	-	-
Ductal cell	3220	18	0.8587	1	14.23	7.79	0.1444	0.2888
Macrophage	1039	3	-	-	2.83	41.51	-	-
B cell	479	2	-	-	6.14	10.90	-	-
Immune cell (other)	664	0	-	-	0	12.49	-	-
T cell	369	0	-	-	0	12.14	-	-

genes upregulated in the cell populations within the TM and Baron et al., 2016 pancreas only datasets.

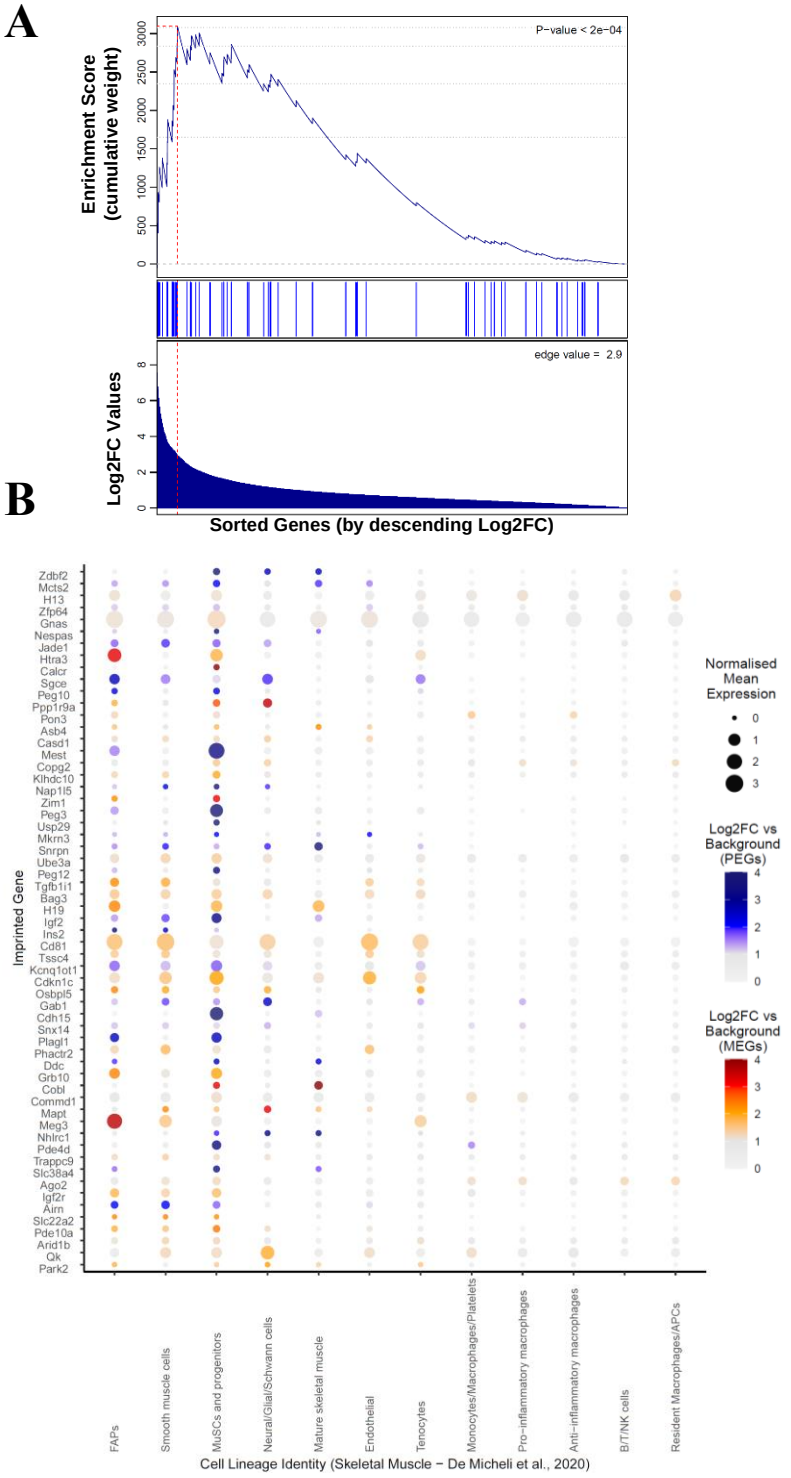
Imprinted genes are over-represented in mesenchymal bladder cells

The bladder was represented by an over-represented stromal cell population in the MCA. The authors of the MCA distinguished bladder stromal cells as expressing *Bmp4* and *Wnt2*, both of which are markers of mesenchymal stromal cells (Mysorekar et al., 2002; Pokrywczynska et al., 2019) with a role in epithelial maintenance and renewal (Mysorekar et al., 2010). Convergently, the mesenchymal bladder cells were the over-represented population in the TM analysis.

Imprinted genes are over-represented in Muscle satellite/stem cells

Despite not being an over-represented tissue in the MCA multi-organ analysis, two muscle-based cell subpopulations were over-represented in the cell subpopulation analysis: a stromal cell population and a muscle progenitor population. In the TM analysis, over-representation was seen in a variety of cells originating from the muscle tissues as the top hits – skeletal muscle satellite cells and mesenchymal stem cells, which mirrors the enrichment of muscle progenitor cells and the mesenchymal cells mirroring the stromal cells in the MCA.

For further analysis, the data from De Micheli et al. (2020) of skeletal muscle during repair were analysed in Appendix Table A2.6. The only significant GSEA was seen in the stem cells and progenitor cell types (also known as the satellite cells) (Appendix Figure A2.3). Additional over-representation was seen in the fibro/adipogenic progenitor cells, also a type of mesenchymal stromal cell which support satellite cell differentiation (Madaro. Mozzetta, Biferali & Proietti, 2019). This shows a very consistent picture with the MCA and TM



Appendix Figure A2.3 Gene Set Enrichment Analysis of Imprinted Genes for genes upregulated in MuSCs and Progenitors from Di Micheli et al. (2020). (*A*) GSEA for imprinted genes upregulated in the 'MuSCs and progenitors' cell type in the Skeletal muscle dataset. See legend of Appendix Figure A2.1 for a description of how to interpret the plot. (*B*) Dot plot of imprinted genes upregulated in the ''MuSCs and progenitors' cell type plotted across all identified cell types in the Di Micheli et al. (2020) skeletal muscle dataset. See legend of Appendix Figure A2.1 for a description of how to interpret the plot.

analyses that the mesenchymal stem cell population in the muscles are the key enriched population.

Imprinted genes are over-represented in Mammary Gland stem cell populations

The mammary gland was not an enriched tissue in either the MCA or TM multi-organ analysis, but the MCA cell subtype analysis saw both populations of stromal cell over-represented and the stromal cells derived from the mammary gland were also enriched in the TM. Both datasets distinguished these stromal cells by expression of matrix metallopeptidases Mammary Stroma consists of a variety of cells including adipocytes, fibroblasts and ECM. The structural ECM is a large part of the stroma and is believed to be an important mammary stem cell niche (Khokha & Werb, 2011; Wiseman & Werb, 2002). However, Xu et al. (2020) reported high expression of imprinted genes in a stem-cell type in the epithelial component and over-representation analysis of Bach et al.'s (2017) scRNA-seq of mammary epithelial cells (Appendix Table A2.7) shows imprinted genes are enriched in the basal compartment of the mammary epithelium, but the top hit was the stem (Procr+) cells as Xu et al. (2020) reported.

Appendix Table A2.6. Imprinted gene over-representation in Muscle cell populations (De Micheli et al., 2020). Column descriptions can be found in the legend of Appendix Table A2.1.

Identity	Up Reg (17,072)	IG (100)	ORA *p*	ORA *q*	Mean FC IG	Mean FC Rest	GSEA *p*	GSEA *q*
FAPs	8270	69	2.43E-05	**0.0003**	5.59	4.25	0.1473	0.4419
Smooth Muscle Cells	5761	45	0.0124	0.1492	5.11	3.73	0.3303	0.9909
MuSCs and Progenitors	8191	59	0.0173	0.2074	8.12	2.91	2.00E-04	**6.00E-04**
Neural, Glial & Schwann Cells	2182	20	0.0269	0.3225	5.60	26.73	-	-
Mature Skeletal Muscle	2221	16	0.2239	1	4.84	10.91	-	-
Endothelial	2888	20	0.2400	1	3.03	5.43	-	-
Tenocytes	2655	17	0.3853	1	2.52	7.33	-	-
Monocytes, Macrophages & Platelets	4754	17	0.9960	1	1.46	1.77	-	-
Pro-inflammatory Macrophages	2984	8	0.9980	1	1.32	1.92	-	-

Anti-inflammatory Macrophages	3686	10	0.9993	1	2.48	2.12	-	-
B, T, NK Cells	2590	5	0.9996	1	4.04	11.61	-	-
Resident Macrophages APCs	3580	8	0.9999	1	1.50	3.77	-	-

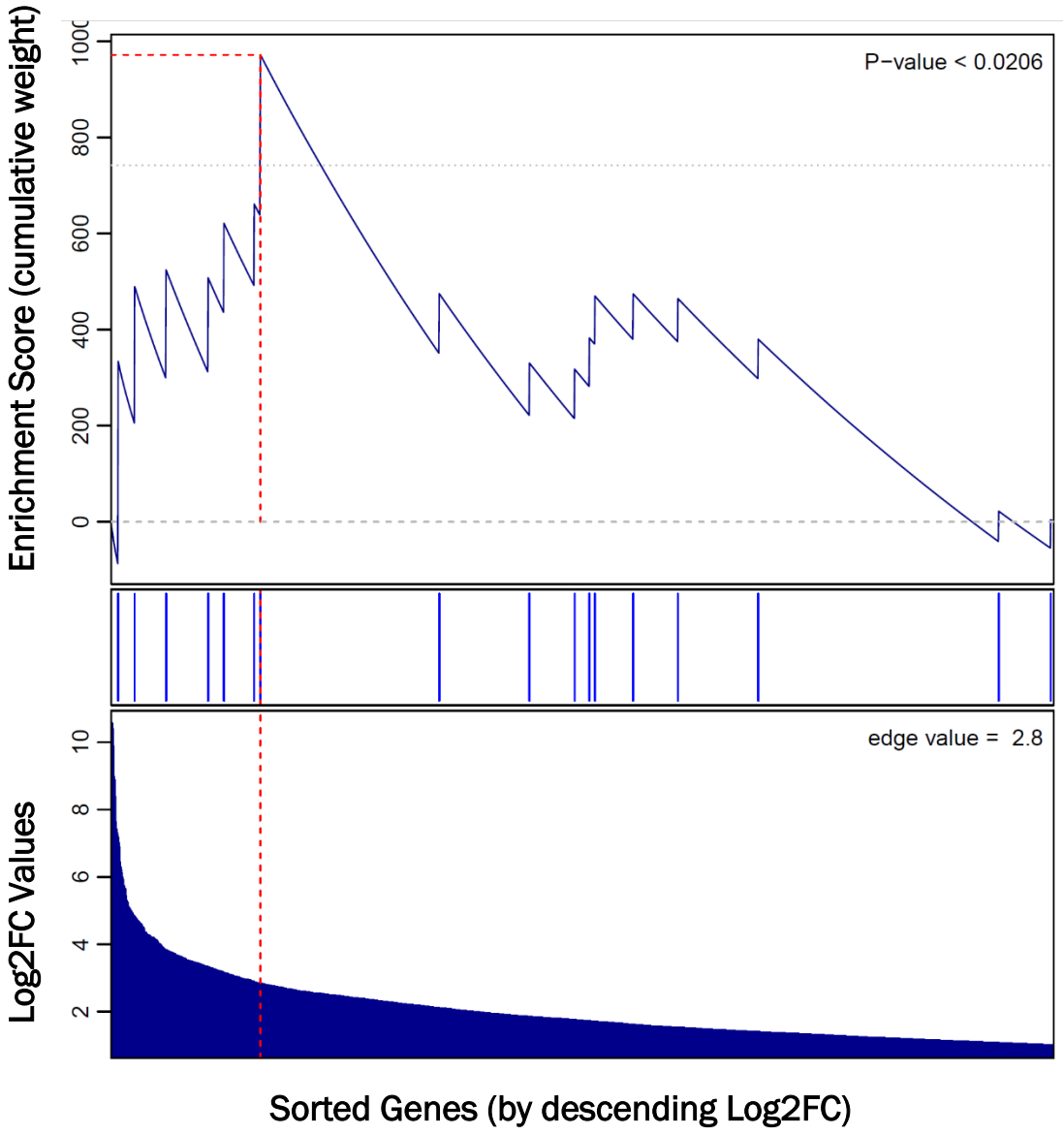
Of further interest in the mammary gland was whether imprinted genes were over-represented at different stages of the mammary gland cycle (Virgin, Gestation, Lactation and Involution) (Hanin & Ferguson-Smith, 2020). The MCA dataset sequencing mammary gland tissue at these different stages allowed me to analyse whether imprinted genes were over-represented in the mammary gland cells at a specific time point. When analysing all cells at the different time points above, there were no over-represented time points (Appendix Table A3.21A). The largest number of IGs were seen during involution and pregnancy but this was not an over-representation. When analysing all cell types at each pregnancy timepoint, stromal cells were the only over-represented cell for each timepoint (Appendix Table A3.21B**).** A larger mean FC was seen for imprinted genes in stromal cells at all time points (bar lactation), but a significant GSEA was only found for pregnancy mammary stromal cells, and this was also the cell type with the largest number of imprinted genes upregulated. Analysis of the epithelial cells only (Bach et al., 2017) at the four mammary periods saw no significant over-representation for any time-point (Appendix Table A3.21C).

Appendix Table A2.7. Imprinted gene over-representation in Mammary Gland Epithelial cells (Bach et al., 2017). Column descriptions can be found in the legend of Appendix Table A2.1.

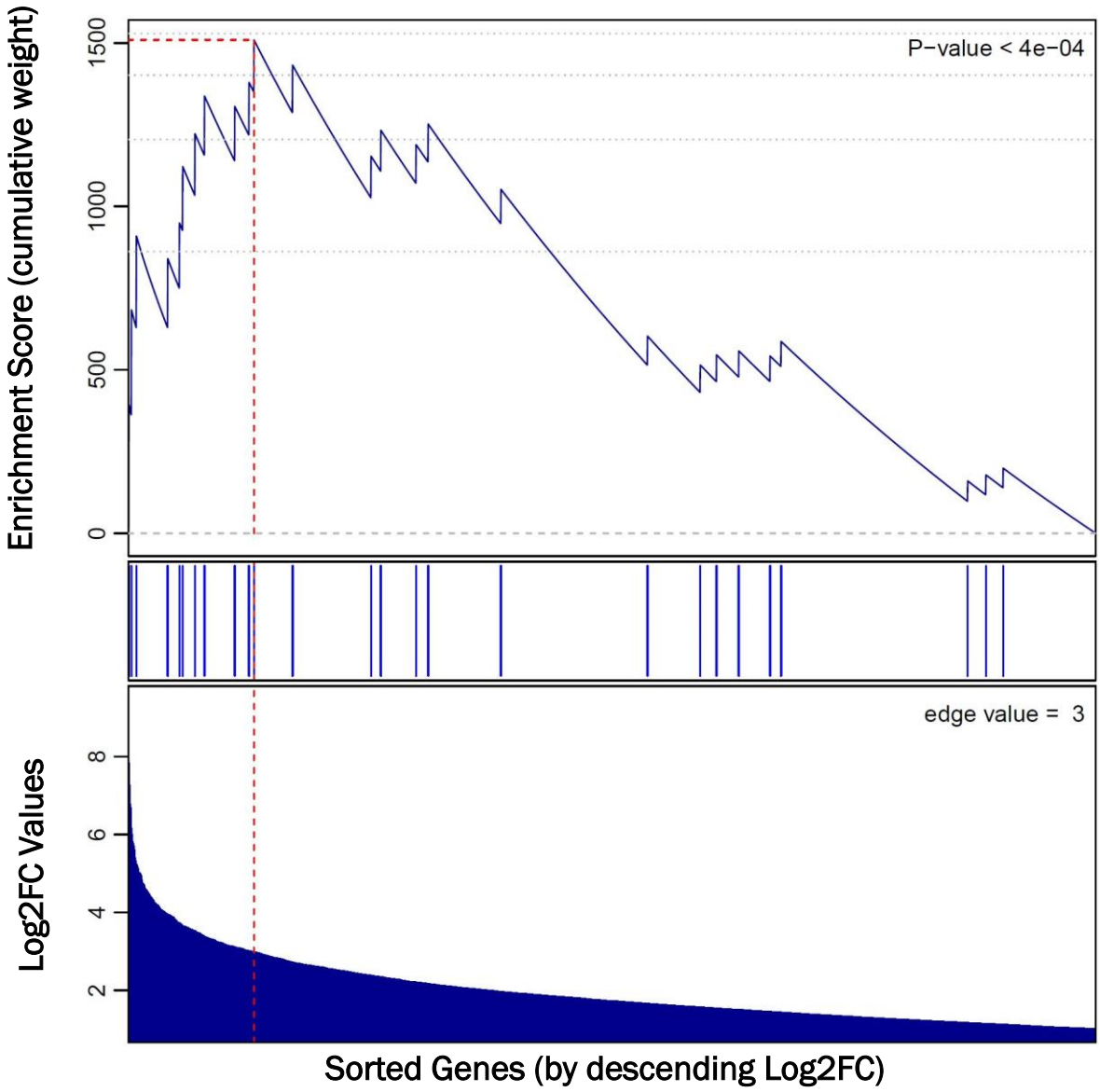
Identity	Up Reg (16,415)	IG (96)	ORA p	ORA q	Mean FC IG	Mean FC Rest	GSEA p
Mammary Gland Stem (Procr+) cells	1267	21	1.06E-05	**8.46E-05**	25.11	30.79	-
Basal cells	1446	18	0.0016	**0.0132**	6.88	6.22	0.20
Luminal hormone sensing differentiated	2120	18	0.0649	0.5194	5.39	6.63	-
Myoepithelial cells	2343	15	0.3952	1	6.42	8.48	-
Luminal hormone sensing progenitors	1217	8	0.4194	1	2.40	2.66	-
Luminal progenitor cells	982	5	0.6876	1	3.74	4.13	-
Luminal alveolar cells differentiated	2228	11	0.7705	1	5.80	8.05	-
Luminal alveolar cells progenitor	1113	5	0.7879	1	3.02	4.88	-

Appendix A3 – Additional Tables & Figures from Chapter 3 and Appendix A2

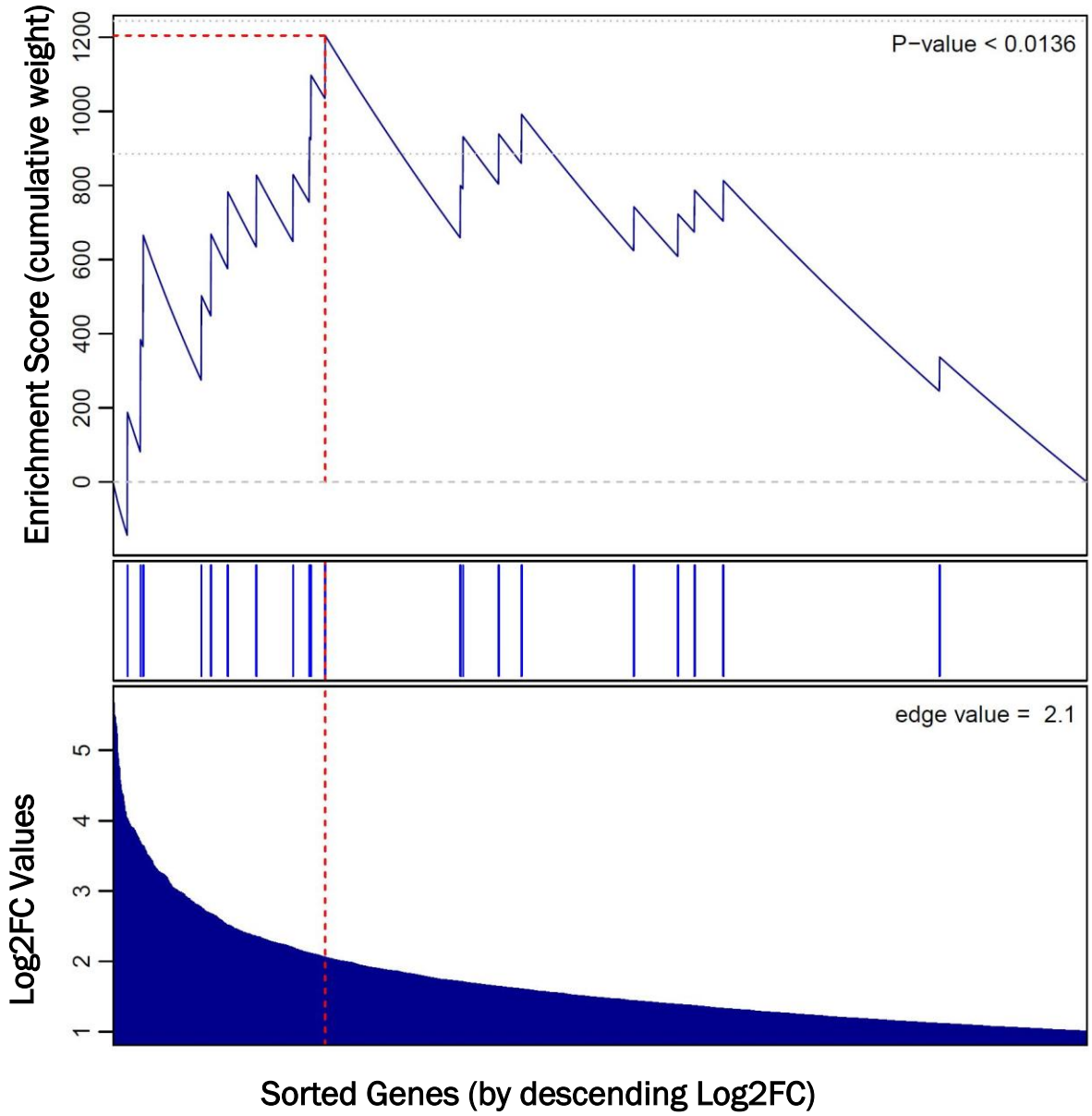
Chapter 3 Additional Figures



Appendix Figure A3.1. Significant Gene Set Enrichment Analysis (GSEA) for Paternally Expressed Genes (PEGs) in the cell of the Pancreas in the Mouse Cell Atlas (Han et al., 2018). In the GSEA analysis, genes are sorted by strength by which they mark this neuronal cluster (sorted by fold change values) indicated by the bar (middle). Fold change values are displayed along the bottom of the graph. The genes are arrayed left (strongest marker) to right and blue lines mark where imprinted genes fall on this array. The vertical axis indicates an accumulating weight, progressing from left to right and increasing or decreasing depending on whether the next gene is an imprinted gene or not. The *p*-value represents the probability of observing the maximum value of the score (red dashed line) if the imprinted genes are distributed randomly along the horizontal axis.



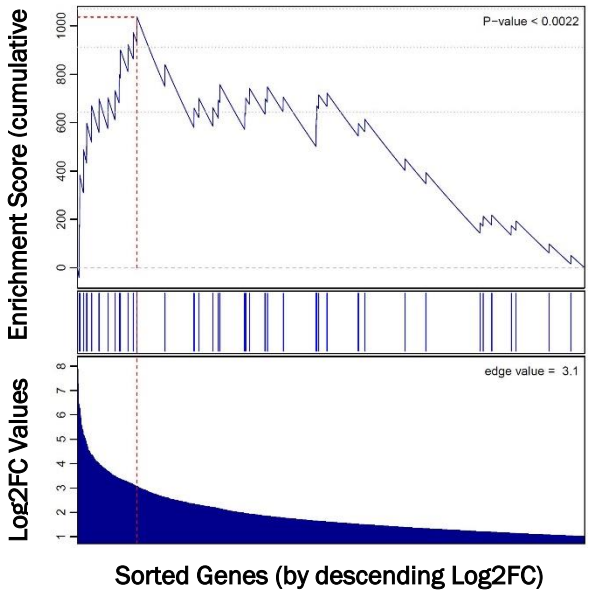
Appendix Figure A3.2 Significant GSEA for PEGs in Neuroendocrine Cells in a Whole Mouse Brain Analysis (Ximerakis et al., 2019). See the legend of Appendix Figure A3.1 for description on how to interpret GSEA analysis and plot.



Appendix Figure A3.3 Significant GSEA for Maternally Expressed Genes (MEGS) in the cells of the Pons in the Mouse Brain Atlas Analysis (Zeisel et al., 2018), *q*-value = 0.0272. See the legend of Appendix Figure A3.1 for description on how to interpret GSEA analysis and plot.

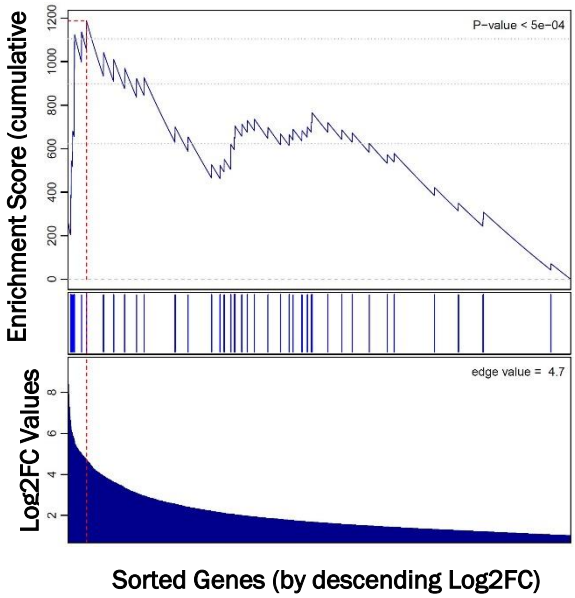
A

HBSER2 (*q* = 0.0462)



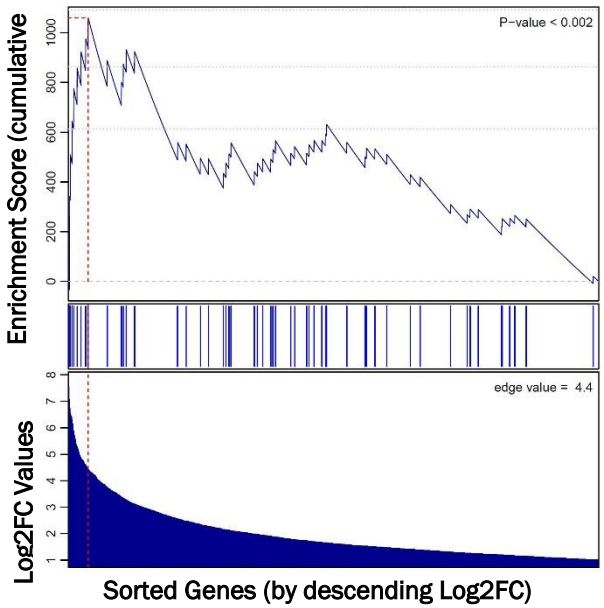
B

HBSER4 (*q* = 0.011)

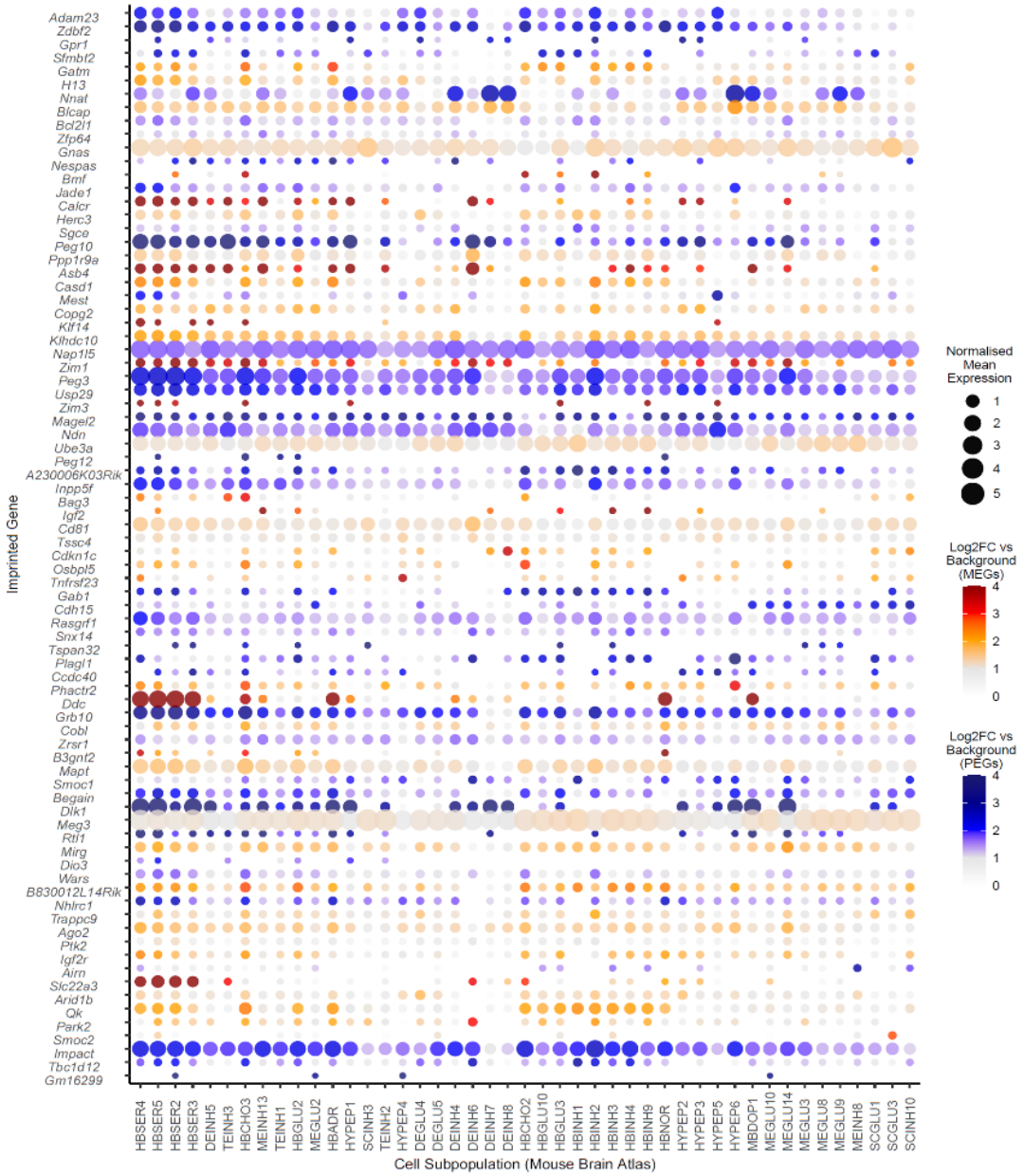


C

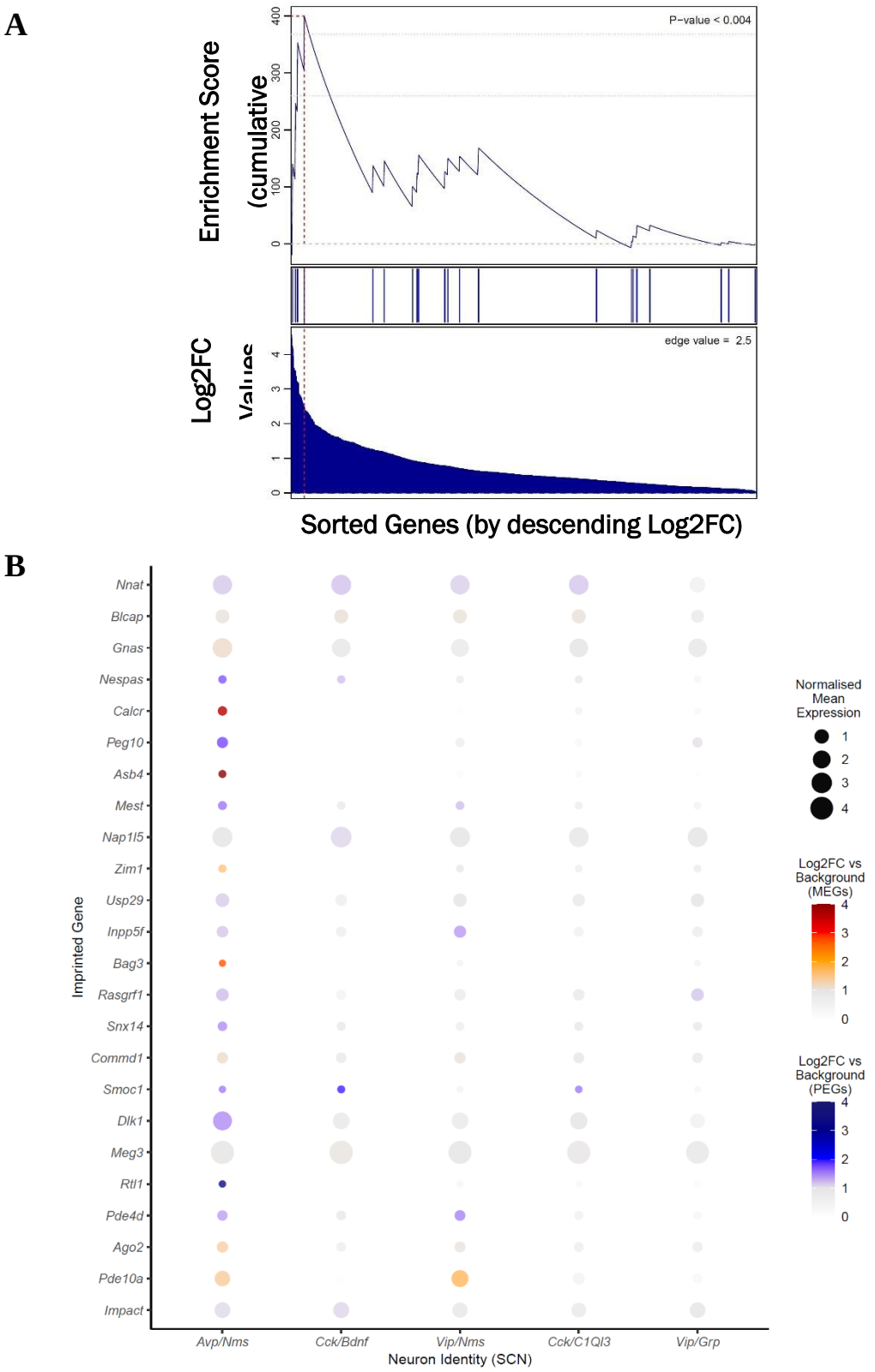
HBSER5 (*q* = 0.042)



D

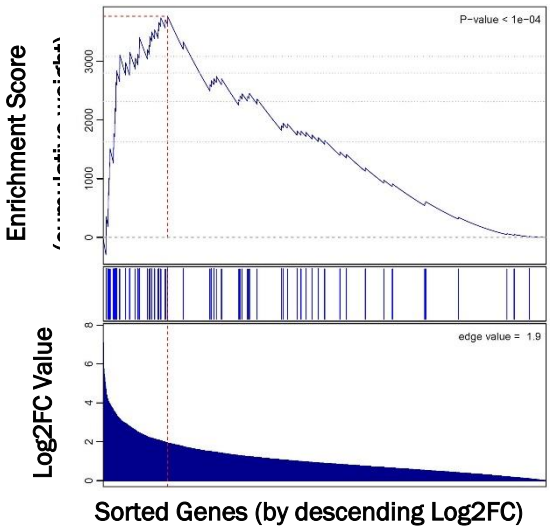


Appendix Figure A3.4. Significant GSEA for imprinted genes in three subpopulations of Hindbrain Serotonin Neurons (HBSER2, HBSER4, HBSER5) in Mouse Brain Atlas (Zeisel et al., 2018). (*A-C***) -** GSEA graph for the three enriched serotonin neuron populations. See the legend of Appendix Figure A3.1 for description on how to interpret GSEA analysis and plot. *(D)* – Dot plot of imprinted genes upregulated in any of the 3 enriched serotonin neuron populations plotted across all over-represented neuron subpopulations found in the Mouse Brain Atlas. Imprinted genes were plotted in chromosomal order. Size of points represented absolute mean expression; colour represented the size of the Log2FC value for the cell identity group (e.g., HBSER4) vs. all other cells. Unique colour scales are used for MEGs (red/orange) and PEGs (blue). Where a gene was not expressed in a cell type, this appears as a blank space in the plot

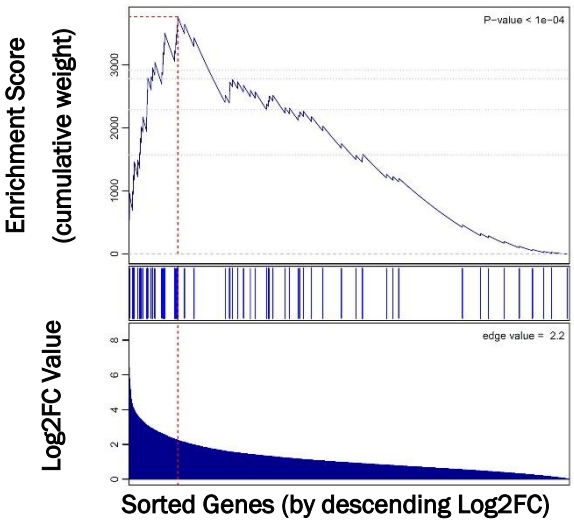


Appendix Figure A3.5. Significant GSEA for imprinted genes in *Avp/Nms* neurons in Suprachiasmatic Nucleus (Wen et al., 2020). *(A)* GSEA graph for *Avp/Nms* neurons. See legend of Appendix Figure A3.1 for description on how to interpret GSEA analysis and plot. *(B)* – Dot plot of imprinted genes upregulated in *Avp/Nms* neurons plotted across the five neuron types identified in the SCN. See legend of Appendix Figure A3.4 for a description of how to interpret the dot plot.

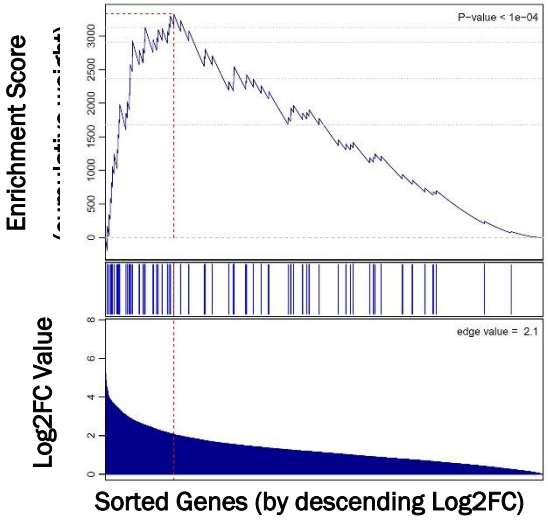
A **Serotonergic Neurons (q = 0.0005)**



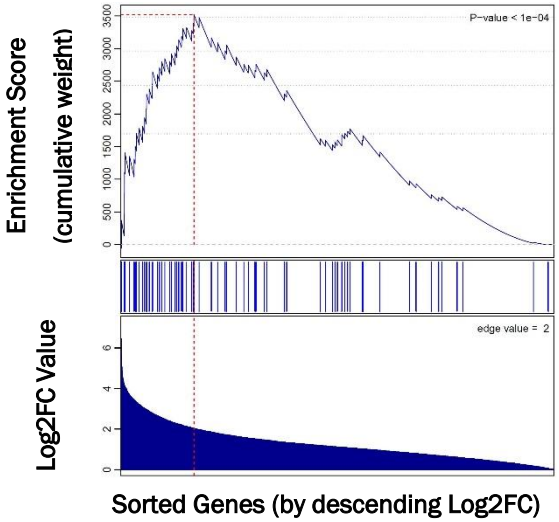
B **Dopaminergic Neurons (*q* = 0.0005)**



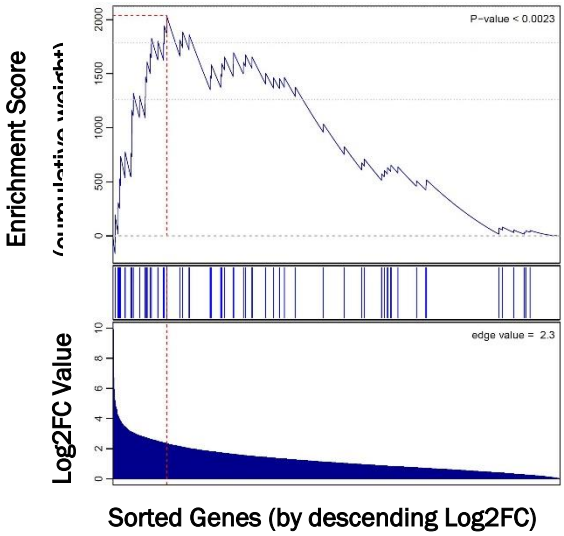
C **GABAergic Neurons (*q* = 0.003)**



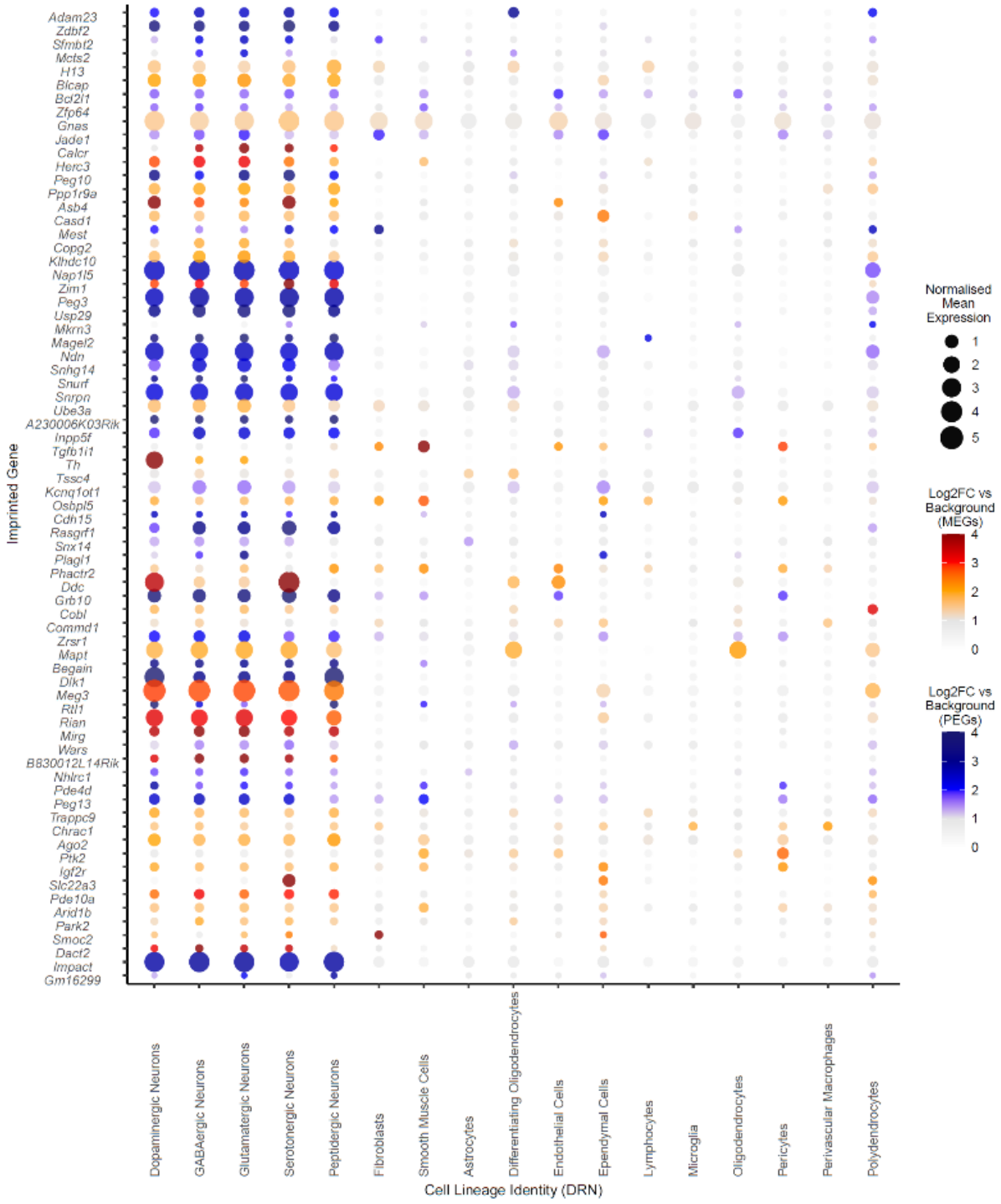
D **Glutamatergic Neurons (*q* = 0.0035)**



E **Peptidergic Neurons (*q* = 0.034)**



F



Appendix Figure A3.6. Significant GSEA for imprinted genes in neuronal types in the Dorsal Raphe Nucleus (Huang et al., 2019) (*A-E*) – GSEA graphs for Serotonin neurons, Dopamine neurons, GABA neurons, Glutamatergic neurons, Peptidergic neurons. See the legend of Appendix Figure A3.1 for description on how to interpret GSEA analysis and plot. *(F)* Dot plot of the expression of all imprinted genes, upregulated in any of the 5 neuron types, plotted across all cells in the DRN. See the legend of Appendix Figure A3.4 for details of the plot.

Chapter 3 Additional Tables

Appendix Table A3.1 – Imprinted gene list used in Analyses

Appendix Table A3.1. Custom List of Imprinted Genes used as the gene set in this study. *Ensembl* – Ensembl ID, *Gene* – Gene name, *Allelic Bias* – Maternal (M) or Paternal (P) bias of the gene, *Chrom* – Chromosome the gene occupies, *Evidence for Imprinting* – References showing this gene is imprinted, genes included in our analysis with only one demonstration of POE are highlighted in red, *POE in brain?* – whether the gene has shown a parent of origin effect in brain tissue (Y/_). Genes were not included in the list based on POE in the brain.

Ensembl	Gene	Allelic Bias	Chrom	Evidence for Imprinting	POE in Brain?
ENSMUSG00000106223	*2400006E01Rik*	P	3	Andergassen 2017, Wanigasuriya 2020	
ENSMUSG00000068151	*A230006K03Rik*	P	7	Shen 2014, Andergassen 2017, Laukoter 2018	Y
ENSMUSG00000109394	*A230057D06Rik*	P	7	Babak 2008, Lorenc 2014	Y
ENSMUSG00000025964	*Adam23*	P	1	Gregg 2010, DeVeale 2012, Perez 2015, Babak 2015	Y
ENSMUSG00000036698	*Ago2*	M	15	Gregg 2010, Bonthius 2015, Andergassen 2017	Y
ENSMUSG00000078247	*Airn*	P	17	Wutz 1997, Babak 2015, Andergassen 2017	Y
ENSMUSG00000031075	*Ano1*	M	7	Okae 2012, Andergassen 2017	
ENSMUSG00000069729	*Arid1b*	M	17	Andergassen 2017, Gigantea 2018	
ENSMUSG00000042607	*Asb4*	M	6	Mizuno 2002, Perez 2015, Babak 2015, Andergassen 2017	Y
ENSMUSG00000009248	*Ascl2*	M	7	Guillemot 1995, Andergassen 2017	
ENSMUSG00000051650	*B3gnt2*	M	11	Perez 2015 - multiple brain regions & body tissues	Y
ENSMUSG00000098202	*B830012L14Rik*	M	12	Hagan 2009, Zhang 2011, Babak 2015	Y
ENSMUSG00000030847	*Bag3*	M	7	Perez 2015, Bonthius 2015	Y
ENSMUSG00000007659	*Bcl2l1*	P	2	Gregg 2010, DeVeale, 2012, Babak 2015	Y
ENSMUSG00000040867	*Begain*	P	12	Tierling 2009, Perez 2015, Babak 2015	Y
ENSMUSG00000067787	*Blcap*	M	2	Schulz 2008, Wang, 2008, Babak 2015, Andergassen 2017	Y
ENSMUSG00000040093	*Bmf*	M	2	Perez 2015 - multiple brain regions & body tissues	Y
ENSMUSG00000023964	*Calcr*	M	6	Hoshiya, 2003, Babak 2015, Andergassen 2017	Y
ENSMUSG00000015189	*Casd1*	M	6	Ono 2003, Babak 2008, Okae 2012	
ENSMUSG00000039963	*Ccdc40*	P	11	Gregg 2010, DeVeale 2012	Y
ENSMUSG00000037706	*Cd81*	M	7	Caspary 1998, Andergassen 2017	
ENSMUSG00000031962	*Cdh15*	P	8	Gregg 2010, Proudhon 2012, Bonthius 2015, Babak 2015	Y
ENSMUSG00000037664	*Cdkn1c*	M	7	Hatada 1995, Babak 2015, Andergassen 2017	Y
ENSMUSG00000068391	*Chrac1*	M	15	Camprubi 2014, Perez 2015	Y
ENSMUSG00000020173	*Cobl*	M	11	Shiura 2009, Bonthius 2015	Y
ENSMUSG00000051355	*Commd1*	M	11	Wang 2004, Babak 2015, Andergassen 2017	Y
ENSMUSG00000025607	*Copg2*	M	6	Lee 2000, Perez 2015, Babak 2015	Y
ENSMUSG00000048826	*Dact2*	M	17	Calabrese 2015, Babak 2015, Andergassen 2017	
ENSMUSG00000020182	*Ddc*	P	11	Menheniott 2008, Bonthius 2015, Babak 2015, Andergassen 2017	Y
ENSMUSG00000075707	*Dio3*	P	12	Tsai 2002, Hernandez 2018	Y
ENSMUSG00000040856	*Dlk1*	P	12	Schmidt 2000, Babak 2015, Andergassen 2017	Y
ENSMUSG00000040013	*Fkbp6*	P	5	Perez 2015, Strogantsev 2015, Andergassen 2017	Y
ENSMUSG00000031714	*Gab1*	P	8	Okae 2012, Inoue 2017, Andergassen 2017	
ENSMUSG00000027199	*Gatm*	M	2	Sandell 2003, Okae 2012	
ENSMUSG00000089679	*Gm16299*	P	19	Perez 2015 - multiple brain regions & body tissues	Y

ENSMUSG00000027523	*Gnas*	M	2	Williamson 1996, Babak 2015, Andergassen 2017	Y
ENSMUSG00000046856	*Gpr1*	P	1	Hiura 2010, Okae 2012	
ENSMUSG00000020176	*Grb10*	M	11	Miyoshi 1998, Arnaud 2003, Perez 2015, Andergassen 2017	Y
ENSMUSG00000019188	*H13*	M	2	Wood 2007, Wang 2011, Babak 2015, Andergassen 2017	Y
ENSMUSG00000000031	*H19*	M	7	Bartolomei 1991, Lorenc 2014, Babak 2015	Y
ENSMUSG00000029804	*Herc3*	M	6	Cowley, 2012, Perez 2015, Babak 2015	Y
ENSMUSG00000034997	*Htr2a*	M	14	Kato 1998, Babak 2015	
ENSMUSG00000029096	*Htra3*	M	5	Nie 2006, Wang 2011	
ENSMUSG00000048583	*Igf2*	P	7	DeChiara 1991, Perez 2015, Babak 2015, Andergassen 2017	Y
ENSMUSG00000086266	*Igf2os*	P	7	Moore 1997, Babak 2015, Andergassen 2017	Y
ENSMUSG00000023830	*Igf2r*	M	17	Barlow 1991, Babak 2015, Perez 2015	Y
ENSMUSG00000024423	*Impact*	P	18	Hagiwara 1997, Babak 2015, Andergassen 2017	Y
ENSMUSG00000042105	*Inpp5f*	P	7	Choi 2005, Andergassen 2017	Y
ENSMUSG00000000215	*Ins2*	P	7	Giddings 1994, Deltour 1995, Schulz 2006	
ENSMUSG00000025764	*Jade1*	P	3	Wang 2011, Inoue 2017, Andergassen 2017	
ENSMUSG00000036760	*Kcnk9*	M	15	Ruf 2007, Perez 2015, Andergassen 2017, Cooper 2020	Y
ENSMUSG00000009545	*Kcnq1*	M	7	Gould 1998, Babak 2015, Andergassen 2017	Y
ENSMUSG00000101609	*Kcnq1ot1*	P	7	Smilinich 1999, Babak 2015, Andergassen 2017	Y
ENSMUSG00000073209	*Klf14*	M	6	Parker-Katiraee 2007, Wang 2011, Babak 2015, Andergassen 2017	
ENSMUSG00000029775	*Klhdc10*	M	6	Babak 2008, Gregg 2010, Lorenc 2014, Perez 2015	Y
ENSMUSG00000056972	*Magel2*	P	7	Boccaccio 1999, Babak 2015, Andergassen 2017	Y
ENSMUSG00000018411	*Mapt*	M	11	Gregg 2010, DeVeale 2012, Babak 2015	Y
ENSMUSG00000042814	*Mcts2*	P	2	Wood 2007, Babak 2015	Y
ENSMUSG00000021268	*Meg3*	M	12	Schuster-Gossler 1996, Schmidt 2000, Andergassen 2017	Y
ENSMUSG00000051855	*Mest*	P	6	Kaneko-Ishino 1995, Babak 2015, Andergassen 2017	Y
ENSMUSG00000097391	*Mirg*	M	12	Seitz, 2003, Babak 2015, Andergassen 2017	Y
ENSMUSG00000070527	*Mkrn3*	P	7	Jong 1999, Babak 2015, Andergassen 2017	Y
ENSMUSG00000055430	*Nap1l5*	P	6	Smith 2003, Babak 2015, Andergassen 2017	Y
ENSMUSG00000087090	*Nctc1*	P	7	Eun 2013, Kim 2019	
ENSMUSG00000033585	*Ndn*	P	7	MacDonald, 1997, Babak 2015, Andergassen 2017	Y
ENSMUSG00000086537	*Nespas*	P	2	Wroe 2000, Babak 2015, Andergassen 2017	Y
ENSMUSG00000044231	*Nhlrc1*	P	13	Perez 2015, Bonthius 2015	Y
ENSMUSG00000067786	*Nnat*	P	2	Kagitani 1997, Okae 2012, Perez 2015	Y
ENSMUSG00000037606	*Osbpl5*	M	7	Engeman 2000, Okae 2012, Andergassen 2017	
ENSMUSG00000023826	*Park2*	M	17	Calabrese 2015, Andergassen 2017	
ENSMUSG00000023868	*Pde10a*	M	17	Wang 2011, Andergassen 2017	
ENSMUSG00000021699	*Pde4d*	P	13	Babak 2008, Perez 2015	Y
ENSMUSG00000092035	*Peg10*	P	6	Ono 2003, Babak 2015, Andergassen 2017	Y
ENSMUSG00000070526	*Peg12*	P	7	Chai 2001, Kobayashi 2012, DeVeale 2012, Babak 2015, Andergassen 2017	Y
ENSMUSG00000106847	*Peg13*	P	15	Smith 2003, Wang 2011, Andergassen 2017	Y
ENSMUSG00000002265	*Peg3*	P	7	Kaneko-Ishino 1995, Babak 2015, Andergassen 2017	Y
ENSMUSG00000062866	*Phactr2*	M	10	Wang 2011, Bonthius 2015, Andergassen 2017	Y
ENSMUSG00000010760	*Phlda2*	M	7	Qian 1997, Babak 2015, Andergassen 2017	
ENSMUSG00000019817	*Plagl1*	P	10	Piras 2000, Babak 2015, Andergassen 2017	Y
ENSMUSG00000086646	*Platr20*	P	11	Calabrese 2015, Andergassen 2017	

ENSMUSG00000029759	*Pon3*	M	6	Monk 2008, Andergassen 2017	
ENSMUSG00000032827	*Ppp1r9a*	M	6	Ono, 2003, Lorenc 2014, Babak 2015	Y
ENSMUSG00000022607	*Ptk2*	M	15	Perez 2015 - multiple brain regions & body tissues	Y
ENSMUSG00000062078	*Qk*	M	17	Calabrese 2015, Andergassen 2017	
ENSMUSG00000032356	*Rasgrf1*	P	9	Plass 1996, Andergassen 2017	Y
ENSMUSG00000097451	*Rian*	M	12	Hatada, 2001, Babak 2015, Andergassen 2017	Y
ENSMUSG00000085925	*Rtl1*	P	12	Seitz, 2003, Lorenc 2014, Babak 2015, Andergassen 2017	Y
ENSMUSG00000061186	*Sfmbt2*	P	2	Kuzmin 2008, Babak 2015 Inoue 2017	
ENSMUSG00000004631	*Sgce*	P	6	Piras 2000, Monk 2008, Babak 2015, Andergassen 2017	Y
ENSMUSG00000000154	*Slc22a18*	M	7	Dao 1998, Andergassen 2017	
ENSMUSG00000040966	*Slc22a2*	M	17	Zwart 2001, Babak 2015, Andergassen 2017	
ENSMUSG00000023828	*Slc22a3*	M	17	Zwart 2001, Babak 2015, Andergassen 2017	
ENSMUSG00000022464	*Slc38a4*	P	15	Mizuno 2002, Babak 2015, Andergassen 2017	Y
ENSMUSG00000021136	*Smoc1*	P	12	Andergassen 2017, Inoue 2017	
ENSMUSG00000023886	*Smoc2*	M	17	Andergassen 2017, Gigante 2018	
ENSMUSG00000100826	*Snhg14*	P	7	Nikaido 2003, Andergassen 2017	Y
ENSMUSG00000102252	*Snrpn*	P	7	Leff 1992, Babak 2015, Andergassen 2017	Y
ENSMUSG00000102627	*Snurf*	P	7	Gray 1999, Andergassen 2017	Y
ENSMUSG00000032422	*Snx14*	P	9	Huang 2014, Thomas 2014	
ENSMUSG00000048720	*Tbc1d12*	P	18	Babak 2008, Perez 2015	
ENSMUSG00000029664	*Tfpi2*	M	6	Monk 2008, Okae 2012	
ENSMUSG00000030782	*Tgfb1i1*	M		Perez 2015, Bonthius 2015	Y
ENSMUSG00000000214	*Th*	M	7	Schulz, 2006, Okae 2012, Bonthius 2015	Y
ENSMUSG00000023885	*Thbs2*	M	17	Andergassen 2017, Inoue 2017	
ENSMUSG00000037613	*Tnfrsf23*	M	7	Clark, 2002, de la Casa Esperon 2012	
ENSMUSG00000047921	*Trappc9*	M	15	Gregg 2010, Perez 2015, Andergassen 2017	Y
ENSMUSG00000000244	*Tspan32*	P	10	Umlauf 2004, Wang 2011, Okae 2012	
ENSMUSG00000045752	*Tssc4*	M	7	Paulsen 2000, Babak 2015, Andergassen 2017	
ENSMUSG00000059585	*Ube2nl*	P	7	Babak 2015, Perez 2015	Y
ENSMUSG00000025326	*Ube3a*	M	7	Albrecht 1997, Babak 2015, Andergassen 2017	Y
ENSMUSG00000051527	*Usp29*	P	7	Szeto 2000, Andergassen 2017	Y
ENSMUSG00000021266	*Wars*	P	12	Gregg 2010, DeVeale 2012, Babak 2015	Y
ENSMUSG00000027520	*Zdbf2*	P	1	Kobayashi 2009, Okae 2012, Babak 2015, Andergassen 2017	Y
ENSMUSG00000109176	*Zfp264*	P	7	Kim 2001, He 2014	Y
ENSMUSG00000027551	*Zfp64*	P	2	Wang, 2011, Andergassen 2017	
ENSMUSG00000002266	*Zim1*	M	7	Kim 1999, Babak 2015, Andergassen 2017	Y
ENSMUSG00000108113	*Zim2*	M	7	Kim 2004, He 2014	Y
ENSMUSG00000108043	*Zim3*	M	7	Kim 2001, Babak 2015	Y
ENSMUSG00000044068	*Zrsr1*	P	11	Hatada 1993, Babak 2015, Andergassen 2017	Y

Appendix Table A3.2. Average Normalised Expression

Appendix Table A3.2. The average normalised expression for all imprinted genes across the identity groups in the datasets for analyses at Levels 1 and 2. Analysis was carried out for the Mouse Cell Atlas (MCA, (Han et al., 2018)) and *Tabula Muris* (Schaum et al., 2018) datasets from Level 1 (Multi-Organ Comparison) and the Mouse Brain Atlas (Zeisel et al., 2018) and Ximerakis et al. (2019) datasets from Level 2 (Whole Brain Comparison). Tables for each dataset are presented in the same spreadsheet. *Tissue/Cell Type/Nervous System Region* – the identity groups that the cells were grouped under. *IG Mean Normalised Expression* – The average normalised expression value for all imprinted genes in cells of that identity group. *Rest Mean Normalised Expression* – The average normalised expression value for all other genes excluding the imprinted genes to provide a baseline expression value.

LEVEL 1

MCA (Han et al., 2018)			*Tabula Muris (Schaum et al., 2018)*		
Tissue	**IG Mean Normalised Expression**	**Rest Mean Normalised Expression**	**Tissue**	**IG Mean Normalised Expression**	**Rest Mean Normalised Expression**
Brain	0.32	0.19	Pancreas	1.06	0.70
Pancreas	0.30	0.18	Bladder	1.02	0.85
Bladder	0.24	0.23	Trachea	0.87	0.64
Uterus	0.21	0.20	Brain (Non-Myeloid)	0.83	0.59
Ovary	0.19	0.21	Limb Muscle	0.75	0.48
Kidney	0.16	0.17	Tongue	0.71	0.85
Lung	0.14	0.17	Heart	0.71	0.53
Liver	0.13	0.16	Mammary Gland	0.71	0.68
Muscle	0.13	0.19	Diaphragm	0.69	0.42
Stomach	0.13	0.21	Fat	0.67	0.57
Small Intestine	0.12	0.20	Large Intestine	0.61	0.69
Mammary Gland	0.11	0.17	Lung	0.57	0.49
Testis	0.11	0.21	Liver	0.56	0.58
Thymus	0.10	0.18	Aorta	0.55	0.45
Bone Marrow	0.10	0.17	Skin	0.54	0.65
Peripheral Blood	0.09	0.17	Brain (Myeloid)	0.50	0.42
Spleen	0.08	0.16	Marrow	0.49	0.59
Prostate	0.07	0.09	Kidney	0.40	0.32
			Thymus	0.33	0.41
			Spleen	0.29	0.38

LEVEL 2

Whole Brain (Ximerakis et al., 2019)			Zeisel Mouse Brain Atlas (Cell Lineages)		
Cell Type	**IG Mean Normalised Expression**	**Rest Mean Normalised Expression**	**Cell Type**	**IG Mean Normalised Expression**	**Rest Mean Normalised Expression**
Neuroendocrine cells (NendC)	0.46	0.25	Neurons	0.19	0.10

Mature Neurons (all types) (mNEUR)	0.29	0.20	Ependymal	0.15	0.10
Oligodendrocyte precursor cells (OPC)	0.29	0.20	Oligos	0.14	0.08
Tanycytes (TNC)	0.29	0.21	Peripheral Glia	0.12	0.09
Choroid plexus epithelial cells (CPC)	0.27	0.24	Vascular	0.11	0.09
Vascular smooth muscle cells (VSMC)	0.27	0.25	Astrocytes	0.10	0.07
Ependymocytes (EPC)	0.27	0.25	Immune	0.08	0.07
Oligodendrocytes (OLG)	0.26	0.20	**Zeisel Mouse Brain Atlas (Nervous System Regions)**		
Arachnoid barrier cells (ABC)	0.26	0.24	**Nervous System Region**	**IG Mean Normalised Expression**	**Rest Mean Normalised Expression**
Vascular and leptomeningeal cells (VLMC)	0.26	0.21	Pons	0.31	0.12
Astrocyte-restricted precursors (ARP)	0.24	0.22	Medulla	0.30	0.11
Olfactory ensheathing glia (OEG)	0.24	0.19	Ventral Midbrain	0.28	0.10
Pericytes (PC)	0.23	0.21	Hypothalamus	0.28	0.12
Hypendymal cells (HypEPC)	0.22	0.23	Dorsal Midbrain	0.25	0.10
Neural stem cells (NSC)	0.22	0.20	Thalamus	0.24	0.09
Immature Neurons (ImmN)	0.22	0.17	CA1	0.23	0.09
Hemoglobin-expressing vascular cells (Hb_VC)	0.21	0.22	Spinal Cord	0.22	0.12
Neuronal-restricted precursor (NRP)	0.21	0.23	Antero-Middle Cortex	0.21	0.09
Astrocytes (ASC)	0.19	0.18	Enteric Nervous System	0.21	0.12
Endothelial cells (EC)	0.18	0.19	SS cortex	0.21	0.09
Microglia (MG)	0.15	0.18	Posterior Cortex	0.20	0.10
Macrophages (MAC)	0.12	0.17	Middle Cortex	0.20	0.10
Dendritic cells (DC)	0.11	0.20	Anterior Cortex	0.20	0.09
			Sympathetic	0.20	0.12
			Ventral Striatum	0.20	0.10
			Amygdala	0.19	0.10
			Dorsal Striatum	0.18	0.09
			Dorsal Root Ganglia	0.17	0.12
			Hippocampus	0.16	0.11
			Olfactory Bulb	0.15	0.12
			Dentate Gyrus	0.14	0.12
			Cerebellum	0.12	0.12

Appendix Table A3.3 (A & B) - MEG/PEG analysis in Mouse Cell Atlas (MCA, (Han et al., 2018))

Appendix Table A3.3A, **Maternally Expressed Genes (MEG) Multi-Organ Level Enrichment Analysis for the Mouse Cell Atlas (MCA, (Han et al., 2018)) dataset.** *Up Reg* – number of upregulated genes with $q \leq 0.05$ and Log2FC ≥ 1 (total number of genes in the dataset in brackets); *MEG* – number of maternally expressed genes upregulated with $q \leq 0.05$ and Log2FC ≥ 1 (total number of MEGs in the dataset in brackets); *ORA p – p* value from over representation analysis on groups with minimum 5% of total MEGs; *ORA q* – Bonferroni corrected *p* value from ORA; *Mean FC MEG* – mean fold change for upregulated MEGs; *Mean FC Rest* – mean fold change for all other upregulated genes; *GSEA p – p* value from Gene Set Enrichment Analysis for identity groups with 15+ MEGs and Mean FC MEG > Mean FC Rest; *GSEA q* – Bonferroni corrected *p* values from GSEA; *No. MEGs with highest expression* – Number of MEGs with highest mean expression value for that identity group.

Identity	Up Reg (20,534)	MEG (53)	ORA *p*	ORA *q*	Mean FC MEG	Mean FC Rest	No. MEGs with highest expression
Pancreas	2737	24	1.39E-08	**1.53E-07**	4.54	10.35	13
Bladder	3183	19	0.0002	**0.0025**	4.26	8.50	5
Kidney	1714	10	0.0113	0.1245	13.76	182.89	5
Uterus	2567	12	0.0281	0.3090	4.44	8.45	4
Lung	1203	6	0.0884	0.9723	4.11	151.16	4
Brain	3401	13	0.0887	0.9753	9.04	124.28	6
Liver	1739	6	0.2909	1	3.30	80.42	1
Small Intestine	1719	5	0.4596	1	7.99	218.64	2
Stomach	1821	5	0.5116	1	4.71	88.50	2
Ovary	2219	4	0.8387	1	2.73	11.26	2
Testis	5212	8	0.9756	1	6.87	5052.59	5
Mammary Gland	902	2	-	-	5.35	4.02	0
Bone Marrow	1095	1	-	-	6.36	4.43	1
Muscle	1127	1	-	-	15.81	15.03	2
Peripheral Blood	1146	1	-	-	4.61	3.57	0
Thymus	1805	1	-	-	5.06	6.75	1
Prostate	369	0	-	-	0.00	478.10	0
Spleen	1501	0	-	-	0.00	4.77	0

Appendix Table A3.3B. Paternally Expressed Genes (PEG) Multi-Organ Level Enrichment Analysis for the MCA (Han et al., 2018) dataset. *Up Reg* – number of upregulated genes with $q \leq 0.05$ and Log2FC ≥ 1 (total number of genes in the dataset in brackets); *PEG* – number of paternally expressed genes upregulated with $q \leq 0.05$ and Log2FC ≥ 1 (total number of PEGs in the dataset in brackets); *ORA p – p* value from over representation analysis on groups with minimum 5% of total PEGs; *ORA q* – Bonferroni corrected *p* value from ORA; *Mean FC PEG* – mean fold change for upregulated PEGs; *Mean FC Rest* – mean fold change for all other upregulated genes; *GSEA p – p* value from Gene Set Enrichment Analysis for identity groups with 15+ PEGs and Mean FC PEG > Mean FC Rest; *GSEA q* – Bonferroni corrected *p* values from GSEA; *No. PEGs with highest expression* – Number of PEGs with highest mean expression value for that identity group

Identity	Up Reg (20,534)	PEG (42)	ORA *p*	ORA *q*	Mean FC PEG	Mean FC Rest	GSEA *p*	No. PEGs with highest express…
Brain	3401	21	5.70E-07	**4.56E-06**	8.59	124.56	-	13
Pancreas	2737	18	2.41E-06	**1.93E-05**	14.33	10.27	**0.021**	9
Uterus	2567	10	0.0312	0.2496	4.93	8.44	-	3
Ovary	2219	9	0.0328	0.2624	9.57	11.26	-	3
Bladder	3183	10	0.1049	0.8389	4.82	8.49	-	3
Thymus	1805	5	0.3082	1	2.32	6.76	-	1
Muscle	1127	3	0.4076	1	6.25	15.05	-	1
Testis	5212	6	0.9730	1	53.94	5050.60	-	5
Liver	1739	2	-	-	8.32	80.24	-	2
Lung	1203	2	-	-	2.96	150.68	-	0
Mammary Gland	902	2	-	-	2.06	4.03	-	0
Peripheral Blood	1146	2	-	-	3.37	3.57	-	0
Stomach	1821	2	-	-	3.07	88.37	-	1
Bone Marrow	1095	1	-	-	4.26	4.43	-	0
Spleen	1501	1	-	-	4.90	4.77	-	1
Kidney	1714	0	-	-	0.00	181.90	-	0
Prostate	369	0	-	-	0.00	478.10	-	0
Small Intestine	1719	0	-	-	0.00	218.03	-	0

Appendix Table A3.4 (A & B) – MEG/PEG analysis in *Tabula Muris* (Schaum et al., 2018).

Appendix Table A3.4A. MEG Multi-Organ Level Enrichment Analysis for the *Tabula Muris* (Schaum et al., 2018). Column descriptions can be found in the legend of Appendix Table A3.3A.

Identity	Up Reg (20,839)	MEG (58)	ORA *p*	ORA *q*	Mean FC MEG	Mean FC Rest	No. MEGs with highest expression
Diaphragm	416	12	1.33E-09	**2.13E-08**	5.28	4.89	3
Limb Muscle	761	14	1.52E-08	**2.43E-07**	8.15	5.16	4
Trachea	1979	15	0.0003	**0.0040**	3.22	4.57	3
Bladder	3338	19	0.0012	**0.0198**	2.98	5.30	10
Pancreas	4104	20	0.0059	0.0938	12.20	12.60	14
Heart	1108	7	0.0334	0.5346	3.03	5.13	0
Mammary Gland	1826	8	0.1321	1	3.36	5.24	2
Fat	1263	6	0.1378	1	2.47	3.69	0
Brain (Non-Myeloid)	3081	12	0.1401	1	4.99	14.19	4
Brain (Myeloid)	1024	5	0.1552	1	3.39	6.80	2
Kidney	584	3	0.2217	1	24.94	22.90	1
Tongue	4295	11	0.6725	1	4.42	7.16	6
Aorta	3515	8	0.7845	1	7.34	16.07	1
Skin	1612	3	0.8365	1	3.88	8.36	1
Large Intestine	4758	10	0.8822	1	6.34	12.22	4
Liver	1808	3	0.8894	1	4.88	54.82	1
Lung	914	2	-	-	2.73	5.41	0
Marrow	1957	2	-	-	10.31	5.25	2
Spleen	625	0	-	-	0.00	4.28	0
Thymus	678	0	-	-	0.00	7.42	0

Appendix Table A3.4B. PEG Multi-Organ Level Enrichment Analysis for the *Tabula Muris* (Schaum et al., 2018). Column descriptions can be found in the legend of Appendix Table A3.3B.

Identity	Up Reg (20,839)	PEG (49)	ORA *p*	ORA *q*	Mean FC PEG	Mean FC Rest	GSEA *p*	GSEA *q*	No. PEGs with highest expression
Limb_Muscle	761	10	8.98E-06	**0.0001**	10.25	5.15	-	-	4
Pancreas	4104	23	1.51E-05	**0.0002**	12.80	12.60	0.1283	0.2566	15
Brain (Non-Myeloid)	3081	19	3.42E-05	**0.0005**	16.69	14.13	0.3785	0.757	10
Diaphragm	416	7	5.01E-05	**0.0007**	8.57	4.84	-	-	1
Trachea	1979	10	0.0154	0.2150	4.69	4.56	-	-	2
Fat	1263	6	0.0744	1	4.46	3.68	-	-	1
Thymus	678	4	0.0746	1	3.46	7.45	-	-	1
Bladder	3338	12	0.0823	1	3.81	5.29	-	-	6
Heart	1108	3	0.4872	1	2.50	5.12	-	-	0
Liver	1808	4	0.6247	1	7.17	54.84	-	-	2
Mammary Gland	1826	4	0.6327	1	3.83	5.23	-	-	1
Marrow	1957	3	0.8512	1	5.88	5.25	-	-	2
Aorta	3515	6	0.8564	1	7.63	16.06	-	-	1
Tongue	4295	4	0.9949	1	3.40	7.16	-	-	2
Large Intestine	4758	1	-	-	2.11	12.21	-	-	1
Skin	1612	1	-	-	4.94	8.35	-	-	0
Spleen	625	1	-	-	4.63	4.28	-	-	0
Brain (Myeloid)	1024	0	-	-	0.00	6.78	-	-	0
Kidney	584	0	-	-	0.00	22.91	-	-	0
Lung	914	0	-	-	0.00	5.40	-	-	0

Appendix Table A3.5 (A & B) MEG/PEG analysis of Brain Cell (Ximerakis et al., 2019)

Appendix Table A3.5A. MEG over-representation in neural lineage types (Ximerakis et al., 2019). Column descriptions can be found in the legend of Appendix Table A3.3A.

Brain Cell Identity (Abbr.)	Up Reg (14,498)	MEG (45)	ORA *p*	ORA *q*	Mean FC MEG	Mean FC Rest	GSEA *p*	GSEA *q*	No. MEGs with highest expression
Arachnoid barrier cells (ABC)	2287	16	0.0009	**0.0139**	19.04	22.60	-	-	5
Neuroendocrine cells (NendC)	3868	21	0.0031	**0.0466**	6.30	5.49	0.3804	0.7608	10
Pericytes (PC)	1801	11	0.0194	0.2910	7.84	8.23	-	-	0
Mature Neurons (all types) (mNEUR)	2968	15	0.0301	0.4519	12.06	9.26	0.2256	0.4512	2
Astrocyte-restricted precursors (ARP)	1445	6	0.2895	1	4.47	5.09	-	-	1

Choroid plexus epithelial cells (CPC)	2602	12	0.0954	1	8.10	19.31	-	-	5
Endothelial cells (EC)	1455	6	0.2950	1	6.11	8.54	-	-	0
Ependymocytes (EPC)	3233	7	0.9022	1	4.58	53.18	-	-	2
Hemoglobin-expressing vascular cells (Hb_VC)	1798	7	0.3218	1	6.20	6.32	-	-	1
Hypendymal cells (HypEPC)	1525	3	0.8660	1	9.59	20.81	-	-	3
Oligodendrocytes (OLG)	1183	5	0.3044	1	3.82	12.88	-	-	3
Oligodendrocyte precursor cells (OPC)	1524	6	0.3339	1	3.11	7.15	-	-	1
Tanycytes (TNC)	1279	5	0.3648	1	6.08	11.98	-	-	3
Vascular and leptomeningeal cells (VLMC)	1714	9	0.0775	1	20.02	13.01	-	-	1
Vascular smooth muscle cells (VSMC)	3006	10	0.4608	1	12.77	6.71	-	-	4
Astrocytes (ASC)	1384	1	-	-	2.38	6.04	-	-	0
Dendritic cells (DC)	1209	1	-	-	3.50	16.02	-	-	1
Immature Neurons (ImmN)	652	1	-	-	3.20	5.78	-	-	0
Macrophages (MAC)	1222	2	-	-	3.47	21.56	-	-	0
Microglia (MG)	1342	2	-	-	26.89	19.21	-	-	3
Monocytes (MNC)	947	0	-	-	0.00	19.12	-	-	0
Neutrophils (NEUT)	519	0	-	-	0.00	61.92	-	-	0
Neuronal-restricted precursor (NRP)	2339	2	-	-	2.51	10.18	-	-	0
Neural stem cells (NSC)	1009	1	-	-	2.11	4.09	-	-	0
Olfactory ensheathing glia (OEG)	1086	2	-	-	12.12	25.90	-	-	0

Appendix Table A3.5B. PEG over-representation in neural lineage types (Ximerakis et al., 2019). Column descriptions can be found in the legend of Appendix Table A3.3B.

Brain Cell Identity (Abbr.)	Up Reg (14,498)	PEG (40)	ORA *p*	ORA *q*	Mean FC PEG	Mean FC Rest	GSEA *p*	No. PEGs with highest expression
Neuroendocrine cells (NendC)	3868	26	4.27E-07	**8.97E-06**	16.39	5.42	**4.00E-04**	16
Mature Neurons (all types) (mNEUR)	2968	17	0.0013	**0.0269**	5.92	9.30	-	0
Olfactory ensheathing glia (OEG)	1086	7	0.0274	0.5757	6.76	26.00	-	1
Tanycytes (TNC)	1279	7	0.0583	1	7.04	11.99	-	5
Oligodendrocyte precursor cells (OPC)	1524	7	0.1211	1	2.97	7.15	-	0
Immature Neurons (ImmN)	652	3	0.2677	1	3.42	5.79	-	0
Neural stem cells (NSC)	1009	4	0.3025	1	4.48	4.08	-	0
Neuronal-restricted precursor (NRP)	2339	8	0.3125	1	3.21	10.20	-	1
Vascular and leptomeningeal cells (VLMC)	1714	6	0.3332	1	7.63	13.07	-	3
Ependymocytes (EPC)	3233	10	0.3998	1	22.37	53.17	-	2
Oligodendrocytes (OLG)	1183	4	0.4141	1	3.60	12.87	-	2
Neutrophils (NEUT)	519	2	0.4220	1	9.18	62.13	-	0
Monocytes (MNC)	947	2	0.7459	1	16.49	19.13	-	1

Hemoglobin-expressing vascular cells (Hb_VC)	1798	4	0.7482	1	3.58	6.33	-	2
Astrocytes (ASC)	1384	3	0.7490	1	2.16	6.04	-	0
Vascular smooth muscle cells (VSMC)	3006	7	0.7517	1	3.46	6.73	-	1
Hypendymal cells (HypEPC)	1525	3	0.8071	1	24.89	20.78	-	2
Choroid plexus epithelial cells (CPC)	2602	5	0.8686	1	5.83	19.28	-	0
Pericytes (PC)	1801	3	0.8893	1	9.50	8.22	-	2
Arachnoid barrier cells (ABC)	2287	4	0.8954	1	8.04	22.60	-	2
Astrocyte-restricted precursors (ARP)	1445	2	0.9189	1	4.57	5.08	-	0
Dendritic cells (DC)	1209	0	-	-	0.00	16.01	-	0
Endothelial cells (EC)	1455	1	-	-	3.97	8.53	-	0
Macrophages (MAC)	1222	0	-	-	0.00	21.53	-	0
Microglia (MG)	1342	1	-	-	4.06	19.23	-	0

Appendix Table A3.6 (A & B) MEG/PEG analysis of Nervous System Cell (Zeisel et al., 2018)

Appendix Table A3.6A. MEG over-representation in nervous system cell types (Zeisel et al., 2018). Column descriptions can be found in the legend of Appendix Table A3.3A.

Nervous System Cell Identity	Up Reg (19,547)	MEG (59)	ORA *p*	ORA *q*	Mean FC MEG	Mean FC Rest	No. MEGs with highest expression
Vascular	2473	19	7.37E-05	**0.0004**	18.52	26.62	14
Neurons	5710	22	0.1122	0.6730	12.60	24.91	21
Ependymal	3683	13	0.3128	1	4.14	66.96	9
Immune	1564	5	0.5150	1	15.51	92.94	3
Oligos	1587	5	0.5283	1	4.42	11.45	5
Peripheral Glia	2820	8	0.6316	1	5.57	12.62	5
Astrocytes	1539	2	-	-	3.14	10.72	2

Appendix Table A3.6B. PEG over-representation in nervous system cell types (Zeisel et al., 2018). Column descriptions can be found in the legend of Appendix Table A3.3B**.**

Nervous System Cell Identity	Up Reg (19,547)	PEG (49)	ORA *p*	ORA *q*	Mean FC PEG	Mean FC Rest	No. PEGs with highest expression
Neurons	5710	22	0.0184	0.0919	10.87	24.92	24
Oligos	1587	6	0.2175	1	4.82	11.45	7
Peripheral Glia	2820	8	0.4353	1	5.28	12.62	7
Ependymal	3683	7	0.8565	1	62.36	66.75	6
Vascular	2473	3	0.9610	1	14.06	26.57	2
Astrocytes	1539	2	-	-	2.63	10.72	1
Immune	1564	2	-	-	8.21	92.80	2

Appendix Table A3.7 (A & B) MEG/PEG analysis of Nervous System Regions (Zeisel et al., 2018)

Appendix Table A3.7A. MEG over-representation in nervous system regions (Zeisel et al., 2018). Column descriptions can be found in the legend of Appendix Table A3.3A.

Nervous System Region Identity	Up Reg (18,335)	MEG (57)	ORA *p*	ORA *q*	Mean FC MEG	Mean FC Rest	GSEA *p*	GSEA *q*	No. MEGs with highest expression
Medulla	3147	21	0.0003	**0.0048**	5.70	4.01	0.0318	0.0636	9
Pons	3581	20	0.0042	0.0680	5.40	3.90	0.0136	**0.0272**	10
Vent. Midbrain	1228	8	0.0352	0.5631	5.95	4.98	-	-	1
Hypothalamus	1040	7	0.0414	0.6625	6.01	5.82	-	-	2
Middle Cortex	623	5	0.0439	0.7017	3.29	3.24	-	-	0
Enteric Nervous System	3885	17	0.0793	1	9.97	120.77	-	-	7
Sympathetic Nervous System	2804	12	0.1523	1	14.88	57.85	-	-	6

Vent. Striatum	689	4	0.1659	1	4.54	4.91	-	-	0
Anterior Cortex	979	5	0.1873	1	2.82	3.30	-	-	1
Somatosensory Cortex	2121	9	0.2085	1	4.46	3.70	-	-	7
Posterior Cortex	1090	5	0.2499	1	2.35	3.20	-	-	0
Hippocampus - CA1	1082	4	0.4362	1	3.48	4.02	-	-	2
Thalamus	1441	5	0.4679	1	3.33	6.35	-	-	0
Dors. Striatum	1196	4	0.5149	1	4.41	5.43	-	-	2
Dorsal Root Ganglion	3607	11	0.5814	1	13.29	75.79	-	-	5
Spinal Cord	972	3	0.5882	1	5.53	12.34	-	-	0
Olfactory Bulb	445	2	-	-	3.92	8.25	-	-	1
Antero-Middle Cortex	646	1	-	-	5.97	4.31	-	-	1
Dentate Gyrus	796	1	-	-	7.62	4.16	-	-	1
Hippocampus	631	1	-	-	4.36	3.83	-	-	0
Amygdala	452	0	-	-	0.00	4.11	-	-	0
Cerebellum	240	0	-	-	0.00	32.30	-	-	0
Dors. Midbrain	1045	0	-	-	0.00	4.83	-	-	2

Appendix Table A3.7B. PEG over-representation in nervous system regions (Zeisel et al., 2018). Column descriptions can be found in the legend of Appendix Table A3.3B.

Nervous System Region Identity	Up Reg (18,335)	PEG (48)	ORA *p*	ORA *q*	Mean FC PEG	Mean FC Rest	No. PEGs with highest expression
Hypothalamus	1040	15	4.67E-08	**6.53E-07**	4.42	5.85	6
Medulla	3147	24	2.93E-07	**4.10E-06**	4.00	4.03	6
Pons	3581	24	3.32E-06	**4.65E-05**	3.21	3.91	12
Vent. Midbrain	1228	10	0.0013	**0.0182**	4.06	4.99	2
Amygdala	452	4	0.0323	0.4522	4.65	4.11	2
Dors. Midbrain	1045	6	0.0586	0.8206	2.20	4.85	1
Vent. Striatum	689	4	0.1115	1	3.31	4.92	0
Hippocampus	631	3	0.2376	1	5.03	3.82	2
Posterior Cortex	1090	4	0.3326	1	3.01	3.20	2
Dentate Gyrus	796	3	0.3585	1	2.52	4.17	1
Enteric Nervous System	3885	9	0.7386	1	7.12	120.55	4
Thalamus	1441	3	0.7515	1	2.19	6.35	0
Sympathetic Nervous System	2804	6	0.7808	1	4.37	57.78	3
Dorsal Root Ganglion	3607	5	0.9758	1	7.74	75.70	4
Hippocampus - CA1	1082	2	-	-	2.06	4.02	0
Antero-Middle Cortex	646	2	-	-	3.91	4.31	0
Olfactory Bulb	445	2	-	-	4.12	8.25	1
Spinal Cord	972	2	-	-	3.13	12.33	1
Somatosensory Cortex	2121	2	-	-	2.45	3.71	1
Dors. Striatum	1196	2	-	-	3.26	5.43	0
Anterior Cortex	979	1	-	-	2.24	3.30	0
Cerebellum	240	0	-	-	0.00	32.30	0
Middle Cortex	623	0	-	-	0.00	3.24	0

Appendix Table A3.8 – Whole Nervous System Neuron Enrichment - Zeisel et al. (2018)

Appendix Table A3.8. Imprinted gene over-representation in neuronal subpopulations in the MBA dataset (Zeisel et al., 2018). *Neuron Description* – more specific classification of neuronal subpopulations as described by Zeisel et al. (2018); *Up Reg* – number of upregulated genes with $q \leq 0.05$ and Log2FC ≥ 1 (total number of genes in the dataset in brackets); *IG* – number of imprinted genes upregulated with $q \leq 0.05$ and Log2FC ≥ 1 (total number of IGs in the dataset in brackets); *ORA p – p* value from over representation analysis on groups with minimum 5% of total IGs; *ORA q* – Bonferroni corrected *p* value from ORA; *Mean FC IG* – mean fold change for upregulated imprinted genes; *Mean FC Rest* – mean fold change for all other upregulated genes; *GSEA p – p* value from Gene Set Enrichment Analysis for identity groups with 15+ IGs and Mean FC IG > Mean FC Rest; *GSEA q* – Bonferroni corrected *p* values from GSEA.

Neuron Identity	Neuron Description	Up Reg (18,335)	IG (105)	ORA *p*	ORA *q*	Mean FC IG	Mean FC Rest	GSEA *p*	GSEA *q*

HBSER5	Serotonergic neurons, hindbrain	3721	53	8.29E-12	**1.19E-09**	13.76	5.98	0.002	**0.042**
DEINH5	Peptidergic neurons, hypothalamus	611	19	1.94E-09	**2.79E-07**	7.47	5.48	0.0639	1
TEINH3	Inhibitory neurons, telencephalon	885	22	5.50E-09	**7.93E-07**	8.60	6.38	0.2076	1
MEINH13	Inhibitory neurons, midbrain	1229	25	2.35E-08	**3.38E-06**	9.85	7.30	0.1213	1
HBSER4	Serotonergic neurons, hindbrain	3510	45	2.83E-08	**4.08E-06**	27.93	6.29	5.00E-04	**0.0105**
HBGLU2	Excitatory neurons, hindbrain	2533	37	3.33E-08	**4.80E-06**	4.08	4.88	-	-
HBCHO3	Afferent nuclei of cranial nerves VI-XII	4647	53	4.03E-08	**5.81E-06**	7.61	5.64	0.1175	1
HYPEP1	Peptidergic neurons, hypothalamus	766	19	7.34E-08	**1.06E-05**	9.92	9.46	0.2588	1
DEINH6	Peptidergic neurons, hypothalamus	466	15	7.43E-08	**1.07E-05**	16.64	18.97	-	-
HBSER3	Serotonergic neurons, hindbrain	2154	33	8.19E-08	**1.18E-05**	12.84	6.30	0.0127	0.2667
HBGLU3	Excitatory neurons, hindbrain	1438	26	1.24E-07	**1.78E-05**	5.62	6.24	-	-
MBDOP1	Dopaminergic neurons, periaqueductal grey	639	17	1.45E-07	**2.09E-05**	12.07	16.84	-	-
HYPEP3	Peptidergic neurons, hypothalamus	814	19	1.89E-07	**2.72E-05**	4.25	7.82	-	-
DEINH4	Inhibitory neurons, thalamus	913	19	1.08E-06	**0.0002**	4.20	10.29	-	-
TEINH2	Inhibitory neurons, septal nucleus	511	14	1.44E-06	**0.0002**	6.32	7.45	-	-
HBSER2	Serotonergic neurons, hindbrain	3497	41	2.00E-06	**0.0003**	10.24	5.32	0.0022	**0.0462**
HBINH9	Inhibitory neurons, hindbrain	1461	24	2.34E-06	**0.0003**	4.62	7.17	-	-
DEINH7	Inhibitory neurons, hypothalamus	259	10	2.58E-06	**0.0004**	7.72	15.76	-	-
HYPEP2	Peptidergic neurons, hypothalamus	465	13	2.83E-06	**0.0004**	4.84	16.74	-	-
HBINH4	Inhibitory neurons, hindbrain	1639	25	5.18E-06	**0.0007**	3.82	6.90	-	-
HBINH2	Inhibitory neurons, hindbrain	3492	40	5.23E-06	**0.0008**	4.67	6.62	-	-
HBNOR	Noradrenergic neurons of the medulla	1611	24	1.25E-05	**0.0018**	10.24	12.90	-	-
HBADR	Adrenergic cell groups of the medulla	2574	32	1.37E-05	**0.0020**	8.27	7.82	0.2581	1
HYPEP6	Orexin-producing neurons, hypothalamus	1542	23	1.93E-05	**0.0028**	6.12	9.58	-	-
TEINH1	Inhibitory neurons, pallidum	991	17	5.16E-05	**0.0074**	3.08	3.74	-	-
MEGLU2	Excitatory neurons, midbrain	1097	18	5.39E-05	**0.0078**	3.31	4.43	-	-
HBINH5	Inhibitory neurons, hindbrain	3288	36	5.43E-05	**0.0078**	3.35	3.74	-	-
DEINH8	Inhibitory neurons, hypothalamus	308	9	7.48E-05	**0.0108**	7.18	15.65	-	-
MBDOP2	Dopaminergic neurons, ventral midbrain (SNc, VTA)	1494	21	0.0001	**0.0161**	11.13	9.29	0.1175	1
MEGLU10	Excitatory neurons, midbrain	792	14	0.0002	**0.0267**	2.83	4.69	-	-
DEINH3	Inhibitory neurons, hypothalamus	1347	19	0.0002	**0.0345**	3.30	3.53	-	-
HYPEP4	Oxytocin-producing cells, hypothalamus	923	15	0.0003	**0.0386**	8.67	9.35	-	-
MEGLU14	Glutamatergic projection neurons of the raphe nucleus	1832	23	0.0003	**0.0388**	7.76	15.15	-	-
MEGLU9	Excitatory neurons, midbrain	661	12	0.0004	0.0626	2.89	6.26	-	-
MEINH14	Inhibitory neurons, midbrain	571	11	0.0005	0.0670	4.42	13.65	-	-
HBINH3	Inhibitory neurons, hindbrain	872	14	0.0005	0.0709	5.31	9.29	-	-
HBGLU6	Excitatory neurons, hindbrain	3518	35	0.0005	0.0719	4.15	4.83	-	-
HBINH6	Inhibitory neurons, hindbrain	3125	31	0.0013	0.1811	4.47	5.65	-	-
HBGLU4	Excitatory neurons, hindbrain	2578	27	0.0013	0.1855	6.87	8.38	-	-
MEINH3	Inhibitory neurons, midbrain	1104	15	0.0017	0.2426	2.76	4.22	-	-
SCGLU1	Excitatory neurons, spinal cord	583	10	0.0020	0.2909	4.04	13.41	-	-
HYPEP5	Vasopressin-producing cells, hypothalamus	820	12	0.0027	0.3958	6.43	7.82	-	-
DECHO1	Cholinergic neurons, septal nucleus, Meissnert and diagonal band	2442	25	0.0028	0.4013	6.23	6.16	0.5821	1
HBCHO2	Cholinergic neurons, hindbrain	3436	32	0.0029	0.4216	4.19	6.44	-	-
HBINH1	Inhibitory neurons, hindbrain	549	9	0.0045	0.6547	6.10	23.68	-	-

HBCHO4	Afferent nuclei of cranial nerves III-V	2967	28	0.0048	0.6951	3.70	6.04	-	-
HBGLU7	Excitatory neurons, hindbrain	3277	30	0.0054	0.7705	3.76	4.72	-	-
SCGLU3	Excitatory neurons, spinal cord	782	11	0.0055	0.7973	3.78	10.65	-	-
HBSER1	Serotonergic neurons, hindbrain	2200	22	0.0069	0.9916	11.02	6.79	0.0218	0.4578
MEGLU1	Excitatory neurons, midbrain	1425	16	0.0077	1	2.65	4.86	-	-
HYPEP7	Pmch neurons, hypothalamus	1088	13	0.0099	1	4.24	13.35	-	-
HBGLU9	Excitatory neurons, hindbrain	2739	25	0.0121	1	5.13	5.17	-	-
DEGLU4	Excitatory neurons, thalamus	1185	13	0.0191	1	3.17	5.53	-	-
HBGLU10	Excitatory neurons, hindbrain	708	9	0.0215	1	3.91	11.03	-	-
HBINH7	Inhibitory neurons, hindbrain	1881	18	0.0221	1	3.90	7.83	-	-
MEINH4	Inhibitory neurons, midbrain	854	10	0.0260	1	5.14	5.34	-	-
HBGLU8	Excitatory neurons, hindbrain	3561	29	0.0293	1	6.04	6.53	-	-
MEGLU3	Excitatory neurons, midbrain	755	9	0.0308	1	2.62	6.77	-	-
OBDOP1	Dopaminergic periglomerular interneuron, olfactory bulb	415	6	0.0334	1	17.15	8.20	-	-
SCINH10	Inhibitory neurons, spinal cord	811	9	0.0452	1	3.61	10.54	-	-
SCINH3	Inhibitory neurons, spinal cord	703	8	0.0510	1	14.95	10.00	-	-
MEINH12	Inhibitory neurons, midbrain	708	8	0.0527	1	2.91	8.24	-	-
TEINH5	Interneuron-selective interneurons, cortex/hippocampus	488	6	0.0638	1	2.23	5.77	-	-
HBCHO1	Cholinergic neurons, hindbrain	2586	21	0.0649	1	4.12	10.35	-	-
DEGLU5	Excitatory neurons, midbrain	619	7	0.0676	1	2.69	4.43	-	-
SCINH5	Inhibitory neurons, spinal cord	1354	12	0.0911	1	6.46	9.99	-	-
MEINH2	Inhibitory neurons, midbrain	1511	13	0.0959	1	2.75	3.62	-	-
HYPEP8	Peptidergic neurons, hypothalamus	962	9	0.1046	1	7.50	24.02	-	-
MEGLU11	Excitatory neurons, midbrain	846	8	0.1166	1	5.02	5.56	-	-
TEINH10	R-LM border Cck interneurons, cortex/hippocampus	1141	10	0.1235	1	4.02	4.23	-	-
SYNOR3	Noradrenergic neurons, sympathetic	2667	20	0.1310	1	9.48	7.66	0.1842	1
ENT4	Cholinergic enteric neurons	3465	25	0.1341	1	13.28	11.93	0.3431	1
DEGLU3	Excitatory neurons, thalamus	1619	13	0.1413	1	3.41	5.56	-	-
ENT7	Cholinergic enteric neurons, VGLUT2	3011	22	0.1416	1	6.73	10.49	-	-
MEGLU4	Excitatory neurons, midbrain	758	7	0.1493	1	6.03	6.26	-	-
HBGLU1	Excitatory neurons, hindbrain	3042	22	0.1528	1	4.74	6.15	-	-
SCGLU5	Excitatory neurons, spinal cord	801	7	0.1814	1	4.79	9.01	-	-
SCINH8	Inhibitory neurons, spinal cord	811	7	0.1892	1	7.91	13.50	-	-
MEINH10	Inhibitory neurons, midbrain	1107	9	0.1902	1	3.22	6.76	-	-
ENT8	Cholinergic enteric neurons, VGLUT2	3029	21	0.2132	1	5.22	8.83	-	-
CBINH2	Granular layer interneurons, cerebellum	696	6	0.2151	1	17.05	14.64	-	-
SCINH11	Central canal neurons, spinal cord	859	7	0.2287	1	53.31	29.03	-	-
SYNOR5	Noradrenergic erector muscle neurons	2259	16	0.2293	1	20.13	9.80	0.0425	0.8925
TEINH8	Interneuron-selective interneurons, hippocampus	870	7	0.2382	1	3.30	7.01	-	-
TEGLU21	Excitatory neurons, hippocampus CA1	1680	12	0.2635	1	4.09	4.42	-	-
SCGLU4	Excitatory neurons, spinal cord	750	6	0.2666	1	3.39	8.48	-	-
TEGLU12	Excitatory neurons, cerebral cortex	913	7	0.2763	1	4.18	5.95	-	-
TEGLU10	Excitatory neurons, cerebral cortex	1394	10	0.2855	1	2.86	3.53	-	-
TEINH19	Hippocamposeptal projection, cortex/hippocampus	1087	8	0.2924	1	5.18	5.04	-	-

TEGLU20	Excitatory neurons, cerebral cortex	1407	10	0.2953	1	7.44	5.82	-	-
MEGLU6	Excitatory neurons, midbrain	1584	11	0.3077	1	2.98	4.83	-	-
MSN1	D1 medium spiny neurons, striatum	1112	8	0.3139	1	4.56	5.13	-	-
DECHO2	Cholinergic neurons, habenula	966	7	0.3253	1	12.45	15.47	-	-
ENT2	Nitrergic enteric neurons	3612	23	0.3385	1	10.44	9.57	0.5114	1
TEGLU22	Excitatory neurons, amygdala	983	7	0.3413	1	5.19	4.76	-	-
SYNOR1	Noradrenergic erector muscle neurons	2637	17	0.3532	1	8.54	6.94	0.1662	1
ENT5	Cholinergic enteric neurons	2475	16	0.3558	1	10.03	10.01	0.3988	1
PSPEP5	Peptidergic (PEP1.2), DRG	2145	14	0.3568	1	6.03	11.34	-	-
TEINH12	Non-border Cck interneurons, cortex/hippocampus	1494	10	0.3626	1	2.89	4.38	-	-
TEGLU8	Excitatory neurons, cerebral cortex	1172	8	0.3666	1	4.56	4.44	-	-
TEGLU13	Excitatory neurons, cerebral cortex	1693	11	0.3889	1	3.54	4.75	-	-
ENT3	Nitrergic enteric neurons	3382	21	0.3965	1	9.61	10.07	-	-
TEGLU16	Excitatory neurons, cerebral cortex	1888	12	0.4098	1	6.37	5.30	-	-
TEGLU23	Excitatory neurons, hippocampus CA3	1261	8	0.4460	1	4.67	4.48	-	-
MBCHO1	Cholinergic neurons, midbrain red nucleus	1624	10	0.4661	1	4.27	9.49	-	-
SCINH6	Inhibitory neurons, spinal cord	1284	8	0.4665	1	5.74	10.16	-	-
HBINH8	Inhibitory neurons, hindbrain	1831	11	0.4938	1	5.74	10.45	-	-
TEGLU14	Excitatory neurons, cerebral cortex	2184	13	0.4999	1	2.90	4.91	-	-
SYCHO2	Cholinergic neurons, sympathetic	2734	16	0.5207	1	6.36	7.41	-	-
ENT9	Cholinergic enteric neurons	3653	21	0.5505	1	9.83	12.43	-	-
TEGLU11	Excitatory neurons, cerebral cortex	1038	6	0.5596	1	2.71	3.92	-	-
TEINH9	Non-border Cck interneurons, hippocampus	1038	6	0.5596	1	4.93	5.23	-	-
SYNOR2	Noradrenergic neurons, sympathetic	2625	15	0.5619	1	9.71	7.59	0.1255	1
SCINH1	Inhibitory neurons, spinal cord	1220	7	0.5640	1	4.60	11.37	-	-
SYNOR4	Noradrenergic erector muscle neurons	2659	15	0.5832	1	9.83	6.98	0.114	1
TECHO	Cholinergic interneurons, telencephalon	1793	10	0.5959	1	2.40	6.80	-	-
PSNF1	Neurofilament (NF1), DRG	2762	15	0.6453	1	6.88	8.27	-	-
TEGLU2	Excitatory neurons, cerebral cortex	1158	6	0.6670	1	2.72	4.98	-	-
PSNP4	Non-peptidergic (NP2.1), DRG	2820	15	0.6782	1	5.98	11.60	-	-
TEGLU15	Excitatory neurons, cerebral cortex	1556	8	0.6867	1	4.18	4.33	-	-
TEINH11	R-LM border Cck interneurons, cortex/hippocampus	1436	7	0.7334	1	4.96	4.30	-	-
TEGLU7	Excitatory neurons, cerebral cortex	1442	7	0.7374	1	3.45	3.93	-	-
ENT1	Nitrergic enteric neurons	2447	12	0.7713	1	4.84	10.91	-	-
PSPEP6	Peptidergic (TrpM8), DRG	1777	8	0.8179	1	5.98	8.68	-	-
ENT6	Cholinergic enteric neurons	2805	13	0.8431	1	5.48	9.36	-	-
DEGLU2	Excitatory neurons, hypothalamus	1645	7	0.8490	1	3.26	4.75	-	-
TEGLU4	Excitatory neurons, cerebral cortex	1471	6	0.8623	1	2.43	3.71	-	-
PSPEP4	Peptidergic (PEP1.1), DRG	2683	12	0.8676	1	7.68	8.55	-	-
HBGLU5	Excitatory neurons, hindbrain	2160	9	0.8902	1	4.60	5.96	-	-
PSNP2	Non-peptidergic (NP1.1), DRG	3065	13	0.9182	1	15.55	8.94	-	-
PSNP5	Non-peptidergic (NP2.2), DRG	3077	13	0.9208	1	6.51	10.15	-	-

PSPEP7	Peptidergic (TrpM8), DRG	1656	6	0.9253	1	11.50	9.76	-	-
SYCHO1	Cholinergic neurons, sympathetic	2721	11	0.9299	1	7.98	7.83	-	-
PSPEP1	Peptidergic (PEP2), DRG	2924	12	0.9302	1	8.84	10.20	-	-
PSPEP3	Peptidergic (PEP1.4), DRG	3131	13	0.9316	1	9.33	8.77	-	-
PSNP1	Non-peptidergic (Th), DRG	3137	13	0.9328	1	7.34	9.26	-	-
PSPEP8	Peptidergic (TrpM8), DRG	2537	10	0.9339	1	11.47	8.74	-	-
PSPEP2	Peptidergic (PEP1.3), DRG	3163	13	0.9375	1	8.27	8.72	-	-
PSNF3	Neurofilament (NF2/3), DRG	2767	11	0.9388	1	7.41	8.48	-	-
PSNP3	Non-peptidergic (NP1.2), DRG	3066	12	0.9539	1	29.76	9.34	-	-
SZNBL	Neuronal intermediate progenitor cells	1802	6	0.9557	1	4.82	23.22	-	-
PSNF2	Neurofilament (NF4/5), DRG	2688	10	0.9587	1	14.99	9.17	-	-
PSNP6	Non-peptidergic (NP3), DRG	3524	14	0.9606	1	11.26	16.83	-	-
DEGLU1	Excitatory neurons, thalamus	2056	6	0.9834	1	3.94	5.08	-	-
DEINH2	Inhibitory neurons, thalamus	991	5	-	-	4.61	5.04	-	-
DGNBL2	Granule neuroblasts, dentate gyrus	715	5	-	-	3.46	5.09	-	-
MEGLU8	Excitatory neurons, midbrain	531	5	-	-	2.62	5.32	-	-
MEINH7	Inhibitory neurons, midbrain	592	5	-	-	3.24	9.11	-	-
MEINH8	Inhibitory neurons, midbrain	332	5	-	-	4.15	24.00	-	-
MSN2	D2 medium spiny neurons, striatum	955	5	-	-	4.87	5.12	-	-
MSN3	D2 medium spiny neurons, striatum	615	5	-	-	5.10	4.57	-	-
MSN4	D1 medium spiny neurons, striatum	710	5	-	-	6.32	4.60	-	-
MSN5	Patch D1/D2 neurons, striatum	443	5	-	-	3.51	7.23	-	-
OBNBL1	Neuroblasts, olfactory	1055	5	-	-	8.07	11.63	-	-
SCGLU2	Excitatory neurons, spinal cord	572	5	-	-	2.88	10.67	-	-
SCINH7	Inhibitory neurons, spinal cord	701	5	-	-	10.21	13.98	-	-
SCINH9	Inhibitory neurons, spinal cord	583	5	-	-	12.71	13.05	-	-
TEGLU5	Excitatory neurons, cerebral cortex	1402	5	-	-	2.61	5.87	-	-
TEINH13	Trilaminar cells, hippocampus	1150	5	-	-	2.79	4.55	-	-
TEINH7	Interneuron-selective interneurons, hippocampus	1053	5	-	-	5.11	5.67	-	-
DEINH1	Inhibitory neurons, thalamus	813	4	-	-	4.03	5.04	-	-
DGGRC1	Granule neuroblasts, dentate gyrus	522	4	-	-	4.70	4.43	-	-
DGNBL1	Granule neuroblasts, dentate gyrus	779	4	-	-	4.04	7.68	-	-
MEGLU5	Excitatory neurons, midbrain	843	4	-	-	2.41	5.85	-	-
MEGLU7	Excitatory neurons, midbrain	646	4	-	-	2.76	6.68	-	-
MEINH5	Inhibitory neurons, midbrain	785	4	-	-	9.36	8.55	-	-
MEINH6	Inhibitory neurons, midbrain	1128	4	-	-	3.52	6.91	-	-
OBINH2	Inhibitory neurons, olfactory bulb	486	4	-	-	3.80	5.61	-	-
OBNBL3	Neuroblasts, olfactory bulb	479	4	-	-	4.01	5.36	-	-
SCGLU6	Excitatory neurons, spinal cord	609	4	-	-	3.94	12.88	-	-
TEGLU19	Excitatory neurons, cerebral cortex	1191	4	-	-	3.29	5.22	-	-
TEGLU9	Excitatory neurons, cerebral cortex	973	4	-	-	2.94	4.69	-	-
TEINH15	CGE-derived neuroglia form cells, cortex/hippocampus	694	4	-	-	6.10	5.10	-	-
TEINH16	Ivy and MGE-derived neuroglia form cells, cortex/hippocampus	788	4	-	-	2.49	4.90	-	-
TEINH18	Basket and bistratified cells, cortex/hippocampus	1103	4	-	-	4.82	4.71	-	-
TEINH21	Sleep-active, long-range projection interneurons, cortex/hippocampus	612	4	-	-	2.93	6.07	-	-

TEINH6	Interneuron-selective interneurons, cortex/hippocampus	511	4	-	-	2.84	7.24	-	-
CBNBL1	Neuroblasts, cerebellum	365	3	-	-	8.27	15.34	-	-
CR	Cajal-Retzius cells, hippocampus	395	3	-	-	20.70	17.25	-	-
DETPH	Neuroblast-like, habenula	877	3	-	-	21.54	106.51	-	-
DGGRC2	Granule neurons, dentate gyrus	659	3	-	-	4.91	5.32	-	-
MSN6	Matrix D1 neurons, striatum	181	3	-	-	4.35	7.06	-	-
OBDOP2	Inhibitory neurons, olfactory bulb	263	3	-	-	5.48	7.32	-	-
SCGLU10	Excitatory neurons, spinal cord	438	3	-	-	4.02	13.85	-	-
SEPNBL	Neuroblasts, septum	383	3	-	-	7.86	13.30	-	-
TEGLU1	Excitatory neurons, cerebral cortex	586	3	-	-	5.57	5.88	-	-
TEGLU18	Excitatory neurons, cerebral cortex	1067	3	-	-	27.77	4.84	-	-
TEGLU24	Excitatory neurons, hippocampus CA1	1053	3	-	-	7.01	4.72	-	-
TEGLU3	Excitatory neurons, cerebral cortex	783	3	-	-	4.38	4.61	-	-
TEGLU6	Excitatory neurons, cerebral cortex	464	3	-	-	2.84	5.12	-	-
TEINH20	Inhibitory interneurons, hippocampus	1074	3	-	-	2.33	5.28	-	-
MEINH11	Inhibitory neurons, midbrain	503	2	-	-	3.68	8.34	-	-
OBINH3	Inhibitory neurons, olfactory bulb	314	2	-	-	5.82	5.53	-	-
OBINH5	External plexiform layer interneuron, olfactory bulb	517	2	-	-	47.62	13.20	-	-
OBNBL2	Neuroblasts, olfactory bulb	724	2	-	-	9.68	12.57	-	-
SCGLU7	Excitatory neurons, spinal cord	337	2	-	-	7.15	23.84	-	-
SCGLU8	Excitatory neurons, spinal cord	577	2	-	-	6.01	20.64	-	-
SCGLU9	Excitatory neurons, spinal cord	483	2	-	-	18.86	12.43	-	-
SCINH2	Inhibitory neurons, spinal cord	645	2	-	-	3.03	8.29	-	-
SCINH4	Inhibitory neurons, spinal cord	526	2	-	-	2.15	12.87	-	-
TEGLU17	Excitatory neurons, cerebral cortex	1043	2	-	-	6.49	3.75	-	-
TEINH17	Axo-axonic, cortex/hippocampus	760	2	-	-	4.47	4.85	-	-
TEINH4	Interneuron-selective interneurons, cortex/hippocampus	432	2	-	-	2.79	5.82	-	-
CBNBL2	Neuroblasts, cerebellum	148	1	-	-	6.52	12.86	-	-
CBPC	Purkinje cells	511	1	-	-	2.55	19.71	-	-
MEINH9	Inhibitory neurons, midbrain	689	1	-	-	2.69	6.51	-	-
OBINH1	Inner horizontal cell, olfactory bulb	242	1	-	-	5.02	5.67	-	-
OBNBL4	Inhibitory neurons, olfactory bulb	252	1	-	-	26.70	14.62	-	-
OBNBL5	Inhibitory neurons, olfactory bulb	313	1	-	-	3.39	8.16	-	-
TEINH14	CGE-derived neuroglia form cells Cxcl14+, cortex/hippocampus	462	1	-	-	4.67	5.33	-	-
CBGRC	Granule neurons, cerebellum	186	0	-	-	0.00	17.88	-	-
CBINH1	Molecular layer interneurons, cerebellum	413	0	-	-	0.00	10.46	-	-
MEINH1	Inhibitory neurons, midbrain	199	0	-	-	0.00	19.00	-	-
OBINH4	Inhibitory neurons, olfactory bulb	204	0	-	-	0.00	13.29	-	-

Appendix Table A3.9 (A & B) - Hypothalamus Enrichment - Chen et al. (2017)

Appendix Table A3.9A. Whole Hypothalamus Level Enrichment Analysis for all imprinted genes in cell lineage populations in the Chen et al. (2017) dataset. *Up Reg* – number of upregulated genes with $q \leq 0.05$ and Log2FC > 0 (total number of genes in the dataset in brackets); *IG* – number of imprinted genes upregulated with $q \leq 0.05$ and Log2FC > 0 (total number of IGs in the dataset in brackets); *ORA p – p* value from over representation analysis on groups with minimum 5% of total IGs; *ORA q* – Bonferroni corrected *p* value from ORA; *Mean FC IG* – mean fold change for upregulated imprinted genes; *Mean FC Rest* – mean fold change for all other upregulated genes; *GSEA p – p* value from Gene Set Enrichment Analysis for identity groups with 15+ IGs and Mean FC IG > Mean FC Rest; *GSEA q* – Bonferroni corrected *p* values from GSEA.

Hypothalamic Cell Identity	Up Reg (14,687)	IG (91)	ORA *p*	ORA *q*	Mean FC IG	Mean FC Rest	GSEA *p*
Neuron	5787	54	9.06E-05	**0.0010**	4.81	3.45	**0.0221**
Oligodendrocyte Progenitor Cell	1602	13	0.1893	1	2.82	3.02	-
Myelinating Oligodendrocyte	1830	14	0.2391	1	2.65	3.9	-
Immature Oligodendrocyte	1577	11	0.3864	1	2.49	3.15	-
Astrocyte	2975	19	0.4827	1	1.87	3.75	-
Epithelial cell 2	1056	7	0.4836	1	6.98	7.34	-
Epithelial cell 1	2249	14	0.5373	1	3.96	6.67	-
SCO	1288	7	0.6962	1	6.48	8.64	-
Macrophage	1834	10	0.7144	1	5.3	10.96	-
Tanycyte	2938	12	0.9665	1	4.49	3.55	-
Ependymocyte	3621	15	0.9776	1	3.41	11.19	-
Microglia	625	3	-	-	5.24	10.72	-
Proliferating Oligodendrocyte Progenitor Cell	1264	1	-	-	1.65	17.41	-

Appendix Table A3.9B. Whole Hypothalamus Level Enrichment Analysis for all imprinted genes in neuronal subpopulations in the Chen et al. (2017) dataset. All column descriptions can be found in the legend of Appendix Table A3.9A. Marker Genes – Genes identified by Chen et al. (2017) with distinct expression in these neuronal subtypes for the purposes of identification.

Hypothalamic Neuron Identity	Marker Genes	Up Reg (12,238)	IG (80)	ORA *p*	ORA *q*	Mean FC IG	Mean FC Rest
GABA17	*Slc6a3*	825	18	4.57E-06	**1.19E-04**	2.73	3.30
GABA8	*Vipr2*	480	11	0.0003	**0.0071**	1.96	2.73
GABA13	*Slc18a2, Gal*	569	12	0.0003	**0.0079**	1.89	2.02
GABA15	*Agrp*	766	13	0.0013	**0.0339**	2.40	2.79
GABA16	*Cox6a2*	258	6	0.0068	0.1762	1.68	2.70
GABA14	*Cbln4*	271	6	0.0085	0.2220	1.68	2.54
GABA2	*Npas1*	278	6	0.0096	0.2499	2.29	4.22
GABA18	*Lhx1, Klhl1*	200	5	0.0099	0.2576	1.55	3.06
GABA9	*Vip*	219	5	0.0142	0.3703	1.39	3.77
GABA10	*Prok2*	605	9	0.0169	0.4391	3.03	4.28
Glu9	*Gng8, Samd3*	345	6	0.0252	0.6556	1.34	3.66
Glu1	*Crh*	182	4	0.0312	0.8120	1.68	4.38
GABA11	*Ghrh*	392	6	0.0430	1	6.50	2.91
GABA5	*Lhx8*	569	7	0.0780	1	1.28	3.09
Glu14	*Avp, Sim1*	945	10	0.0875	1	2.07	3.30
Glu15	*Sst, Prdm8*	650	7	0.1319	1	1.69	3.79
Glu12	*Vgll2*	339	4	0.1811	1	1.70	3.88
Glu13	*Pomc*	471	5	0.1945	1	3.36	4.04
GABA12	*Crabp1*	354	4	0.2010	1	1.63	2.61
Glu6	*Tac1*	607	6	0.2052	1	2.39	3.58
Glu7	*Fezf1, Lbhd2*	609	6	0.2072	1	1.44	2.13
Glu11	*Kiss1*	468	4	0.3663	1	7.78	3.68
GABA3	*Bcl11b*	484	4	0.3902	1	1.41	3.36
Glu8	*Lbhd2, Cartpt*	534	4	0.4641	1	3.34	2.97
Glu4	*Shox2*	1690	8	0.8792	1	2.25	3.09
Glu5	*Foxb1*	984	4	0.8942	1	2.20	3.69
GABA6	*Pax6*	228	3	-	-	1.03	3.86
GABA7	*Trh*	194	3	-	-	3.87	3.46
Glu3	*Fezf2, Samd3*	203	3	-	-	1.07	5.35
GABA1	*Pvalb*	738	2	-	-	3.81	4.57
GABA4	*Gm13498*	254	2	-	-	3.97	4.43
Glu10	*Trh*	228	2	-	-	1.17	4.88
Hista	*Hdc*	585	2	-	-	3.98	5.14
Glu2	*Sln*	376	0	-	-	0.00	4.38

Appendix Table A3.10 (A & B) – Hypothalamus Enrichment - Romanov et al. (2017)

Appendix Table A3.10A. Whole Hypothalamus Level Enrichment Analysis for all imprinted genes in cell lineage populations in the Romanov et al. (2017) dataset. All column descriptions can be found in the legend of Appendix Table A3.9A

Hypothalamic Cell Identity	Up Reg (14,914)	IG (94)	ORA *p*	ORA *q*	Mean FC IG	Mean FC Rest	GSEA *p*
neurons	8806	68	0.0050	**0.0202**	2.53	2.07	**0.0105**
endothelial	2077	14	0.4377	1	9.84	7.62	-
ependymal	2232	13	0.6656	1	4.13	5.91	-
oligos	3047	16	0.8288	1	2.05	2.16	-
astrocytes	769	4	-	-	1.95	4.67	-
vascular smooth muscle	632	4	-	-	8.18	7.85	-
microglia	692	1	-	-	9.84	27.66	-

Appendix Table A3.10B. Whole Hypothalamus Level Enrichment Analysis for all imprinted genes in neuronal subpopulations in the Romanov et al. (2017) dataset. All column descriptions can be found in the legend of Appendix Table A3.9A. Marker Genes – Genes identified by Romanov et al. (2017) with distinct expression in these neuronal subtypes for the purposes of identification.

Hypothalamic Neuron Identity	Marker Genes	Up Reg (12,243)	IG (86)	ORA *p*	ORA *q*	Mean FC IG	Mean FC Rest
GABA 5	*Calcr, Lhx1*	40	6	2.47E-07	**1.73E-06**	3.47	4.95
GABA 14	*Npy, Agrp*	193	6	0.0019	**0.0132**	2.94	6.12
Ghrh	*Ghrh*	234	6	0.0049	**0.0342**	18.31	17.22
Dopamine 4	*Th, Slc6a3, Slc18a2*	272	6	0.0100	0.0697	5.62	5.64
Avp 2, high	*Avp*	217	5	0.0154	0.1078	3.74	5.26
GABA 11	*Nts*	319	5	0.0636	0.4450	2.25	2.82
Vglut2 11	*Slc17a6*	1840	12	0.5924	1	3.25	2.97
GABA 12	*Nts*	162	4	-	-	1.78	3.08
Qrfp	*Qrfp*	257	4	-	-	10.77	7.58
Dopamine 1	*Th, Slc18a2*	217	3	-	-	5.46	4.26
Dopamine 2 (low *Slc18a2*)	*Th*	135	3	-	-	6.97	6.47
Dopamine 3	*Th, Slc18a2*	335	3	-	-	7.79	4.26
GABA 2	*Gucy1a3*	198	3	-	-	2.34	3.70
Hcrt	*Hcrt*	129	3	-	-	5.46	6.30
Hmit+/-	*Slc2a13*	16	3	-	-	1.88	5.24
Vglut2 15	*Hcn1, 6430411K18Rik*	213	3	-	-	8.84	8.05
circadian 2	*Nms, Vip+/-*	53	2	-	-	2.53	8.31
GABA 4	*Crh, Pgr15l*	21	2	-	-	1.35	3.55
Oxytocin 1	*Oxt*	331	2	-	-	3.02	4.10
Oxytocin 4	*Oxt*	245	2	-	-	2.93	4.81
Pmch	*Pmch*	239	2	-	-	2.51	6.04
Trh 3 (high)	*Trh, Cartpt*	178	2	-	-	3.83	6.38
Vglut2 1	*Penk*	294	2	-	-	9.13	15.24
Vglut2 10	*Morn4,Prrc2a*	252	2	-	-	9.92	8.29
Vglut2 17	*A930013F10Rik, Pou2f2*	122	2	-	-	11.42	10.93
Adcyap1 2	*Adcyap1*	236	1	-	-	1.81	3.21
Avp 1, high	*Avp*	141	1	-	-	3.10	5.71
Avp 3, medium	*Avp*	105	1	-	-	3.49	6.29
circadian 3	*Per2*	19	1	-	-	9.08	6.08
GABA 10	-	81	1	-	-	1.17	3.45
GABA 13	*Gal*	126	1	-	-	4.40	3.61
GABA 3	*Crh+/-, Lhx6*	144	1	-	-	1.66	3.66
GABA 9	-	141	1	-	-	1.72	3.23
GABA1	-	135	1	-	-	6.25	4.07
Sst 1, low	*Sst*	63	1	-	-	7.74	4.49
Sst 3, medium	*Sst*	202	1	-	-	6.36	5.05
Vglut2 12	*Mgat4b*	199	1	-	-	7.09	6.34
Vglut2 13	*Ninl, Rfx5, Zfp346*	149	1	-	-	6.90	6.59
Vglut2 16	*Zfp975, Tnr*	371	1	-	-	9.30	11.02
Vglut2 18	*Zfp458, Ppp1r12b,*	191	1	-	-	27.13	22.64
Vglut2 2	*Crh+/-*	117	1	-	-	1.89	4.32
Vglut2 5	*Myt1, Lhx9*	109	1	-	-	7.33	12.30
Vglut2 6	*Prmt8, Ugdh*	46	1	-	-	5.98	5.29

Vglut2 7	Pgam, Snx12	93	1	-	-	7.17	10.26
Vglut2 8	-	47	1	-	-	2.49	5.01
Adcyap1 1	*Adcyap1, Tac1*	183	0	-	-	0.00	3.91
circadian 1	*Vip, Grp+/-*	30	0	-	-	0.00	12.44
GABA 15	*Npy*	67	0	-	-	0.00	10.03
GABA 6 (Otof, Lhx1)	*Otof, Lhx1*	33	0	-	-	0.00	3.54
GABA 7	*Pomc+/-*	16	0	-	-	0.00	5.59
GABA 8	-	74	0	-	-	0.00	4.10
Gad-low, *Gnrh-/+*	*Gnrh-/+*	44	0	-	-	0.00	9.53
Npvf	*Npvf*	147	0	-	-	0.00	6.95
Oxytocin 2	*Oxt*	254	0	-	-	0.00	5.14
Oxytocin 3	*Oxt*	166	0	-	-	0.00	5.37
Sst 2, high	*Sst*	215	0	-	-	0.00	5.16
Trh 1 (low)	*Trh*	31	0	-	-	0.00	5.35
Trh 2 (medium)	*Trh*	48	0	-	-	0.00	4.80
Vglut2 14	*Col9a2*	101	0	-	-	0.00	8.21
Vglut2 3	*Crh-+/-, low*	1	0	-	-	0.00	5.19
Vglut2 4	-	218	0	-	-	0.00	7.70
Vglut2 9	*Gpr149*	61	0	-	-	0.00	8.77

Appendix Table A3.11 – Marker Genes of GABA5 and GABA17

Appendix Table A3.11. Comparison of Upregulated Genes for GABA5 in Romanov et al. (2017) and GABA17 in Chen et al. (2017) . 21 genes were upregulated in both subgroups which is over 50% of the upregulated genes found in GABA5 including the top 5 marker genes. Position of genes organised by *p* value is included for both subtypes alongside the *p* values, Bonferroni corrected *q* values, and the fold change (FC) values for expression in this neuronal subpopulation vs. background. *Prlr* did not have a significant q-value in GABA5 but is included due to its significance to TIDA neurons.

Gene	Position in GABA5 sorted by *p* (38 upreg genes)	GABA5 *p*	GABA5 *q*	GABA5 FC	Position in GABA17 sorted by *p* (825 upreg genes)	GABA17 *p*	GABA17 *q*	GABA17 FC
Calcr	1	2.09E-38	1.30E-36	8.13	41	4.12E-13	4.67E-12	4.96
Lhx1	2	1.82E-10	1.13E-08	9.56	29	1.35E-16	1.53E-15	4.74
Slc6a3	3	6.44E-09	1.33E-07	15.71	1	2.20E-111	7.47E-110	69.29
Asb4	4	6.76E-08	4.19E-06	2.82	445	0.000408	0.002311	1.99
Lhx1os	5	1.13E-07	3.50E-06	4.74	23	1.25E-19	1.42E-18	4.61
Npbwr1	7	8.22E-06	0.000510	9.02	819	0.008319	0.047138	3.55
Diap3	9	4.79E-05	0.002967	7.51	12	1.73E-27	5.87E-26	9.97
Gpr83	11	0.000118	0.003456	4.27	14	3.49E-25	1.19E-23	6.27
Slc18a2	13	0.000158	0.001961	3.24	3	9.52E-44	1.62E-42	9.43
Peg10	16	0.000240	0.004955	2.25	40	9.95E-14	3.38E-12	2.46
Dlk1	20	0.000324	0.009853	1.89	124	1.24E-07	1.05E-06	1.93
Fam159b	21	0.000371	0.007659	4.98	72	5.46E-10	6.19E-09	4.19
Rab3b	24	0.000462	0.009550	1.85	30	1.39E-16	4.72E-15	1.96
Six6	25	0.000526	0.008157	4.35	11	1.01E-27	3.45E-26	6.41
AA388235	30	0.000863	0.036304	3.79	83	2.36E-09	8.01E-08	10.61
Hnrnpul1	33	0.001399	0.043371	3.87	602	0.001031	0.035041	2.20
Arhgap36	36	0.001821	0.028222	2.19	46	2.61E-12	4.44E-11	3.50
Th	37	0.002050	0.014126	2.24	8	1.45E-35	2.46E-34	8.29
Dlx2	38	0.002801	0.034730	4.09	139	4.42E-07	3.01E-06	3.23
Dlx6	39	0.003047	0.047231	9.07	193	5.23E-06	5.93E-05	3.53
Dlx1	40	0.006400	0.036074	2.07	34	1.23E-15	2.08E-14	3.60
Prlr	66	0.011087	0.114386	1.86	13	3.21E-26	1.09E-24	4.36

Appendix A3.12 – Arcuate Nucleus Enrichment (Campbell et al., 2017)

Appendix Table A3.12. Specific Hypothalamic Nuclei Level Enrichment Analysis for all imprinted genes in neuronal subpopulation in the Arcuate Nucleus dataset (Campbell et al., 2017). All column descriptions can be found in the legend of Appendix Table A3.9A.

ARC Neuron Identity	Up Reg (16,098)	IG (95)	ORA *p*	ORA *q*	Mean FC IG	Mean FC Rest	GSEA *p*	GSEA *q*
Th/Slc6a3	1113	26	7.49E-10	**1.72E-08**	2.14	4.04	-	-
Arx/Nr5a2	599	16	3.98E-07	**9.16E-06**	1.33	3.15	-	-
Agrp/Sst	1704	23	0.0001	**0.0026**	1.61	1.78	-	-
Pomc/Anxa2	3162	34	0.0002	**0.0039**	1.70	1.82	-	-
Ghrh	1265	18	0.0004	**0.0088**	2.02	2.82	-	-
Gpr50	578	11	0.0006	**0.0136**	2.52	5.02	-	-
Pomc/Glipr1	1536	19	0.0014	**0.0328**	1.91	1.90	0.2851	0.8553
Slc17a6/Trhr	1547	19	0.0016	**0.0357**	3.22	2.35	0.1143	0.3429
Th/Cxcl12	818	12	0.0031	0.0710	1.95	3.40	-	-
Agrp/Gm8773	3629	33	0.0044	0.1021	2.08	1.84	0.1032	0.3096
Fam19a2	1237	15	0.0057	0.1321	1.62	3.42	-	-
Th/Lef1	600	9	0.0089	0.2052	1.54	3.98	-	-
Tbx19	584	8	0.0220	0.5063	1.36	3.71	-	-
Pomc/Ttr	1114	11	0.0634	1	1.92	2.21	-	-
Gm8773/Tac1	1628	14	0.0963	1	1.51	3.14	-	-
Kiss1/Tac2	1087	9	0.1912	1	1.55	3.86	-	-
Nfix/Htr2c	1719	13	0.2112	1	2.93	4.36	-	-
Sst/Unc13c	832	7	0.2200	1	1.39	2.35	-	-
Th/Sst	1605	12	0.2357	1	3.00	7.62	-	-
Qrfp	751	6	0.2828	1	4.58	14.58	-	-
Htr3b	1514	10	0.4034	1	1.29	2.17	-	-
Sst/Nts	3861	23	0.5191	1	1.62	2.31	-	-
Sst/Pthlh	1609	7	0.8493	1	1.81	2.38	-	-
Tmem215	498	2	-	-	1.08	2.18	-	-

Appendix Table A3.13 (A, B & C) – POA Enrichment (Moffitt et al., 2018)

Appendix Table A3.13A. Specific Hypothalamic Nuclei Level Enrichment Analysis for all imprinted genes in neuronal subpopulation in the Preoptic Area (Moffitt et al., 2018). Suspected non-POA neuronal populations included (text coloured red). All column descriptions can be found in the legend of Appendix Table A3.9A.

POA Neuron Identity	Up Reg (16,402)	IG (101)	ORA *p*	ORA *q*	Mean FC IG	Mean FC Rest	GSEA *p*	GSEA *q*
i22:*Gal/Pmaip1*	1490	28	4.16E-08	**2.54E-06**	1.99	2.03	-	-
i17:*Th/Nos1*	1402	23	9.05E-06	**0.0006**	1.86	1.95	-	-
i8:*Gal/Amigo2*	3700	42	1.08E-05	**0.0007**	1.77	1.49	0.1096	1
e15:*Ucn3/Brs3*	1039	18	5.06E-05	**0.0031**	5.07	3.41	0.0473	0.5203
i12:GABA	1266	20	6.66E-05	**0.0041**	1.78	1.65	0.1911	1
e19:*Ghrh/Trh*	1981	26	0.0001	**0.0063**	2.07	3.8	-	-
i6:*Avp/Nms*	1773	24	0.0001	**0.0078**	2.19	3.03	-	-
h1:GABA/*Slc17a6*	1259	19	0.0002	**0.0116**	1.91	3.49	-	-
i24:*Nmu*	3092	34	0.0002	**0.0139**	1.68	2	-	-
e13:*Ghrh/C1ql1*	3639	38	0.0002	**0.0147**	2.49	2.79	-	-
e4:*Trh/Angpt1*	2844	32	0.0002	**0.0152**	1.85	2.24	-	-
e20:*Crh*	2466	29	0.0003	**0.0158**	2.36	3.13	-	-
i37:*Bdnf/Chrm2*	2119	25	0.0008	**0.0458**	2.7	3.15	-	-
i1:GABA	673	12	0.0008	**0.0486**	1.27	1.47	-	-
e1:Glut	1536	20	0.0009	0.0524	1.6	1.85	-	-
h3:*Slc32a1/Gsc*	2402	27	0.0009	0.0564	1.79	2.44	-	-
i35:*Crh/Tac2*	1547	20	0.0009	0.0573	1.93	2.36	-	-
i23:*Crh/Nts*	1553	20	0.0010	0.0601	1.62	1.92	-	-
e17:*Th/Adcyap1*	4517	42	0.0013	0.0763	2.71	2.92	-	-
i38:*Kiss1/Th*	2074	24	0.0013	0.0787	2.1	5.3	-	-
i42:*Pthlh*	653	11	0.0021	0.1270	2.59	5.42	-	-
e22:*Gal/Ucn3*	1907	22	0.0022	0.1361	2.65	6.26	-	-
e7:*Reln/C1ql1*	1678	20	0.0025	0.1529	1.67	2.68	-	-
i16:*Gal/Th*	5447	47	0.0028	0.1718	2.39	1.83	0.0226	0.2486
e2:*Tac1/Fezf1*	4042	37	0.0039	0.2394	1.61	1.76	-	-
i15:GABA	3478	33	0.0040	0.2457	1.47	1.77	-	-
e12:*Nos1/Foxp2*	1762	20	0.0044	0.2682	1.85	2.53	-	-
i9:GABA	1639	19	0.0044	0.2699	1.31	1.67	-	-

i31:*Calca*	427	8	0.0046	0.2785	1.41	3.21	-	-
i14:*Avp/Cck*	1303	16	0.0053	0.3254	2.74	2.7	0.5429	1
i28:GABA/*Six6*	2849	28	0.0055	0.3336	2.25	2.23	0.4027	1
e21:Glut/*Rxfp3*	649	10	0.0062	0.3778	1.94	4.58	-	-
h2:*Nts/Slc17a8*	1270	15	0.0099	0.6012	1.42	1.94	-	-
e6:*Nos1/Trp73*	1324	15	0.0141	0.8580	2.48	2.67	-	-
e3:*Cartpt/Isl1*	4658	39	0.0141	0.8594	3	2.01	0.1428	1
i20:*Gal/Moxd1*	2002	20	0.0171	1	1.42	2.14	-	-
i41:*Npy/Penk*	652	9	0.0180	1	4.07	8.28	-	-
i25:*Npy/Etv1*	912	11	0.0230	1	2.55	3.39	-	-
i44:*Th/Cxcl14*	795	10	0.0230	1	6.42	5.24	-	-
i39:GABA	594	8	0.0286	1	1.55	4.56	-	-
i13:GABA	2248	21	0.0287	1	1.7	1.74	-	-
e10:*Glut/Meis2*	2153	20	0.0345	1	1.94	2.39	-	-
e16:*Sst/Cartpt*	3906	32	0.0383	1	2.39	2.89	-	-
i29:GABA/*Igsf1*	1774	17	0.0393	1	1.48	2.28	-	-
i27:*Th/Trh*	1512	15	0.0403	1	3.71	3.11	0.0339	0.3729
i5:GABA/*Pou3f3*	3249	27	0.0499	1	1.41	1.85	-	-
e5:*Adcyap1/Nkx2-1*	2857	23	0.0926	1	1.89	1.94	-	-
i11:GABA	781	8	0.1039	1	1.91	1.94	-	-
i32:*Sst/Npy*	660	7	0.1082	1	1.76	4.71	-	-
i7:GABA	3387	26	0.1164	1	1.54	1.51	0.2443	1
i18:*Gal/Tac2*	4682	34	0.1361	1	1.6	1.69	-	-
i30:*Vip*	976	9	0.1406	1	3.21	4.85	-	-
e23:*Reln/Etv1*	1869	15	0.1626	1	2.67	5.44	-	-
e11:*Glut/Shox2*	2018	16	0.1636	1	2.22	3.44	-	-
i2:Tac1/Pdyn	1988	15	0.2268	1	2.42	2.22	0.0757	0.8327
e9:*Glut/Tcf7l2*	1352	10	0.3093	1	1.66	3.02	-	-
i3:*Penk/Nts*	1621	11	0.4013	1	1.88	2.5	-	-
i21:*Sst/Pou3f3*	2628	17	0.4362	1	2.25	2.14	0.2163	1
i43:*Chat*	2069	13	0.4988	1	3.58	9.07	-	-
e24:*Gal/Rxfp1*	913	5	0.6603	1	22.59	8.39	-	-
i10:*Tac1/Nts*	2013	9	0.8791	1	2.7	2.19	-	-
i40:*Sst/Reln*	383	4	-	-	5.3	5.59	-	-
i4:GABA/*Mylk*	613	3	-	-	2.24	2.67	-	-
i26:*Tac1/Prok2*	185	1	-	-	6.49	3.59	-	-
e8:*Cck/Ebf3*	1008	1	-	-	4.73	6.07	-	-
i45:*Bdnf/Pmaip1*	202	1	-	-	25.51	11.19	-	-

Appendix Table A3.13B. Specific Hypothalamic Nuclei Level Enrichment Analysis for all imprinted genes in neuronal subpopulation in the Preoptic Area (Moffitt et al., 2018) with extra-neuronal populations excluded and a Log2FC > 0 criteria used. All column descriptions can be found in the legend of Appendix Table A3.9A.

POA Neuron Identity	Up Reg (16,402)	IG (100)	ORA *p*	ORA *q*	Mean FC IG	Mean FC Rest	GSEA *p*	GSEA *q*
i22:*Gal/Pmaip1*	1569	27	4.64E-07	**2.41E-05**	2.02	2.14	-	-
i8:*Gal/Amigo2*	3625	42	6.29E-06	**0.0003**	1.78	1.53	0.1628	1
e15:*Ucn3/Brs3*	996	18	2.90E-05	**0.0015**	6.15	3.43	0.04	0.32
i12:GABA	1229	20	4.40E-05	**0.0023**	1.76	1.68	0.2236	1
i35:*Crh/Tac2*	1464	22	5.69E-05	**0.0030**	1.98	2.44	-	-
i17:*Th/Nos1*	1358	21	5.76E-05	**0.0030**	2.02	2	0.4519	1
e4:*Trh/Angpt1*	2770	33	6.13E-05	**0.0032**	1.9	2.28	-	-
e19:*Ghrh/Trh*	1974	26	9.75E-05	**0.0051**	2.14	3.88	-	-
e1:Glut	1518	22	9.76E-05	**0.0051**	1.58	1.87	-	-
e12:*Nos1/Foxp2*	1635	22	0.0003	**0.0148**	1.81	2.64	-	-
i1:GABA	602	12	0.0003	**0.0154**	1.2	1.55	-	-
i24:*Nmu*	3020	33	0.0003	**0.0173**	1.71	2.02	-	-
e13:*Ghrh/C1ql1*	3588	37	0.0004	**0.0207**	2.58	2.81	-	-
i37:*Bdnf/Chrm2*	2078	25	0.0006	**0.0292**	2.8	3.17	-	-
i38:*Kiss1/Th*	2078	25	0.0006	**0.0292**	2.09	5.11	-	-
e2:*Tac1/Fezf1*	3900	38	0.0010	0.0526	1.62	1.78	-	-
i16:*Gal/Th*	5313	47	0.0016	0.0830	2.51	1.87	0.0259	0.2072
i25:*Npy/Etv1*	839	13	0.0017	0.0898	2.26	3.67	-	-
e7:*Reln/C1ql1*	1652	20	0.0021	0.1084	1.78	2.79	-	-
h1:GABA/*Slc17a6*	1192	16	0.0022	0.1157	1.97	3.71	-	-
i23:*Crh/Nts*	1437	18	0.0025	0.1291	1.73	1.94	-	-
i20:*Gal/Moxd1*	1938	22	0.0027	0.1421	1.72	2.22	-	-
e17:*Th/Adcyap1*	4418	40	0.0030	0.1568	2.79	2.97	-	-
i42:*Pthlh*	601	10	0.0036	0.1889	2.8	5.7	-	-

e22:*Gal/Ucn3*	1870	21	0.0039	0.2049	2.82	6.25	-	-
i9:GABA	1640	19	0.0045	0.2316	1.28	1.75	-	-
e3:*Cartpt/Isl1*	4568	40	0.0057	0.2950	2.89	2.05	0.1545	1
i11:GABA	657	10	0.0067	0.3501	1.68	2.01	-	-
h2:*Nts/Slc17a8*	1104	14	0.0070	0.3632	1.43	2.02	-	-
i13:GABA	2126	22	0.0083	0.4305	1.65	1.74	-	-
i15:GABA	3368	31	0.0088	0.4550	1.49	1.77	-	-
i29:*GABA/Igsf1*	1701	18	0.0139	0.7244	1.45	2.33	-	-
e6:*Nos1/Trp73*	1219	14	0.0158	0.8211	2.47	2.78	-	-
i44:*Th/Cxcl14*	760	10	0.0174	0.9063	6.33	5.29	-	-
e5:*Adcyap1/Nkx2-1*	2783	25	0.0261	1	1.81	1.94	-	-
e16:*Sst/Cartpt*	3867	32	0.0337	1	2.46	3.04	-	-
i41:*Npy/Penk*	623	8	0.0364	1	4.21	9.52	-	-
e10:Glut/*Meis2*	2059	19	0.0415	1	1.96	2.39	-	-
i7:GABA	3158	26	0.0598	1	1.51	1.51	-	-
i18:*Gal/Tac2*	4513	35	0.0608	1	1.59	1.71	-	-
i39:GABA	575	7	0.0617	1	1.77	4.77	-	-
i32:*Sst/Npy*	635	7	0.0930	1	1.84	4.76	-	-
i2:*Tac1/Pdyn*	1858	16	0.0973	1	2.26	2.18	0.0985	0.788
i5:GABA/*Pou3f3*	3073	24	0.1123	1	1.42	1.84	-	-
e11:Glut/*Shox2*	1905	16	0.1146	1	2.19	3.59	-	-
e23:*Reln/Etv1*	1851	15	0.1538	1	2.82	5.46	-	-
e24:*Gal/Rxfp1*	924	8	0.2016	1	14.75	8.62	-	-
i3:*Penk/Nts*	1547	12	0.2318	1	1.79	2.44	-	-
i21:*Sst/Pou3f3*	2490	18	0.2525	1	2.17	2.12	0.2095	1
e9:*Glut/Tcf7l2*	1255	8	0.5015	1	1.81	3.12	-	-
i43: *Chat*	2039	12	0.5978	1	4.16	8.75	-	-
i10:*Tac1/Nts*	1877	7	0.9488	1	2.87	2.18	-	-
i40:*Sst/Reln*	363	4	-	-	6.16	6.04	-	-
i4:GABA/*Mylk*	574	2	-	-	2.95	2.69	-	-
i26:*Tac1/Prok2*	180	1	-	-	7.04	3.73	-	-
e8:*Cck/Ebf3*	1009	1	-	-	5.84	6.19	-	-
i45:*Bdnf/Pmaip1*	194	1	-	-	23.06	11.47	-	-

Appendix Table A3.13C. Specific Hypothalamic Nuclei Level Enrichment Analysis for all imprinted genes in neuronal subpopulation in the Preoptic Area (Moffitt et al., 2018) with extra-neuronal populations excluded and a Log2FC > 1 criteria used. All column descriptions can be found in the legend of Appendix Table A3.9A

POA Neuron Identity	Up Reg (16,402)	IG (100)	ORA *p*	ORA *q*	Mean FC IG	Mean FC Rest	GSEA *p*
i16:*Gal/Th*	825	15	0.000137	**0.002594**	4.89	4.16	0.416
i35:*Crh/Tac2*	393	9	0.000662	**0.012572**	2.84	5.20	-
i8:*Gal/Amigo2*	353	7	0.005834	0.110848	3.81	3.67	-
e13:*Ghrh/C1ql1*	1088	14	0.006165	0.117129	4.28	5.79	-
e4:*Trh/Angpt1*	606	9	0.011653	0.221404	3.23	5.41	-
e15:*Ucn3/Brs3*	433	7	0.016667	0.316681	13.73	5.98	-
e17:*Th/Adcyap1*	1659	17	0.02239	0.425404	4.39	5.21	-
i23:*Crh/Nts*	270	5	0.024755	0.470349	2.66	4.22	-
e22:*Gal/Ucn3*	1033	11	0.049722	0.944712	3.89	10.00	-
i44:*Th/Cxcl14*	557	7	0.053803	1	8.31	6.63	-
i37:*Bdnf/Chrm2*	806	9	0.057269	1	5.34	5.74	-
i22:*Gal/Pmaip1*	417	5	0.111348	1	4.76	4.26	-
e3:*Cartpt/Isl1*	799	8	0.11433	1	8.73	4.94	-
e19:*Ghrh/Trh*	937	9	0.117593	1	3.23	6.39	-
e2:*Tac1/Fezf1*	569	6	0.134146	1	2.77	4.13	-
i24:*Nmu*	610	6	0.168524	1	2.90	4.45	-
e10:Glut/*Meis2*	511	5	0.201534	1	3.94	5.28	-
e16:*Sst/Cartpt*	1518	12	0.212905	1	3.96	5.29	-
e23:*Reln/Etv1*	1054	8	0.313293	1	3.93	8.37	-
e1:Glut	283	3	-	-	3.39	4.22	-
e11:Glut/*Shox2*	676	4	-	-	4.44	7.37	-
e12:*Nos1/Foxp2*	482	4	-	-	3.77	5.42	-
e24:*Gal/Rxfp1*	730	4	-	-	27.99	10.43	-
e5:*Adcyap1/Nkx2-1*	552	2	-	-	6.55	4.10	-
e6:*Nos1/Trp73*	368	1	-	-	17.60	5.84	-
e7:*Reln/C1ql1*	601	4	-	-	3.40	5.18	-
e8:*Cck/Ebf3*	526	1	-	-	5.84	10.53	-
e9:Glut/*Tcf7l2*	459	1	-	-	4.58	5.86	-
h1:GABA/*Slc17a6*	563	3	-	-	3.30	5.99	-

h2:*Nts/Slc17a8*	237	1	-	-	2.76	4.37	-
i1:GABA	74	0	-	-	0.00	3.38	-
i10:*Tac1/Nts*	662	3	-	-	4.73	3.48	-
i11:GABA	172	1	-	-	5.43	3.69	-
i12:GABA	177	2	-	-	6.02	4.01	-
i13:GABA	322	2	-	-	4.93	3.64	-
i15:GABA	412	2	-	-	3.26	4.97	-
i17:*Th/Nos1*	290	4	-	-	4.71	4.41	-
i18:*Gal/Tac2*	516	4	-	-	2.51	3.75	-
i2:*Tac1/Pdyn*	620	4	-	-	4.65	3.63	-
i20:*Gal/Moxd1*	444	4	-	-	3.47	5.26	-
i21:*Sst/Pou3f3*	602	4	-	-	4.97	4.21	-
i25:*Npy/Etv1*	396	4	-	-	4.19	6.07	-
i26:*Tac1/Prok2*	110	1	-	-	7.04	5.15	-
i29:GABA/*Igsf1*	418	2	-	-	2.52	5.21	-
i3:*Penk/Nts*	591	3	-	-	2.95	3.99	-
i32:*Sst/Npy*	350	2	-	-	3.14	7.38	-
i38: *Kiss1/Th*	636	3	-	-	6.45	13.13	-
i39:GABA	331	2	-	-	2.74	7.13	-
i4:GABA/*Mylk*	232	2	-	-	2.95	4.53	-
i40:*Sst/Reln*	270	3	-	-	7.75	7.55	-
i41:*Npy/Penk*	376	4	-	-	7.15	14.72	-
i42:*Pthlh*	375	3	-	-	5.75	8.11	-
i43:*Chat*	1174	3	-	-	11.91	13.93	-
i45:*Bdnf/Pmaip1*	182	1	-	-	23.06	12.12	-
i5:GABA/*Pou3f3*	525	1	-	-	3.23	3.93	-
i7:GABA	261	2	-	-	4.52	4.42	-
i9:GABA	284	1	-	-	2.09	4.12	-

Appendix Table A3.14 – Suprachiasmatic Nucleus Enrichment (Wen et al., 2020)

Appendix Table A3.14. Specific Hypothalamic Nuclei Level Enrichment Analysis for all imprinted genes in neuronal subpopulation in the Suprachiasmatic Nucleus (Wen et al., 2020). All column descriptions can be found in the legend of Appendix Table A3.9A

SCN Neuron Identity	Up Reg (11,391)	IG (77)	ORA *p*	ORA *q*	Mean FC IG	Mean FC Rest	GSEA *p*
Avp/Nms	909	24	3.01E-09	**1.51E-08**	3.12	1.94	**0.004**
Cck/Bdnf	831	11	0.0233	0.1167	2.17	3.09	-
Vip/Nms	816	10	0.0473	0.2367	2.54	1.95	-
Cck/C1ql3	1287	12	0.1554	0.7769	1.69	1.97	-
Vip/Grp	2026	12	0.7385	1	1.67	2.43	-

Appendix Table A3.15 – Dopaminergic Neuron Enrichment (Hook et al., 2018)

Appendix Table A3.15. Monoaminergic Nuclei Level Enrichment Analysis for all imprinted genes in dopaminergic subpopulations located throughout the mouse brain at P7 and E15.5 (Hook et al., 2018). All column descriptions can be found in the legend of Appendix Table A3.9A

Dopaminergic Cell Identity	Up Reg (13,095)	IG (87)	ORA *p*	ORA *q*	Mean FC IG	Mean FC Rest	GSEA *p*
Arcuate Nucleus (*Th, Slc6a3, Prlr*) (P7)	931	25	8.82E-10	**1.15E-08**	1.97	1.97	0.5795
Arcuate Nucleus (*Th, Ghrh, Gal*) (P7)	560	15	3.69E-06	**4.79E-05**	2.96	3.32	-
Periaqueductal Gray (P7)	667	14	0.0001	**0.0015**	2.67	3.54	-
Midbrain Neuroblast (P7)	1100	18	0.0003	**0.0036**	2.26	2.66	-
Ventral Tegmental Area (P7)	1872	25	0.0004	**0.0046**	1.88	2.33	-
Olfactory Bulb Mature Th+ (P7)	691	7	0.1750	1	3.54	3.79	-
Olfactory Bulb immature Th+ (P7)	128	2	0.2090	1	3.06	4.18	-
Olfactory Bulb Maturing Th+ (P7)	360	4	0.2173	1	4.19	4.42	-
Substantia Nigra (P7)	2553	20	0.2409	1	2.19	2.51	-
Post-mitotic midbrain neuron (E15.5)	1362	11	0.2929	1	2.30	2.06	-
Forebrain neuroblast (E15.5)	3192	19	0.7474	1	1.51	1.79	-
Midbrain Neuroblast (E15.5)	5085	29	0.8789	1	1.70	2.16	-
Post-mitotic forebrain Th+ (E15.5)	1177	5	0.9013	1	1.84	2.36	-

Appendix Table A3.16 – Ventral Midbrain Enrichment (La Manno et al., 2016)

Appendix Table A3.16. Monoaminergic Nuclei Level Enrichment Analysis for all imprinted genes in cell populations isolated from E11.5- E18.5 ventral midbrain (La Manno et al., 2016). All column descriptions can be found in the legend of Appendix Table A3.9A

Ventral Midbrain Cell Identity	Up Reg (14,008)	IG (93)	ORA *p*	ORA *q*	Mean FC IG	Mean FC Rest	GSEA *p*	GSEA *q*
Serotonergic Neurons	1781	33	1.47E-08	**3.09E-07**	4.78	2.36	0.0272	0.0544
Pericytes	1409	22	0.0001	**0.0022**	8.57	8.04	0.2209	0.4418
Dopaminergic 1 (*Th*)	347	9	0.0005	**0.0103**	3.17	2.84	-	-
Dopaminergic 0 (Immature)	218	7	0.0006	**0.0129**	4.03	2.90	-	-
Lateral Neuroblasts 2	2158	26	0.0014	**0.0290**	2.11	2.48	-	-
Oculomotor and trochlear nucleus	1994	22	0.0101	0.2116	1.83	2.28	-	-
Mediolateral Neuroblasts 3	707	10	0.0187	0.3918	2.25	2.27	-	-
GABAergic neurons 1b	943	12	0.0218	0.4585	1.84	2.03	-	-
GABAergic neurons 1a	743	10	0.0253	0.5317	2.05	2.12	-	-
Dopaminergic 2 (*Aldh1a1*)	463	7	0.0342	0.7187	3.22	3.42	-	-
Endothelial cell	2554	24	0.0432	0.9075	7.38	8.31	-	-
Mediolateral Neuroblasts 1	302	5	0.0505	1	1.76	2.24	-	-
GABAergic neurons 2	5711	46	0.0550	1	1.89	1.95	-	-
Radial Glia-like Cells 3	1141	9	0.3455	1	3.43	5.44	-	-
Red Nucleus	1021	8	0.3668	1	2.21	2.01	-	-
Mediolateral Neuroblasts 4	811	6	0.4533	1	1.54	2.42	-	-
Microglia	1283	7	0.7594	1	6.06	20.70	-	-
Radial Glia-like Cells 1	1309	7	0.7777	1	3.08	2.79	-	-
Ependymal	1388	7	0.8270	1	3.49	12.31	-	-
Neuronal Progenitor	1774	9	0.8485	1	2.32	2.74	-	-
Radial Glia-like Cells 2	1726	8	0.9002	1	2.89	3.41	-	-
Lateral Neuroblasts 1	529	4	-	-	1.76	2.93	-	-
Mediolateral Neuroblasts 5	125	4	-	-	1.86	1.92	-	-
Neuroblast Dopaminergic	196	3	-	-	3.55	2.14	-	-
Medial Neuroblast	313	2	-	-	1.48	2.34	-	-
Mediolateral Neuroblasts 2	206	1	-	-	1.13	1.82	-	-

Appendix Table A3.17 (A, B & C) – Dorsal Raphe Enrichment (Huang et al., 2019)

Appendix Table A3.17A. Monoaminergic Nuclei Level Enrichment Analysis for all imprinted genes in cell populations isolated from the Dorsal Raphe Nucleus (Huang et al., 2019). All column descriptions can be found in the legend of Appendix Table A3.9A

Dorsal Raphe Cell Identity	Up Reg (14,391)	IG (85)	ORA *p*	ORA *q*	Mean FC IG	Mean FC Rest	GSEA *p*	GSEA *q*
Dopaminergic Neurons	9408	65	0.0006	**0.0089**	11.23	3.02	1.00E-04	**7.00E-04**
Serotonergic Neurons	9496	65	0.0008	0.0125	5.83	2.92	1.00E-04	**7.00E-04**
Glutamatergic Neurons	10828	68	0.0087	0.1398	5.59	3.14	1.00E-04	**7.00E-04**
GABAergic Neurons	10588	66	0.0145	0.2323	5.10	3.06	1.00E-04	**7.00E-04**
Peptidergic Neurons	7066	56	0.0000	**0.0008**	4.91	3.60	0.0023	**0.0161**
Fibroblasts	2430	19	0.0468	0.7481	9.56	8.20	0.2697	1
Smooth Muscle Cells	2242	16	0.1241	1	7.69	7.46	0.5246	1
Polydendrocytes	4880	38	0.0036	0.0574	2.11	2.64	-	-
Pericytes	1226	14	0.0046	0.0742	7.28	7.67	-	-
Oligodendrocytes	1812	13	0.1530	1	2.14	4.42	-	-
Ependymal Cells	6214	32	0.5970	1	5.14	7.43	-	-
Endothelial Cells	3147	15	0.7042	1	3.06	6.80	-	-
Differentiating Oligodendrocytes	6484	31	0.7812	1	1.92	1.99	-	-
Astrocytes	2688	9	0.9560	1	1.87	2.94	-	-
Microglia	3003	9	0.9841	1	3.08	6.14	-	-
Perivascular Macrophages	3354	7	0.9995	1	2.23	7.45	-	-
Lymphocytes	1121	0	-	-	0.00	14.87	-	-

Appendix Table A3.17B. Monoaminergic Nuclei Level Enrichment Analysis for all imprinted genes in neuronal major populations isolated from the Dorsal Raphe Nucleus (Huang et al., 2019). All column descriptions can be found in the legend of Appendix Table A3.9A

Dorsal Raphe Cell Identity	Up Reg (14,391)	IG (85)	ORA *p*	ORA *q*	Mean FC IG	Mean FC Rest	GSEA *p*	GSEA *q*
Serotonergic Neurons	1274	18	0.0004	**0.0019**	4.03	2.66	0.1603	0.3206
Dopaminergic Neurons	1519	15	0.0316	0.1579	4.93	2.93	0.1004	0.2008

Glutamatergic Neurons	3458	28	0.0391	0.1955	1.67	1.82	-	-
Peptidergic Neurons	1298	11	0.1416	0.7082	1.83	6.42	-	-
GABAergic Neurons	2796	19	0.2860	1	1.34	1.76	-	-

Appendix Table A3.17C. Monoaminergic Nuclei Level Enrichment Analysis for all imprinted genes in neuronal subpopulations populations isolated from the Dorsal Raphe Nucleus (Huang et al., 2019). All column descriptions can be found in the legend of Appendix Table A3.9A

Dorsal Raphe Cell Identity	Up Reg (14,391)	IG (85)	ORA *p*	ORA *q*	Mean FC IG	Mean FC Rest
Neuron_GABA_*Sst_Asb4*_vlPAG_dl	1001	20	1.06E-06	**1.70E-05**	1.66	2.25
Neuron_Glu_*Vsx2_Pnoc*_vlPAG_vl	633	13	8.14E-05	**0.0013**	2.23	2.56
Neuron_5-HT_*Hcrtr1_Asb4*_DRN_vl	464	11	8.62E-05	**0.0014**	2.06	2.10
Neuron_DA_*Cck_Npw*_DRN_rm	556	12	9.85E-05	**0.0016**	2.28	2.29
Neuron_5-HT_*Prkcq_Trh*_DRN_dl	848	14	0.0004	**0.0066**	1.89	2.10
Neuron_5-HT_*Met_Trpc3*_DRN_cm	1249	15	0.0061	0.0968	3.42	5.30
Neuron_DA_*Vip_Prkcd*_DRN_dl	959	12	0.0104	0.1667	2.64	2.84
Neuron_5-HT_*Pdyn_Nos1*_DRN_dm	541	8	0.0145	0.2313	2.29	2.26
Neuron_Peptide_*Cartpt_Ucn*_EW	963	11	0.0259	0.4139	1.83	8.09
Neuron_Glu_*Pax6_Penk*_vlPAG	1905	17	0.0517	0.8279	1.53	2.06
Neuron_Glu_*Fign_Pdyn*_vlPAG_cd	2063	17	0.0940	1	1.73	2.32
Neuron_GABA_*Kit_Ebf3*_vlPAG	925	9	0.0949	1	1.53	2.23
Neuron_5-HT_*Slc17a8_Cbln2*_DRN_vm	641	6	0.1777	1	1.69	2.01
Neuron_DA_*Chrna4_Chrnb3*_CLi	808	7	0.1998	1	2.72	2.45
Neuron_Glu_*Slc17a8_Gata3*_DRN_dm	1744	12	0.3320	1	1.96	3.01
Neuron_GABA_*Asic4_Calb2*_DRN_dl	1116	8	0.3379	1	1.18	2.06
Neuron_GABA/Glu_*Crhbp_Ret*_CLi	888	1	-	-	2.91	3.19
Neuron_Glu_*Nifa_Nfix*_vlPAG_dl	524	3	-	-	3.84	2.66

Appendix Table A3.18 (A & B) – Pituitary Cell Enrichment (Ho et al., 2020)

Appendix Table A3.18A. Imprinted gene over-representation in Anterior Pituitary cells – 10x (Ho et al., 2020). All column descriptions can be found in the legend of Appendix Table A3.9A

Pituitary Cell Identity	Up Reg (11,175)	IG (70)	ORA *p*	ORA *q*	Mean FC IG	Mean FC Rest
Somatotrope	617	14	2.30E-05	**0.000184**	2.97	2.92
Lactotrope	1179	15	0.005571	**0.044565**	1.78	2.40
Melanotrope	1678	18	0.013409	0.107273	2.19	3.54
Multi-Hormone Cluster	1233	9	0.366262	1	3.11	3.11
Stem Cell (*Mki67+*)	1003	6	0.608454	1	4.12	9.80
Gonadotrope	1755	9	0.790887	1	3.97	4.53
Stem Cell (*Sox2+*)	1776	9	0.802769	1	11.59	9.00
Macrophage	2186	8	0.975755	1	51.60	75.60
Corticotrope	320	4	-	-	2.29	5.71
Folliculo-stellate cells	498	3	-	-	7.51	16.96

Appendix Table A3.19 - Pituitary Cell Enrichment (Cheung et al., 2018).

Appendix Table A3.18B. Imprinted gene over-representation in Anterior Pituitary cells – Dropseq (Ho et al., 2020). All column descriptions can be found in the legend of Appendix Table A3.9A

Pituitary Cell Identity	Up Reg (11,444)	IG (77)	ORA *p*	ORA *q*	Mean FC IG	Mean FC Rest	GSEA *p*
Lactotrope	401	11	8.71E-06	**7.84E-05**	1.72	2.14	-
Somatotrope	1036	17	2.58E-05	**0.000232**	1.97	1.87	0.4987
Corticotrope	3291	28	0.004954	**0.044589**	2.75	3.44	-
Melanotrope	2926	23	0.029694	0.267250	2.03	2.90	-
Gonadotrope	2604	19	0.090364	0.813277	2.78	3.28	-
Stem Cell (*Sox2+*)	2773	18	0.215493	1	3.24	4.84	-
Folliculo-stellate cells	1290	9	0.250136	1	32.62	6.96	-
Multi-Hormone Cluster	1312	9	0.265892	1	1.99	2.38	-
Macrophage	2666	6	0.997677	1	75.17	46.03	-
Stem Cell (*Mki67+*)	1280	1	-	-	3.19	8.86	-

Appendix Table A3.19. Imprinted gene over-representation in pituitary cells (Cheung et al., 2018). All column descriptions can be found in the legend of Appendix Table A3.9A

Pituitary Cell Identity	Up Reg (15,777)	IG (92)	ORA *p*	ORA *q*	Mean FC IG	Mean FC Rest	GSEA *p*	GSEA *q*
Somatotropes	1277	22	2.98E-06	**3.57E-05**	1.99	2.60	-	-
Thyrotropes	486	9	0.002094	**0.025123**	4.39	7.32	-	-
Posterior pituitary	3010	26	0.020724	0.248685	14.38	14.15	0.3636	1
Melanotropes	1759	17	0.024650	0.295798	1.85	4.47	-	-
Corticotropes	2728	23	0.038664	0.463974	2.37	4.66	-	-
Connective tissue	2832	22	0.090352	1	18.95	28.50	-	-
Lactotropes	4535	30	0.237678	1	1.98	2.02	-	-
Gonadotropes	8039	49	0.367498	1	8.11	3.70	0.0319	0.0957
WBCs	1620	8	0.740475	1	26.02	137.56	-	-
Endothelia	4603	22	0.892176	1	20.80	15.85	0.4886	1
Proliferating Pou1f1-cells	5103	22	0.970235	1	1.53	5.08	-	-
Pituitary stem cells	10000	49	0.982304	1	3.40	7.36	-	-

Appendix A2 Additional Tables

Appendix Table A3.20 – Pancreatic Cell Enrichment

Table A3.20A. Imprinted gene enrichment analysis in pancreatic cells in the MCA (Han et al., 2018). All column descriptions can be found in the legend of Appendix Table A3.9A

Pancreatic Cell Identity	Up Reg (9,982)	IG (64)	ORA p	ORA q	Mean FC IG	Mean FC Rest	GSEA p
Stromal cell_Smoc2 high(Pancreas)	977	20	1.87E-06	**1.68E-05**	4.24	3.32	0.2468
Stromal cell_Fn1 high(Pancreas)	714	13	0.001	**0.005**	3.44	3.54	-
Stromal cell_Mfap4 high(Pancreas)	600	8	0.039	0.350	2.30	3.89	-
Endocrine cell(Pancreas)	393	6	0.041	0.370	4.94	6.39	-
Î²-cell(Pancreas)	1088	11	0.089	0.798	21.79	5.58	-
Endothelial cell_Tm4sf1 high(Pancreas)	431	4	0.305	1	2.23	4.40	-
Dividing cell(Pancreas)	698	5	0.476	1	4.88	7.81	-
Ductal cell(Pancreas)	1779	12	0.489	1	9.46	4.81	-
Endothelial cell_Fabp4 high(Pancreas)	790	4	0.764	1	1.68	4.24	-
Acinar cell(Pancreas)	199	2	-	-	9.26	11.59	-
B cell(Pancreas)	278	2	-	-	8.50	13.99	-
Dendritic cell(Pancreas)	384	0	-	-	0.00	6.41	-
Endothelial cell_Lrg1 high(Pancreas)	362	0	-	-	0.00	6.64	-
Erythroblast_Hbb-bt high(Pancreas)	373	0	-	-	0.00	9.06	-
Erythroblast_Igkc high(Pancreas)	158	3	-	-	5.42	7.49	-
Granulocyte(Pancreas)	401	3	-	-	16.15	16.10	-
Macrophage(Pancreas)	632	1	-	-	8.05	7.88	-
Macrophage_Ly6c2 high(Pancreas)	337	0	-	-	0.00	9.21	-
Smooth muscle cell_Acta2 high(Pancreas)	261	2	-	-	9.12	7.35	-
Smooth muscle cell_Rgs5 high(Pancreas)	257	2	-	-	5.11	9.01	-
T cell(Pancreas)	277	1	-	-	5.85	10.27	-

Table A3.20B. Imprinted gene enrichment analysis in pancreatic cells (Baron et al., 2016). All column descriptions can be found in the legend of Appendix Table A3.9A

Pancreatic Cell Identity	Up Reg (11,418)	IG (77)	ORA *p*	ORA *q*	Mean FC IG	Mean FC Rest	GSEA *p*	GSEA q
pancreatic D cell	4234	40	0.000593	**0.004745**	3.35	2.69	0.3523	1
pancreatic stellate cell	2067	22	0.004166	**0.03333**	13.72	18.68	-	-
endothelial cell	1883	16	0.089188	0.713501	3.79	18.29	-	-
pancreatic A cell	4903	34	0.164903	1	2.96	1.86	0.0619	0.1857
pancreatic PP cell	3224	23	0.194995	1	1.66	1.86	-	-
pancreatic ductal cell	3510	24	0.253517	1	10.31	8.91	0.0322	0.0966
type B pancreatic cell	7525	48	0.284249	1	2.41	2.91	-	-
leukocyte	1897	6	0.976333	1	5.77	27.37	-	-
pancreatic acinar cell	411	2	-	-	2.38	25.78	-	-

Appendix Table A3.21 – Mammary Gland Cell Enrichment

Table A3.21A. Imprinted gene enrichment analysis in the mammary gland at different timepoints in the MCA (Han et al., 2018). All column descriptions can be found in the legend of Appendix Table A3.9A

Mammary Gland Timepoint	Up Reg (12,252)	IG (62)	ORA *p*	ORA *q*	Mean FC IG	Mean FC Rest	GSEA *p*	GSEA q
Involution	3842	27	0.0287	0.0860	2.72	2.70	0.2286	0.4572
Lactation	305	1	-	-	2.26	30.96	-	-
Pregnancy	6213	26	0.9351	1	2.89	2.31	0.1838	0.3676
Virgin	4993	14	0.9992	1	1.90	2.22	-	-

Table A3.21B. Imprinted gene enrichment analysis in mammary gland cells different timepoints in the MCA (Han et al., 2018). All column descriptions can be found in the legend of Appendix Table A3.9A

Mammary Gland Cell Identity	Up Reg (12,252)	IG (62)	ORA *p*	ORA *q*	Mean FC IG	Mean FC Rest	GSEA *p*	GSEA q
Stromal cell(Pregnancy)	3260	36	1.66E-07	**4.64E-06**	8.96	5.87	1.00E-04	**6.00E-04**
Stromal cell_Col3a1 high(Virgin)	1846	25	1.22E-06	**3.41E-05**	7.36	4.58	0.0099	0.0594
Stromal cell_Pi16 high(Virgin)	1282	17	0.000148	**0.00413**	7.04	6.92	0.3485	1
Stromal cell(Involution)	1243	16	0.000343	0.009599	6.18	4.95	0.1128	0.6768
Stromal cell(Lactation)	133	5	0.000551	0.015427	4.14	4.25	-	-
Muscle cell_Pi16 high(Involution)	1252	15	0.001161	0.032517	9.67	6.53	0.0882	0.5292
Endothelial cell_Fabp4 high(Involution)	727	10	0.003302	0.092462	6.64	12.29	-	-
Muscle cell_Inmt high(Involution)	690	7	0.058614	1	7.25	10.37	-	-
Luminal cell(Involution)	1116	9	0.107729	1	5.40	6.21	-	-
Endothelial cell_Aqp1 high(Involution)	726	6	0.159844	1	16.12	14.51	-	-
B cell_Jchain high(Involution)	420	4	0.162773	1	7.49	12.02	-	-
Endothelial cell(Pregnancy)	636	5	0.218465	1	11.62	18.07	-	-
Macrophage_Apoe high(Involution)	734	5	0.313576	1	7.75	7.87	-	-
Luminal progenitor(Virgin)	1170	7	0.380089	1	12.51	7.64	-	-
Secretory alveoli cell_Csn2 high(Involution)	1225	7	0.427246	1	4.38	5.93	-	-
B cell_Cd79a&Fcer2a high(Virgin)	685	4	0.459221	1	4.45	4.25	-	-
Dendritic cell_Fscn1 high(Pregnancy)	710	4	0.487329	1	15.08	11.50	-	-
Myoepithelial cell(Pregnancy)	2509	13	0.512325	1	10.42	7.51	-	-
Dendritic cell_Cst3 high(Virgin)	752	4	0.533244	1	8.01	5.32	-	-
Macrophage_C1qc high(Virgin)	892	4	0.670297	1	4.90	6.93	-	-
Secretory alveoli cell_Trf high(Involution)	1746	8	0.674328	1	5.07	4.37	-	-
Luminal cell_Krt19 high (Virgin)	2587	12	0.681868	1	6.14	5.21	-	-
Macrophage_Pf4 high(Involution)	914	4	0.689232	1	5.98	6.80	-	-
Dendritic cell_Fscn1 high(Virgin)	1190	4	0.864731	1	6.70	6.11	-	-
Macrophage(Pregnancy)	2783	9	0.961082	1	6.86	5.59	-	-
Secretory alveoli cell(Pregnancy)	6973	28	0.976838	1	6.01	4.49	0.0321	0.1926
Dividing cell(Virgin)	2793	4	0.999878	1	2.87	3.71	-	-
NK_cells_Gzmb high(Pregnancy)	4520	9	0.999977	1	3.89	4.53	-	-

Table A3.21C. Imprinted gene enrichment analysis in the mammary gland at different timepoints (Bach et al., 2017). All column descriptions can be found in the legend of Appendix Table A3.9A.

Mammary Gland Timepoint	Up Reg (16,415)	IG (96)	ORA *p*	ORA *q*	Mean FC IG	Mean FC Rest
Gestation	1885	16	0.0803	0.3212	5.55	7.40
Post Involution	1062	9	0.1680	0.6719	4.14	4.81
Lactation	2810	20	0.1997	0.7988	5.52	11.01
Nulliparous	1464	9	0.4892	1	3.06	4.22

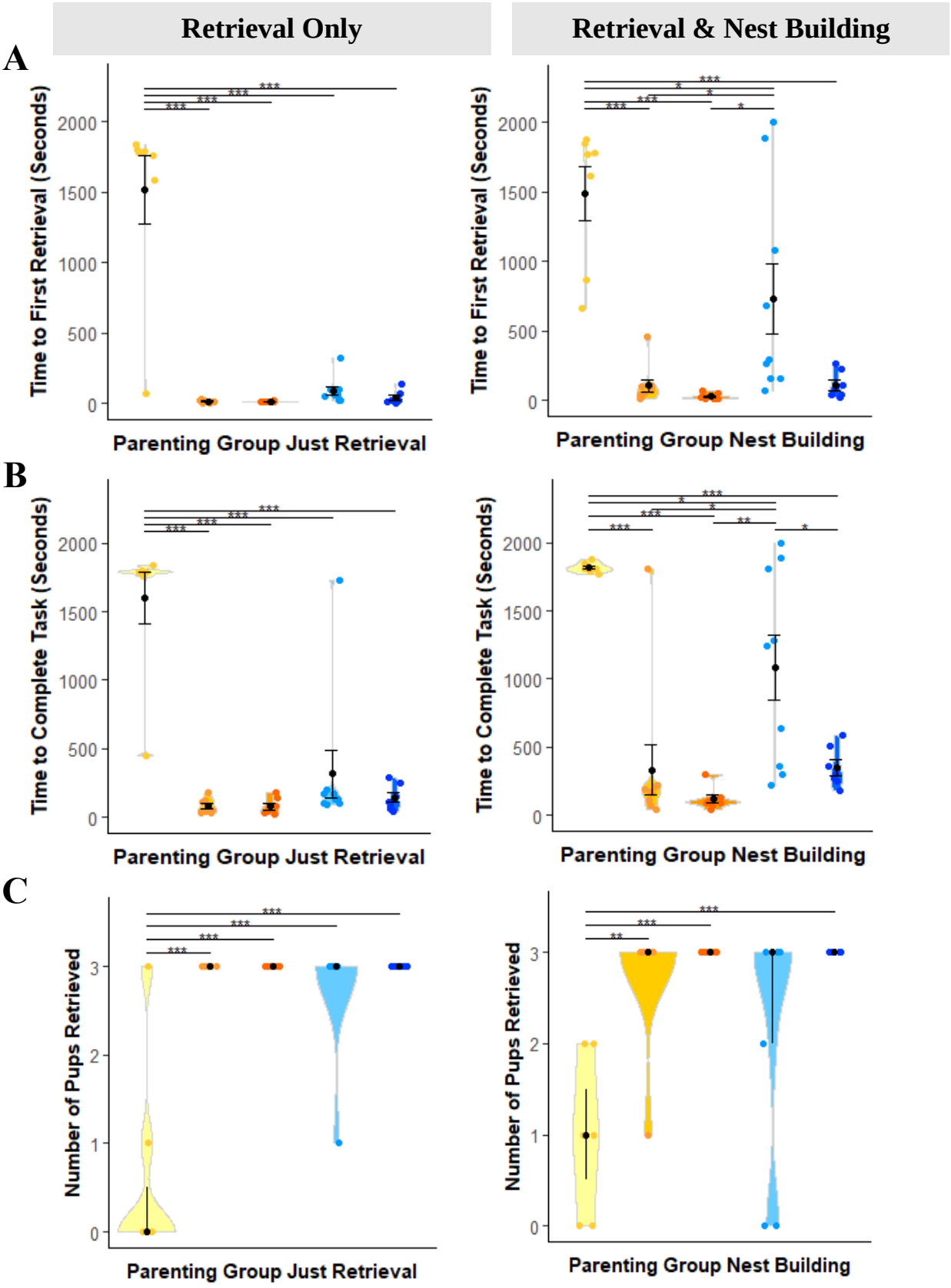
Appendix A4 – Optimisation of Retrieval/Nest Building Assessment

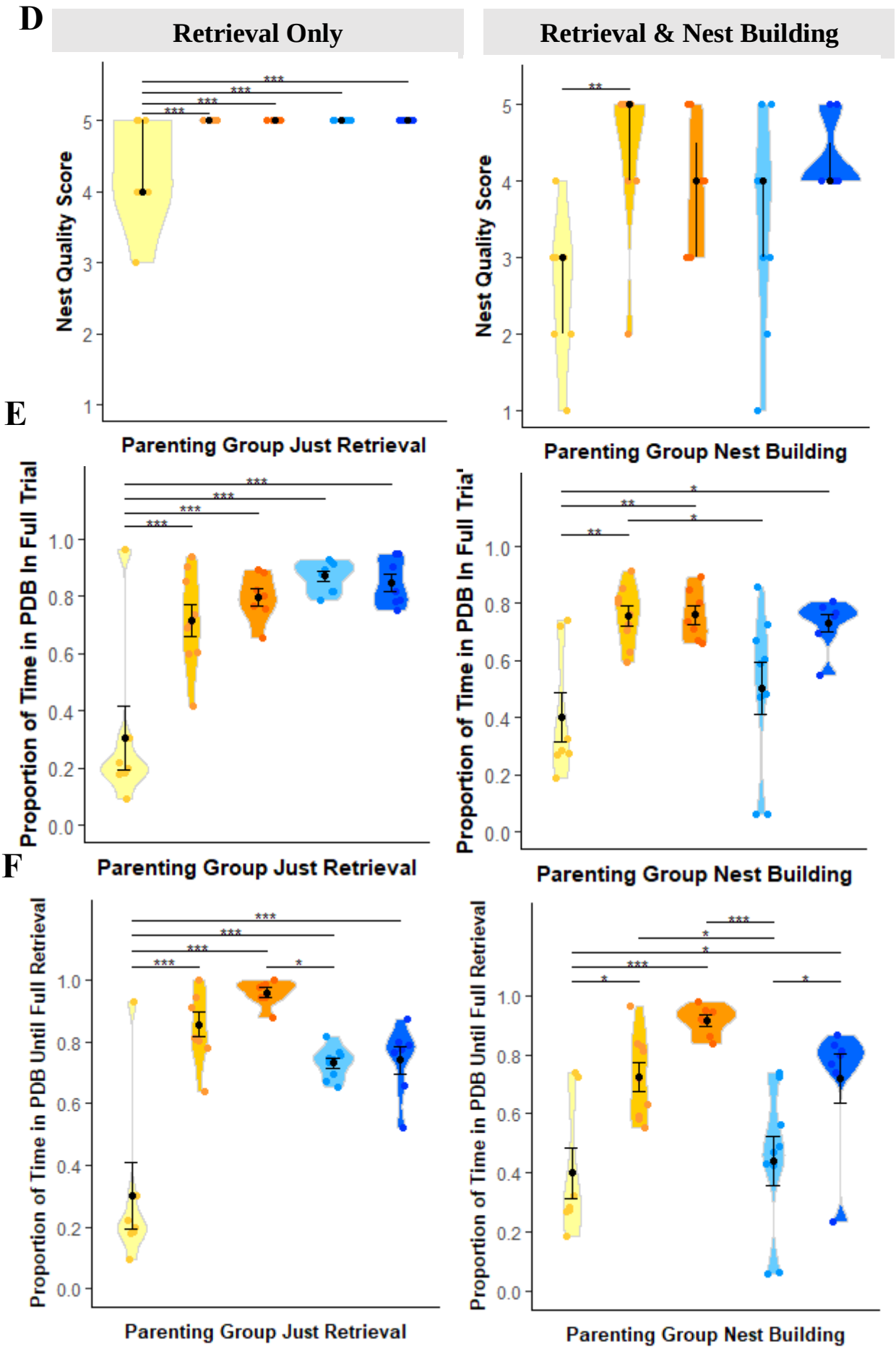
18 first-time and 14 second-time mothers and fathers along with 14 virgin females were assessed in the retrieval/best building assessment. This assessment was carried out exactly as detailed in section 2.2.4 but with a time limit of 30 minutes rather than the 60 minutes used in Chapter 5 due to animal project licensing constraints at the time of this optimisation.

Animals were split evenly between two conditions, the Retrieval/Nest Building set up as detailed in section 2.2.4 and a Retrieval assessment without the nest building. In this case, before the test in the home cage, either the home nest was left intact or was pulled apart and distributed evenly across one side of the home cage while 3 pups were distributed on the opposite side of the home cage. Both conditions allowed me to assess many of the metrics detailed in chapter 5 (see next two pages). The Figure on the next page displays data gathered from this optimisation task. Retrieval-only data for the five parenting group is presented on the left while Retrieval and nest building data is presented on the right. Data were scored for time taken to retrieve the firs pup (A) and to retrieve the last pup (B) as well as total number of pups retrieved (C). The quality of nest at the end of the assessment was scored (D) and finally the proportion of PDB engaged in by the animal up until full retrieval (E) and within the 30-minute task duration (F)

Analysis of the data shows that in the retrieval assessment, virgin females were consistently poorer than all other groups, however this test lacked the sensitivity to detect any differences between the groups of fathers and mothers. The retrieval and nest building assessment did a much better job at pulling apart differences between the groups, differences between mothers and fathers were seen in retrieval times (A,B) and in pup directed behaviour (E,F) suggesting this test would be much more useful in detecting subtle performance differences between groups all capable of displaying adequate parenting behaviour. Virgin females were statistically the poorest at all parenting behaviour and multiparous mothers are the most reliable. Interestingly second-time fathers saw improvements in parenting behaviour making their performance equivalent to mothers. Overall, this optimization experiment highlighted the value of including nest building with retrieval to increase sensitivity of measurements of parenting ability. Standard retrieval is potentially too ‘easy’ for parentally motivated animals and can be completed in a few minutes, making differences in performance and motivation hard to tease out.


Female Parenting Group
Virgin Female
Primiparous Mother
Multiparous Mother
Male Parenting Group
First Time Father
Second Time Father

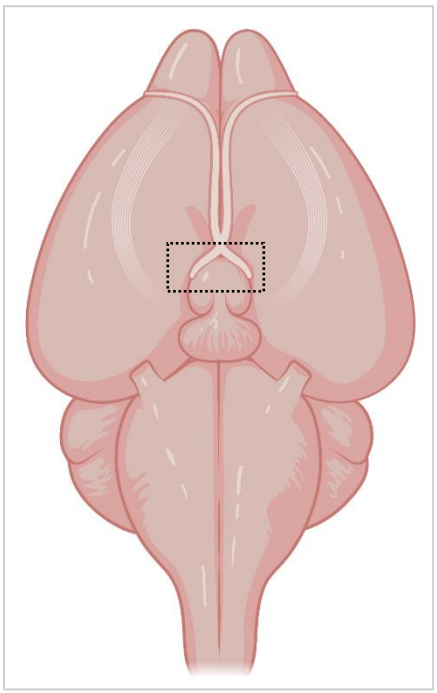

Retrieval Only
Retrieval & Nest Building
A
Time to First Retrieval (Seconds)
Parenting Group Just Retrieval
Parenting Group Nest Building
B
Time to Complete Task (Seconds)
Parenting Group Just Retrieval
Parenting Group Nest Building
C
Number of Pups Retrieved
Parenting Group Just Retrieval
Parenting Group Nest Building


D
Retrieval Only
Retrieval & Nest Building
Nest Quality Score
Parenting Group Just Retrieval
Parenting Group Nest Building
E
Proportion of Time in PDB In Full Trial
Proportion of Time in PDB In Full Tria'
Parenting Group Just Retrieval
Parenting Group Nest Building
F
Proportion of Time in PDB Until Full Retrieval
Parenting Group Just Retrieval
Parenting Group Nest Building

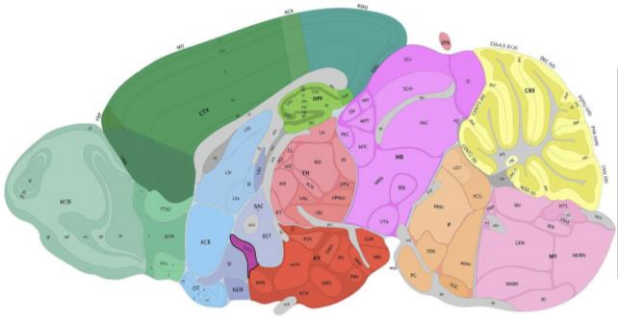
Appendix Figure A4. Retrieval/Nest Building Optimisation Metrics. Virgin females, primiparous and multiparous mothers and fathers underwent either a Retrieval-only assessment in which they must retrieve three pups to their home nest or a Retrieval/Nest Building assessment in which the three pups had to be retrieved but the home nest had been shredded which required its reconstruction during the assessment. (A) Time to retrieve the first pup in seconds, (B) Time to retrieve the final pup and hence complete the retrieval, (C) Number of pups retrieved by the end

Appendix A5 - Section guide for the POA

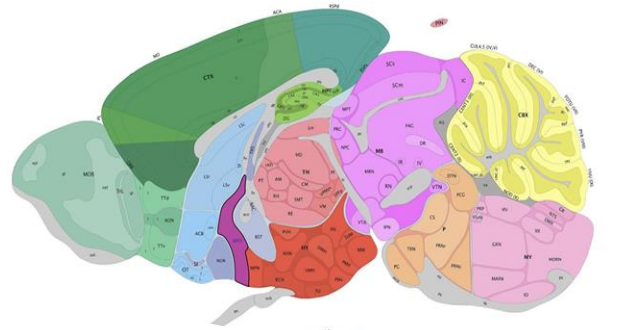
Appendix Figure A5. *Right* - 5 consecutive 200 µm Sagittal Sections, taken from the midpoint (section 5), encompassing the POA taken from the Allen Mouse Brain Reference Atlas (2011). *Bottom* – Inferior view of the mouse brain with the POA highlighted. *Next Page*. 16 coronal sections encompassing the POA (100 µm) taken from the Allen Mouse Brain Reference Atlas (2011). MPOA is highlighted in purple



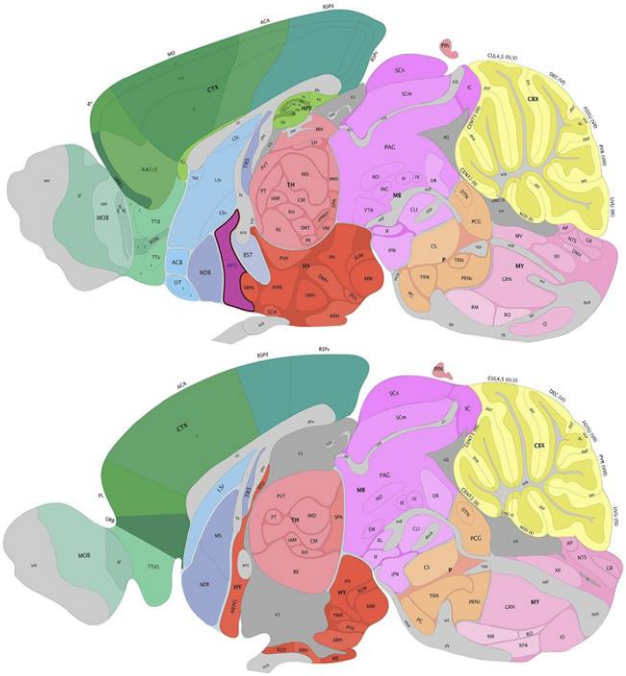
1



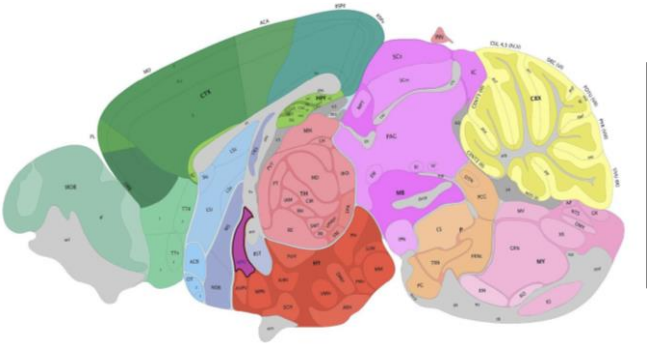
2



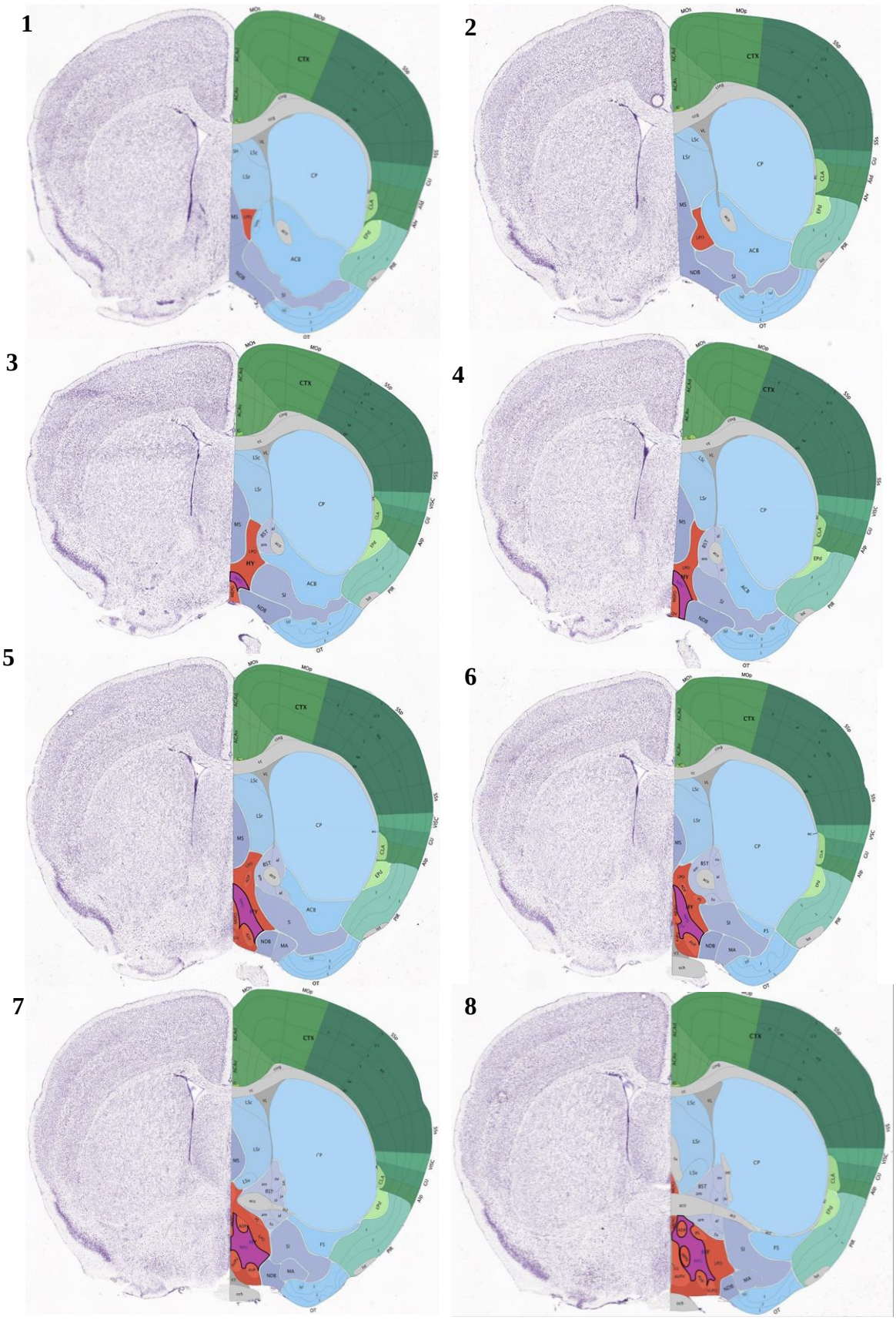
3



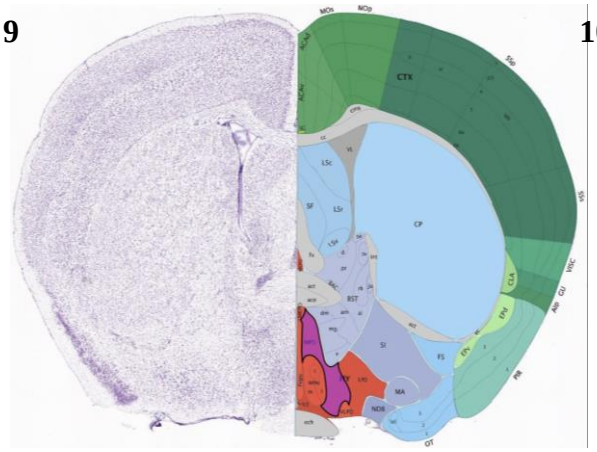
4



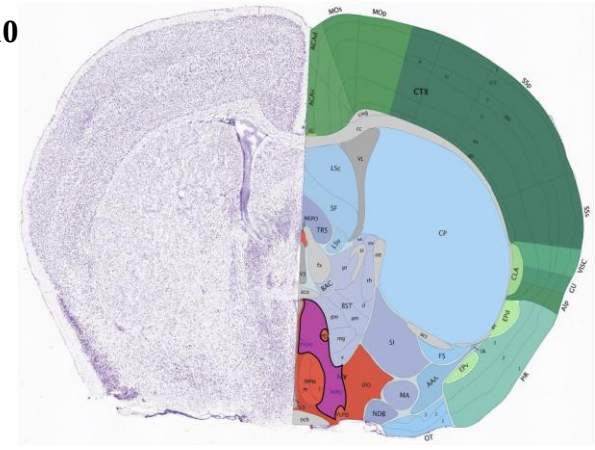
5


1
2
3
4
5
6
7
8

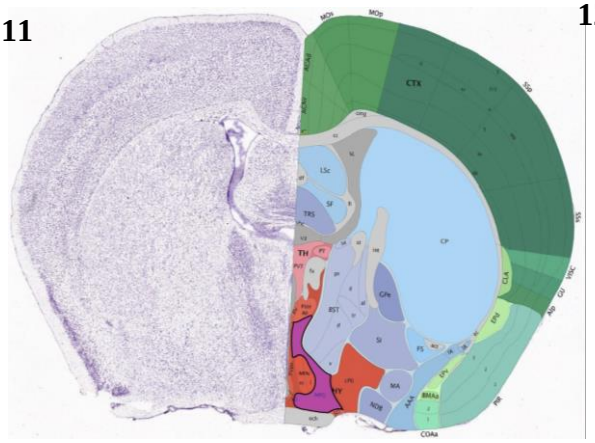
9



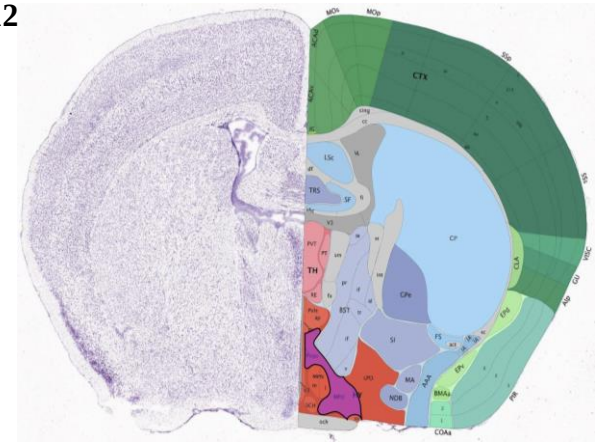
10



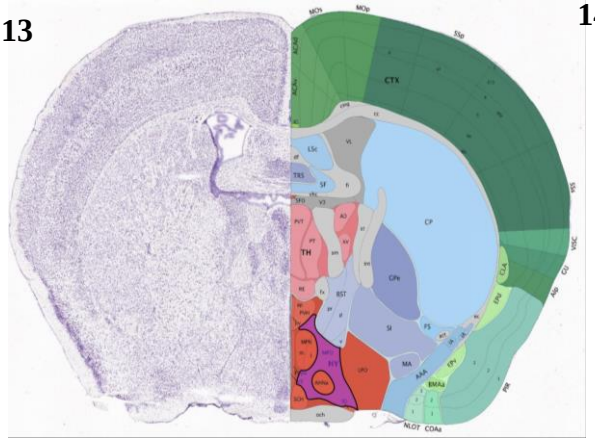
11



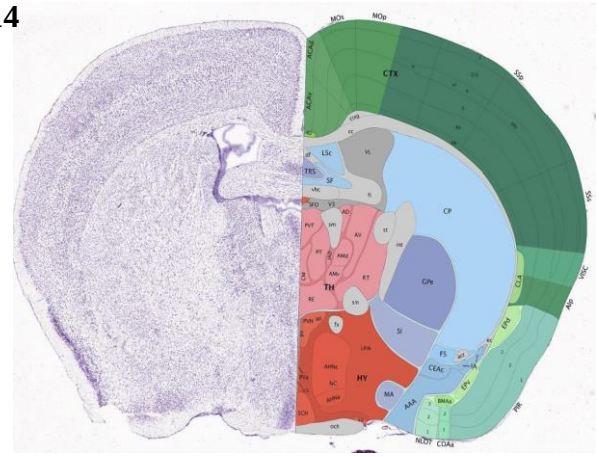
12



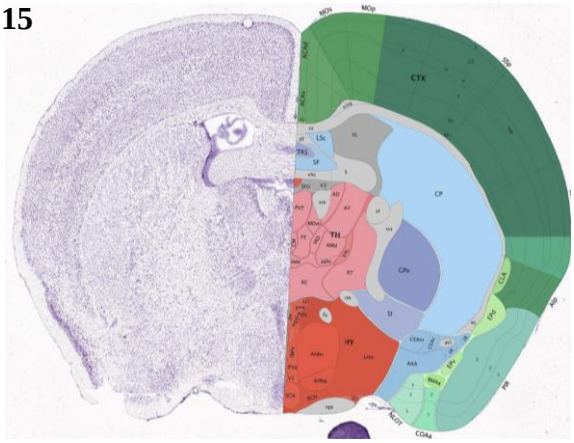
13



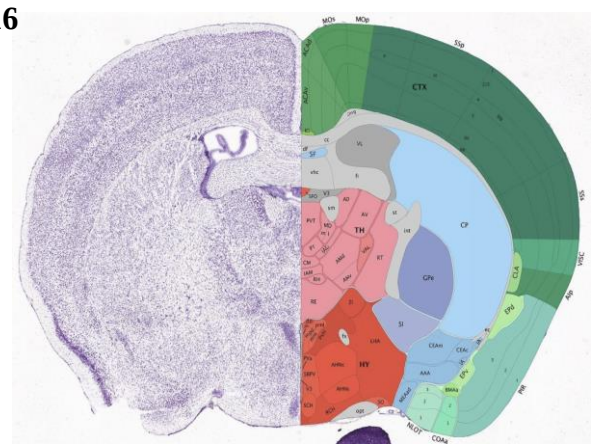
14



15



16



Appendix A6 – Chapter 4 Statistics & H-scores

Appendix Table A6.1. Post-hoc Dunn test results from *Magel2* expression across brain regions. The average number of Magel2 molecules in cells from regions of the POA, PVN, SCN, SON and surrounding hypothalamus (Rest) were compared. Kruskal-Wallis chi-squared = 2821.3, d.f. = 4, *p*-value < 2.2e-16 indicated differences between the regions and the outcome of the Bonferroni-adjusted post-hoc Dunn tests is presented in the table for all comparisons.

Comparison	Z	*P* unadjusted	*P* adjusted (Bonferroni)
POA - PVN	5.900572	3.62E-09	3.62E-08
POA - Rest	27.9454	7.50E-172	7.50E-171
POA - SCN	-35.280327	1.18E-272	1.18E-271
POA - SON	25.423861	1.37E-142	1.37E-141
PVN - Rest	5.890242	3.86E-09	3.86E-08
PVN - SCN	-30.756209	1.01E-207	1.01E-206
PVN - SON	17.721673	2.85E-70	2.85E-69
SCN - SON	42.247224	0.00E+00	0.00E+00
Rest - SCN	-44.32336	0.00E+00	0.00E+00
Rest - SON	16.578941	9.90E-62	9.90E-61

Appendix Table A6.2. Outcome of statistical tests for *Gal/Th* and *Gal/Calcr Magel2* expression comparisons. For the *Gal/Th* (top) and *Gal/Calcr* (bottom) experiments, below are the results of comparing the number of *Magel2* molecules present in these cells to the rest of all POA cells, the other *Gal* positive POA cells as well as the *Th/Calcr* positive cells. Finally, a comparison was made when removing all Magel2 non-expressing cells from the analysis and so assessing whether *Magel2* was more strongly expressed in target cells vs. other *Magel2* expressing cells.

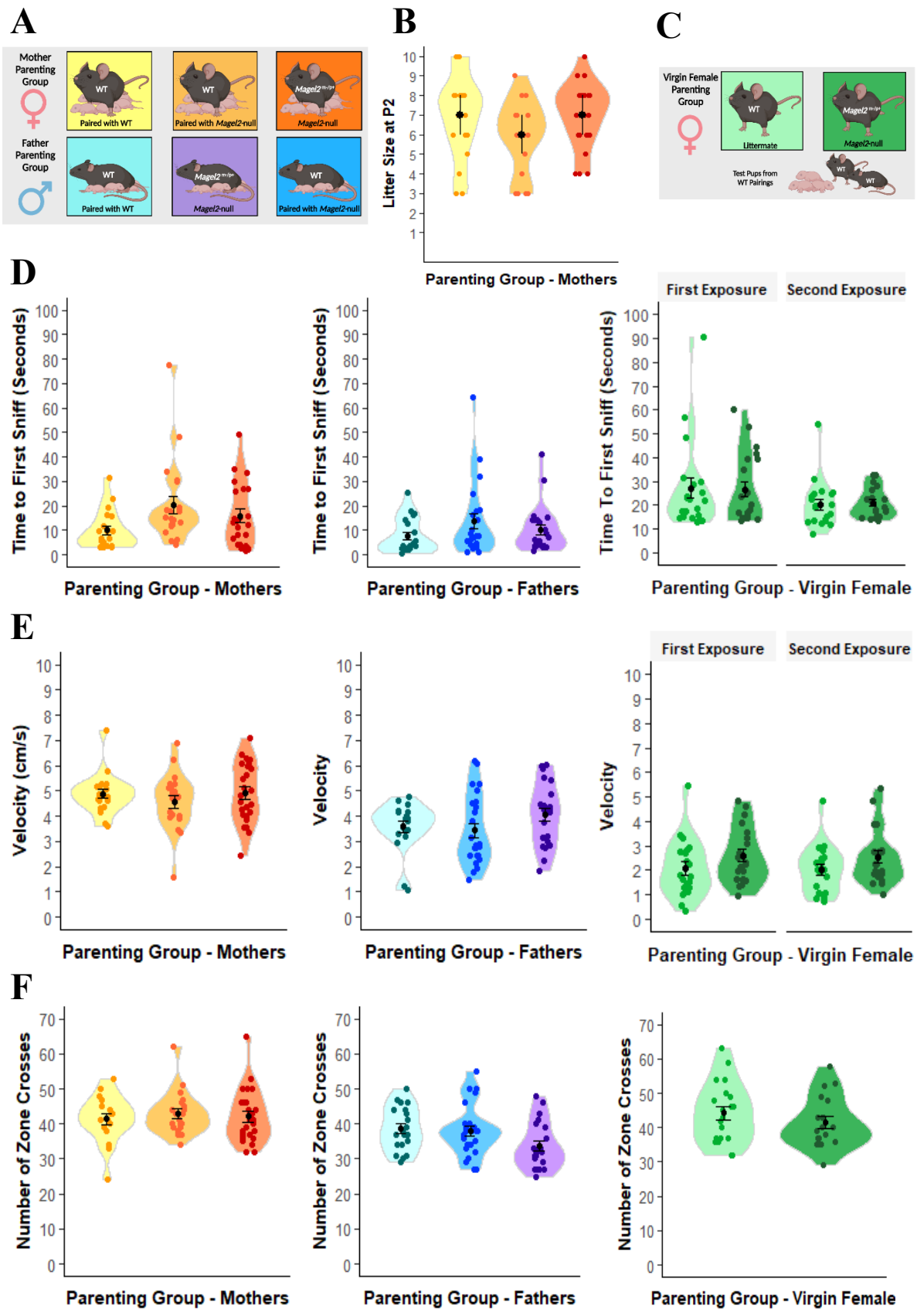
Group (N)	No. of Cells	% Cells *Magel2* Positive	Avg. No. *Magel2* molecules	Statistical Test	Test statistic vs. target group	*p* value vs. target group
***Gal/Th* POA (N=2) compared to ...**	2579	88.29%	5.28			
Rest of All POA Cells (N=2)	66,981	61.26%	2.11	Mann Whitney U	W = 47192483	< 2.2x10-16
Rest of POA *Gal* (N=2)	2,511	82.99%	4.11	Bonferroni Corrected Dunn Test	Z = 6.87	3.85x10-11
Rest of POA *Th* (N=2)	12,912	77.17%	3.22	Bonferroni Corrected Dunn Test	Z = 19.91	2.14x10-87
Rest of POA Cells (No *Gal*/No *Th*) (N=2)	51,558	56.22%	1.74	Bonferroni Corrected Dunn Test	Z = 46.5	0.00
***Gal/Th Magel2* +ve POA (N=2) compared to …**	2,277	100.00%	5.99			
Rest of All POA Cells *Magel2* +ve (N=2)	41,033	100.00%	3.45	Mann Whitney U	W = 30882369	< 2.2x10-16
***Gal/Calcr* POA (N=4) compared to …**	3846	92.23%	5.81			
Rest of All POA Cells (N=4)	152,323	54.48%	1.59	Mann Whitney U	W = 110992285	< 2.2x10-16
Rest of POA *Gal* (N=4)	23,512	76.72%	3.06	Bonferroni Corrected Dunn Test	Z = 32.64	7.34x10-233
Rest of POA *Calcr* (N=4)	4,818	84.81%	3.30	Bonferroni Corrected Dunn Test	Z = 19.08	3.46x10-81
Rest of POA *Cells* (No *Gal*/No *Calcr*) (N=4)	123,993	49.09%	1.24	Bonferroni Corrected Dunn Test	Z = 77.54	0.00
***Gal/Calcr* Magel2 +ve POA (N=4) compared to …**	3,547	100%	6.30			

Rest of All POA Cells *Magel2* +ve (N=4)	82,990	100%	2.91	Mann Whitney U	W = 75812992	< 2.2x10-16

Appendix Table A6.3. Proportion of cells with varying numbers of *Magel2* reads and the corresponding H-scores. For the experiments in Chapter 4, the proportion of different cell types in the POA/PVN which have a binned number of *Magel2* RNA molecules. H-scores are calculated from weighted calculation of proportions

Cell Type	**0 *Magel2* reads**	***1-3* *Magel2* reads**	**4-9 *Magel2* reads**	**10-15 *Magel2* reads**	**16+ *Magel2* reads**	**H-Score**	**H-score - only expressors**
POA - *Gal/Calcr*							
Gal/Calcr	0.08	0.37	0.43	0.13	0.06	173.583	188.215
Gal	0.30	0.59	0.33	0.06	0.02	115.379	150.385
Calcr	0.18	0.60	0.34	0.05	0.01	125.633	148.140
Rest	0.84	0.74	0.23	0.03	0.01	71.029	130.370
POA - *Gal/Th*							
Gal/Th	0.13	0.40	0.42	0.12	0.06	162.311	183.838
Gal	0.20	0.47	0.40	0.10	0.03	85.941	168.762
Th	0.30	0.58	0.34	0.06	0.02	140.064	152.579
Rest	0.63	0.66	0.29	0.04	0.01	117.743	140.287
PVN							
Oxt	0.31	0.51	0.16	0.02	0.01	92.031	132.547
Avp	0.16	0.42	0.33	0.08	0.02	138.336	164.023
Rest	0.49	0.37	0.12	0.01	0.00	66.550	130.956

Appendix A7 – Parenting Confounders



Appendix Figure A7. (Previous Page) Confounders of parenting measured during the retrieval and nest building task and the three chambers task in mothers, fathers and virgin females. (A) The cohorts of mothers and fathers assessed for these measures. The pairings are indicated by the male below the corresponding female. (B) Litter size recorded at P2 for the mother/father pairings. (C) The cohorts of virgin female assessed for these measures. (D) Time taken to first sniff a pup in the retrieval and nest building task for all three groups of animals. (E) Velocity of each group of animals during the 60-minute retrieval and nest building task measured via automated tracking on Ethovision. (F) The number of chamber crosses in the three-chamber assessment of the three groups of animals measured via automated tracking on Ethovision.

Appendix A8 – Retrieval & Nest Building analysis when removing litters with mutant pups

Appendix Table A8. Statistical results from mother and father analyses carried out without including parents with test litters containing mutant pups. Statistical tests used are the same as reported in the main text ***p*** values in red indicate disparities with the original statistical analysis.

Mothers (N=11)			**Pairwise p**		
Measure	**Test Stat**	**p**	**WT(WT) vs. WT(HET)**	**WT(WT) vs. HET**	**WT(HET) vs. HET**
Task Fin	20.72	**2.59×10^{-7}**	1	**6.7×10^{-7}**	**0.00021**
Task Status	15.727	**0.00038**	1	**0.001**	**0.01**
No. Pups Ret	NA	NA	NA	NA	NA
First Ret	1.34	0.271	NA	NA	NA
Last Ret	2.804	0.069	NA	NA	NA
Level 3 Nest	20.72	**2.59×10^{-7}**	1	**6.7×10^{-7}**	**0.00021**
Nest Quality	17.363	**0.00017**	1	**0.0003**	**0.015**
PDB Ret	6.862	**0.002**	1	**0.0043**	**0.035**
PDB Fin	19.26	**5.88×10^{-7}**	1	**1.8×10^{-6}**	**0.00024**
Fathers (N=11)			**Pairwise p**		
Measure	**Test Stat**	**p**	**WT(WT) vs. WT(HET)**	**WT(WT) vs. HET**	**WT(HET) vs. HET**
Task Fin	14.771	**0.0006**	0.676	**0.0004**	**0.0093**
Task Status	22.391	**1.374×10^{-5}**	0.88	**1.4×10^{-5}**	**0.0003**
No. Pups Ret	18.315	**0.0001**	1	**0.0002**	**0.00038**
First Ret	6.53	**0.003**	1	**0.0026**	**0.0159**
Last Ret	15.833	**0.0004**	1	**0.0005**	**0.001**
Level 3 Nest	16.512	**0.0003**	0.19	**0.0001**	**0.02**
Nest Quality	11.327	**0.003**	**0.044**	**0.004**	0.61
PDB Ret	4.597	**0.0145**	1	**0.014**	**0.052**
PDB Fin	6.29	**0.0036**	0.466	**0.0025**	**0.047**

Appendix A9 – Chapter 6 c-*Fos* H-scores

Appendix Table A9. Proportion of cell types (*Gal/Calcr, Gal, Calcr,* Rest) expressing the specified numbers of c-Fos molecules. Resulting H-scores are displayed for Pup-exposed mice and Controls as well as for *Magel2*-null and WT.

Group	***Gal/Calcr* 0 c-*Fos* reads**	***Gal/Calcr* 1-3 *c-Fos* reads**	***Gal/Calcr* 4-9 c-*Fos* reads**	***Gal/Calcr* 10-15 c-*Fos* reads**	***Gal/Calcr* 16+ c-*Fos* reads**	**H-score *Gal/Calcr***
Pup-Exposed	0.30	0.36	0.19	0.08	0.08	128.716443
Control	0.44	0.40	0.11	0.03	0.02	80.04342422
WT	0.35	0.39	0.15	0.05	0.05	107.0470328
Magel2-null	0.38	0.36	0.15	0.05	0.06	103.8031543
Group	***Gal* 0 reads**	***Gal* 1-3 reads**	***Gal* 4-9 reads**	***Gal* 10-15 reads**	***Gal* 16+ reads**	**H-score *Gal***
Pup-Exposed	0.49	0.38	0.09	0.02	0.03	72.37556027
Control	0.50	0.37	0.09	0.02	0.02	69.90844156
WT	0.47	0.39	0.09	0.02	0.03	75.52161532
Magel2-null	0.52	0.36	0.08	0.02	0.02	65.42610572
Group	***Calcr* 0 reads**	***Calcr* 1-3 reads**	***Calcr* 4-9 reads**	***Calcr* 10-15 reads**	***Calcr* 16+ reads**	**H-score *Calcr***
Pup-Exposed	0.44	0.39	0.11	0.03	0.03	82.174319
Control	0.50	0.40	0.08	0.02	0.02	66.4110805
WT	0.44	0.41	0.10	0.02	0.02	77.42189551
Magel2-null	0.50	0.37	0.09	0.02	0.02	70.27330912
Group	**Rest 0 reads**	**Rest 1-3 reads**	**Rest 4-9 reads**	**Rest 10-15 reads**	**Rest 16+ reads**	**H-score Rest**
Pup-Exposed	0.53	0.36	0.07	0.02	0.02	62.41413464
Control	0.57	0.35	0.05	0.01	0.01	54.9638301
WT	0.52	0.38	0.07	0.02	0.02	64.02143421
Magel2-null	0.58	0.33	0.05	0.01	0.01	53.56306974