

UC Davis

UC Davis Electronic Theses and Dissertations

Title

Single-Nucleus Molecular Signatures of Rett Syndrome

Permalink

<https://escholarship.org/uc/item/54b7x3wq>

Author

Sharifi, Osman

Publication Date

2024

Peer reviewed|Thesis/dissertation

Single-Nucleus Molecular Signatures of Rett Syndrome

By

OSMAN SHARIFI

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Biochemistry, Molecular, Cellular, and Developmental Biology

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Janine LaSalle, Chair

Ian Korf

Gerald Quon

Committee in Charge

2024

Contents

ABSTRACT.....	VI
ACKNOWLEDGEMENTS	VIII
CHAPTER 1	1
INTRODUCTION	1
1.1 PREFACE	1
1.2 ABSTRACT	1
1.3 INTRODUCTION	2
1.4 FUNDAMENTAL ASPECTS OF MECP2 FUNCTION IN MAMMALS	3
1.4.1 <i>MECP2/Mecp2 Mutations and Resulting MeCP2 Deficiencies in Brain Underlie the Majority of Rett Syndrome Like Phenotypes.....</i>	3
1.4.2 <i>Mutations in MECP2/Mecp2 Affect Virtually All Organs and Tissues.....</i>	6
1.4.3 <i>MeCP2 Is a DNA Binding Protein.....</i>	8
1.4.4 <i>MeCP2 Is Involved in Higher Order Chromatin Organization</i>	11
1.4.5 <i>MeCP2 Regulates RNA Splicing.....</i>	12
1.5 UNRESOLVED DETAILS OF MECP2 FUNCTION.....	13
1.5.1 <i>MeCP2 Functions as a Transcriptional Repressor and Activator</i>	13
1.5.2 <i>MeCP2e1 and MeCP2e2 Protein Isoforms Have Distinct Functions</i>	14
1.5.3 <i>Glial Cells Are Critical for Rett Syndrome Pathology.....</i>	16
1.5.4 <i>MECP2/Mecp2 Has Both Cell Autonomous and Non-cell Autonomous Effects</i>	18
1.5.5 <i>MeCP2 Preferentially Regulates Long Genes Over Short Genes.....</i>	19
1.5.6 <i>At What Developmental Stage Do MeCP2 Defects Impair Development</i>	19
1.5.7 <i>MeCP2 Is an RNA Binding Protein</i>	20
1.6 RECENT DEVELOPMENTS IN THE MECP2 FIELD	21
1.6.1 <i>MeCP2 Enables Liquid Phase Separation Events in Nuclear Compartmentalization</i>	21
1.6.2 <i>MeCP2 Functions with DNMT3A/Dnmt3a to Regulate Gene Expression</i>	22
1.6.3 <i>MeCP2 May Function in DNA Repair Processes.....</i>	22
1.6.4 <i>Potential Treatments Under Development for Rett Syndrome</i>	23
1.7 THE INTERACTION BETWEEN ENVIRONMENTAL FACTORS AND MECP2	25
1.8 DISCUSSION	25

1.9	REFERENCES.....	29
CHAPTER 2		50
SEX-SPECIFIC SINGLE CELL-LEVEL TRANSCRIPTOMIC SIGNATURES OF RETT SYNDROME DISEASE PROGRESSION		50
2.1	PREFACE	50
2.1.1	AUTHORSHIP	50
2.2	ABSTRACT	51
2.3	INTRODUCTION	51
2.4	RESULTS	53
2.4.1	<i>Experimental design to test longitudinal, cellular, and sex-specific transcriptional dysregulation in a symptomatically progressive mouse model of Rett syndrome.....</i>	53
2.4.2	<i>Sexually dimorphic trajectories of transcriptional dysregulation across cortical cell types</i>	54
2.4.3	<i>Co-expression networks of dysregulated genes within cortical cell types correlate with Mecp2e1 genotype, time point, sex, body weight, and disease score.....</i>	57
2.4.4	<i>X chromosome expression mosaicism in female cortical cell populations reveals dynamic non- cell-autonomous transcriptional homeostasis.....</i>	59
2.4.5	<i>Human RTT cortical cell transcriptional dysregulation is recapitulated by the female but not the male RTT mouse model</i>	62
2.5	METHODS	63
2.5.1	<i>Mecp2-e1 mouse breeding and phenotyping</i>	63
2.5.2	<i>Single nuclei isolation for mouse and human post-mortem cortex.....</i>	63
2.5.3	<i>Single nuclei-RNA sequencing</i>	64
2.5.4	<i>Pre-processing and quality control</i>	64
2.5.5	<i>Cell type identification by dimensionality reduction.....</i>	65
2.5.6	<i>DEG analysis.....</i>	65
2.5.7	<i>KEGG analysis.....</i>	66
2.5.8	<i>hdWGCNA analysis</i>	66
2.5.9	<i>WT and mutant cell parsing in mosaic female mouse cortices</i>	66
2.5.10	<i>Overlap of human and mouse DEGs and KEGG pathways.....</i>	66
2.5.11	<i>Statistics and Reproducibility.....</i>	67
2.6	DATA AVAILABILITY	67
2.7	CODE AVAILABILITY	68

2.8	DISCUSSION	68
2.9	REFERENCES.....	72
2.10	MAIN FIGURES AND TABLES.....	75
2.10.1	<i>Main Tables.....</i>	75
2.10.2	<i>Main Figures</i>	78
2.11	SUPPLEMENTAL FIGURES AND SUPPLEMENTAL TABLES	85
2.11.1	<i>Supplemental Tables.....</i>	85
2.12.1	<i>Supplemental Figures.....</i>	86
CHAPTER 3		99
EFFECTS OF PCB MIXTURES ON GENE EXPRESSION IN DISTINCT CELL TYPES OF THE		
RETT SYNDROME MOUSE MODEL CORTEX.....99		
3.1	PREFACE	99
3.1.1	<i>Authorship</i>	99
3.2	ABSTRACT	99
3.3	INTRODUCTION	100
3.4	RESULTS	102
3.4.1	<i>Experimental design to test transcriptional dysregulation in PCB-exposed WT and Mecp2e1 heterozygous female cortical cell types.....</i>	102
3.4.2	<i>Development of a novel bioinformatic pipeline for cell classification in a mosaic brain.....</i>	103
3.4.3	<i>Clustering and characterization of cortical cell types.....</i>	105
3.4.4	<i>Cell type proportions in PCB-treated versus Vehicle-treated groups</i>	107
3.4.5	<i>Differential gene expression (DEG) analysis of cell types in PCB-treated versus vehicle-treated groups</i>	108
3.4.6	<i>Dysregulated KEGG pathways across experimental conditions</i>	110
3.4.7	<i>Investigating PCB effects on Mecp2e1 mosaic cortices and cell non-autonomous interactions</i>	112
3.4.8	<i>Network analysis using hdWGCNA to investigate PCB and Mecp2 genotype effects on RTT mouse phenotypes</i>	115
3.5	DISCUSSION	118
3.6	REFERENCES.....	120
CHAPTER 4		123

SINGLE CELL-LEVEL TRANSCRIPTOMIC SIGNATURES OF RETT SYNDROME DISEASE

PROGRESSION IN THE HYPOTHALAMUS 123

4.1	PREFACE	123
4.1.1	<i>Authorship</i>	<i>123</i>
4.2	ABSTRACT	123
4.4	RESULTS	126
4.4.1	<i>Experimental design to test longitudinal, cellular, and sex-specific transcriptional dysregulation in hypothalamus of a symptomatically progressive mouse model of Rett syndrome.....</i>	<i>126</i>
4.4.2	<i>Cell type labeling and identification.....</i>	<i>127</i>
4.4.3	<i>Cell clustering based on genotype, sex, and disease time point.....</i>	<i>129</i>
4.4.3	<i>Differential gene expression in hypothalamic cell types across postnatal development.....</i>	<i>131</i>
4.4.4	<i>Functional pathway analysis of male and female excitatory and inhibitory Neurons</i>	<i>135</i>
4.5	DISCUSSION	138
4.6	REFERENCES.....	140
4.7	SUPPLEMENTAL FIGURES.....	143

Abstract

Rett Syndrome (RTT) is a neurodevelopmental disorder primarily affecting females, caused by mutations in the *MECP2* gene, which encodes the Methyl CpG binding protein (MeCP2). This dissertation explores various facets of MeCP2 function and its role in RTT, integrating findings from multiple studies to provide a comprehensive overview of the current understanding and research gaps.

Chapter 1 reviews the known functions of MeCP2, emphasizing its role in transcriptional regulation and the impact of *MECP2* mutations on RTT. The chapter summarizes key findings from human cell culture and rodent models, discusses controversies and gaps in the research, and highlights recent discoveries related to potential treatments and therapies.

Chapter 2 investigates the effects of *Mecp2e1* and MECP2 mutations on gene expression using longitudinal single-nucleus RNA sequencing (snRNA-seq 5') in a *Mecp2e1* mutant mouse model and human postmortem cortices. The study reveals significant sex differences in the number and nature of differentially expressed genes (DEGs), with mutant females showing more pronounced changes prior to symptom onset. These findings contribute to understanding the progression of RTT and the influence of non-cell-autonomous effects on transcriptional homeostasis.

Chapter 3 examines the interaction between mutant *Mecp2* and environmental factors, specifically persistent organic pollutants (POPs) like PCBs. By exposing wild-type and *Mecp2e1* mutant dams to a PCB mixture, the study identifies overlaps in dysregulated genes and pathways between PCB exposure and *Mecp2* mutations. This highlights the relevance of environmental factors in RTT pathology.

Chapter 4 focuses on transcriptional dysregulation in the hypothalamus of *Mecp2e1* mutant mice, using snRNA-seq 5' to explore cell type-specific and sex-specific variations. The analysis reveals

distinct disruptions in synaptic function, cellular metabolism, and neural development, with notable differences between the hypothalamus and cortex. This chapter provides insights into metabolic and neurodevelopmental changes associated with RTT and underscores the importance of considering both cellular and sex-specific factors.

Together, these chapters offer a detailed examination of MeCP2's role in RTT, integrating genetic, environmental, and cellular perspectives to advance the understanding of the disorder and inform future research directions.

Acknowledgements

I would like to thank my dissertation committee Dr. Janine LaSalle, Dr. Ian Korf and Dr. Gerald Quon. I am thankful to Dr. Janine LaSalle, my major advisor and principal investigator, for her exceptional mentorship throughout my PhD journey. Her expertise as an established scientist, combined with her kindness and unwavering support, has been a constant source of inspiration. Before making my decision to join UC Davis, I reached out to Dr. LaSalle to inquire about the potential for me to do a research rotation in her lab. Her enthusiastic response and her research interests were crucial in my decision to come to UC Davis. Throughout my time in her lab, Dr. LaSalle has not only guided my research but also provided me with the chance to present my work at national and international conferences in Washington DC, Nashville, Vancouver, Japan, and Australia. Moreover, Dr. LaSalle has always supported my career aspirations, including my interest in starting a startup and continuing my research as a postdoctoral fellow in her lab. Her encouragement and insightful feedback have been instrumental in my growth as a scientist. I have thoroughly enjoyed working with her and am profoundly thankful for her dedication and support. I could not have asked for a better mentor and for that, I will forever be grateful.

I also want to express my heartfelt thanks to Dr. Ian Korf, my co-mentor, who has been a cornerstone in my bioinformatics journey. After making the decision to return to graduate school from my industry job to shift my focus to bioinformatics, I had the opportunity to rotate in Dr. Korf's lab despite having no prior experience in programming or bioinformatics. From the moment I arrived, Dr. Korf welcomed me with open arms and demonstrated a remarkable commitment to my growth. He meticulously guided me through the fundamentals of computational biology and bioinformatics, offering invaluable insights and support along the way. His willingness to be available at any time to address my questions and his patient mentorship have been critical to my

understanding and development in this field. Moreover, Dr. Korf played a pivotal role in my research project, particularly in parsing mutant *Mecp2*-expressing cells from wild-type *Mecp2*-expressing cells. His expertise was indispensable in enabling us to study mosaicism in Rett syndrome brains using our single-cell datasets. Without his help, this crucial aspect of our research would not have been possible. I truly owe everything I know in bioinformatics to him.

I would also like to extend my sincere gratitude to Dr. Gerald Quon, the third member of my dissertation committee, whose expertise in single-cell biology has been invaluable to my research. Given that my PhD work centered around single-cell genomics, Dr. Quon's insights and guidance were crucial in advancing my projects. I am also grateful to his student, Nelson Johansen, whose support and collaboration further enhanced the development of my research especially regarding cell type labeling of the mouse cortex. The collective input from Dr. Quon and Nelson was instrumental in driving my projects forward, and I deeply appreciate their contributions.

I would also like to express my profound gratitude to Dr. Dag Yasui, my mentor in the lab, who has been an indispensable part of my PhD journey. Dr. Yasui's expertise in Rett syndrome biology and his support were crucial in interpreting data and writing manuscripts. His patience and willingness to entertain all my ideas, whether they were fully formed or still in the brainstorming phase, were greatly appreciated. Dr. Yasui went above and beyond in providing the resources I needed, including new pipets, single-cell kits, and laptops, which were essential for my research. Beyond his exceptional scientific acumen, Dr. Yasui has been a remarkable person who genuinely cares about his colleagues. He has been like a father figure to me throughout my PhD, always having my best interests at heart. His support extended beyond the lab; he generously offered to store my belongings in his garage during my move from Davis to Woodland, and he wrote the kindest letters of recommendation for my fellowships and awards. His generosity also shone through in organizing numerous lab outings, which made our work environment both enjoyable and uplifting. Dr. Yasui is

one of the most generous individuals I have encountered, and I cannot say enough good things about him. I am deeply grateful for his mentorship, kindness, and consistent support.

I would also like to extend my gratitude to all the current and former members of the LaSalle lab: Kari Neier, PhD, Jessica Huang, PhD, George Koudza, MD, Ensi Habibi, PhD, Julia Mouat, AJ Mendiola, Sophia Hakam, Viki Haghani, Logan Williams, Oran Gutierrez, Kelly Chau and Christina Torres. I would also like to my thanks to the current and former members of the Korf lab: Keith Fraga, PhD, Ismael Acedo, and Hannah Lyman, PhD. Your camaraderie greatly enriched my experience and contributed to the success of my research. Thank you for being such an integral part of my journey.

I also want to extend my thanks to my friends who made my journey through graduate school both enjoyable and memorable. Nick Decuzzi, PhD; Tracy Tan, PhD; Chunyang Dong (Jason), PhD; Bill Lin; Cuong Lieu, MD; and Ryan Keivanfar have each contributed to making this experience fun and fulfilling. Our countless shared moments of laughter and support have been a bright spot in my academic journey. Whether through stimulating conversations, collaborative projects, or simply unwinding together, your presence has made a significant difference, and I am deeply grateful for your friendship throughout this journey.

I am most grateful to my family for their unwavering support and love throughout my academic journey. My mother, Shafiq Sharifi, had an unshakeable belief in me even before I believed in myself. Her faith, along with the sacrifices of my father, Haider Sharifi, who worked tirelessly to secure our family's future, enabled us to migrate to the United States and provided me with the incredible opportunity to pursue an education. My sister, Frishta Sharifi, has been like a second mother to me, offering not only emotional support but also help in every way possible such as cooking specifically for me so that I could focus on my studies without the added stress of meal preparation. Her consistent support has been a tremendous comfort. I am equally grateful to my

other siblings—Tamim Durani, Hameeda Sharifi, Ali Durani, Farkhanda Sharifi, Khalida Sharifi, and Masuda Sharifi. Each of you has played a significant role in my life, offering encouragement and understanding. I especially want to thank Masuda for her pivotal role in encouraging me to apply to UC Davis and join the BMCDB program. Your support and belief in my potential have been instrumental in guiding me towards this path. Additionally, I am thankful for my nephews and nieces—Baheer Durani, Zala Durani, Sana Durani, Faraz Shahnawaz, Shazad Shahnawaz, and Zoya Shahnawaz. Their presence has brought joy and a sense of grounding throughout my journey. I owe everything to my family for their constant support, love, and sacrifices. You have all been there for me in countless ways, and I am deeply appreciative of the role each of you has played in helping me achieve my goals and reach this milestone.

Chapter 1

Introduction

1.1 Preface

Sections 1.2 -1.6 and 1.8-1.9 contains text that was co-written and published with the project scientist (Dag Yasui, PhD) in the Lasalle lab. doi.org/10.3389/fgene.2021.624290. Section 1.7 was added to this chapter to show the environmental aspect of Rett syndrome biology.

1.1.1 Authorship

Osman Sharifi, Dag Yasui

This article has been modified to satisfy the formatting requirements of graduate studies and this dissertation.

1.2 Abstract

MeCP2 protein, encoded by the *MECP2* gene, binds to DNA and affects transcription. Outside of this activity, the true range of MeCP2 function is still not entirely clear. As *MECP2* gene mutations cause the neurodevelopmental disorder Rett syndrome in 1 in 10,000 female births, much of what is known about the biologic function of MeCP2 comes from studying human cell culture models and rodent models with *Mecp2* gene mutations. In this review, the full scope of MeCP2 research available in the NIH Pubmed (<https://pubmed.ncbi.nlm.nih.gov/>) database to date is considered. While not all original research can be mentioned due to space limitations, the main aspects of MeCP2 and Rett syndrome research are discussed while highlighting the work of individual researchers and research groups. First, the primary functions of MeCP2 relevant to Rett syndrome are summarized and

explored. Second, the conflicting evidence and controversies surrounding emerging aspects of MeCP2 biology are examined. Next, the most obvious gaps in MeCP2 research studies are noted. Finally, the most recent discoveries in MeCP2 and Rett syndrome research are explored with a focus on the potential and pitfalls of novel treatments and therapies.

1.3 Introduction

MeCP2 research efforts have been extensive and have therefore greatly advanced the fields of epigenetics, neuroscience, and chromatin research. MeCP2, encoded by the *MECP2* gene in humans and the *Mecp2* gene in rodents, was first characterized as a methyl CpG DNA binding protein in 1992, thereby establishing it as an epigenetic reader of DNA methylation (Meehan et al., 1992). It was not until the 1999 discovery that mutations in *MECP2* contribute to the pathology of the rare disease Rett syndrome (RTT, OMIM #312750) that the relevance of MeCP2 to normal human development was established (Amir et al., 1999). Virtually all Rett patients are female (Reichow et al., 2015) and mosaic for cells with *MECP2* expression, as point mutations on the X chromosome are almost always paternally inherited (Trappe et al., 2001). Despite this fact, the function of MeCP2 has been predominantly studied in the nervous system of *Mecp2* null male mice, as ablation of MeCP2 in all brain cells accounts for the most severe disease phenotypes (Guy et al., 2001). Since the initial characterization of MeCP2 as a DNA binding protein in 1992 (Meehan et al., 1992), a whole field of biochemical research has emerged, culminating in the discovery that MeCP2 can organize chromatin by liquid phase separation (Fan et al., 2020; Li et al., 2020; Wang et al., 2020). Since 1999 and the discovery that mutations in *MECP2* contribute to Rett syndrome, at least 3,194 papers mentioning MeCP2 have accumulated in the NIH Pubmed database. In this review, we will examine the biologic function of MeCP2 in mammals, highlight controversial aspects of MeCP2 research, point out significant gaps in knowledge, and report on paradigm-shifting advances in the Rett syndrome field.

1.4 Fundamental Aspects of MeCP2 Function in Mammals

1.4.1 *MECP2/Mecp2* Mutations and Resulting MeCP2 Deficiencies in Brain

Underlie the Majority of Rett Syndrome Like Phenotypes

The key to defining the functions of MeCP2 is to understand the effects that hypomorphic *MECP2* has at the whole organism level. Prior to the identification of *MECP2* mutations in Rett patients, a general model of *MECP2*/MeCP2 function had been developed from early studies on the *Mecp2* gene and MeCP2 protein in mice (Lewis et al., 1992). Evidence from in vitro studies concluded that MeCP2 acted as a transcriptional repressor of genes in cis (Nan et al., 1997). In 1999 a team led by Huda Zoghbi at Baylor College of Medicine and Uta Francke at Stanford University identified the common mutations for Rett syndrome and mapped them to the *MECP2* gene located on the X chromosome (Amir et al., 1999). A Rett syndrome model with a germline *Mecp2* exon 3 and 4 deletion allele produced healthy mice at birth but male pups (*Mecp2*^{/y} null) soon displayed impaired motor skills and premature lethality, while female mice (*Mecp2*^{+/+} deficient heterozygotes) became hypoactive and exhibited motor and breathing defects after 3 months of age. Post mortem examination of male *Mecp2*^{/y} null mice revealed reduced brain and neuronal size (Guy et al., 2001). These *Mecp2* exon 3 and 4 deletion or “Bird *Mecp2* deletion” mice have been the primary animal model of Rett syndrome research. A simultaneous germline *Mecp2* exon3 deletion mouse which will be henceforth referred to as the “Jaenisch *Mecp2* deletion” model displayed very similar motor defects and premature death as the Bird *Mecp2* null mice in males and motor, hypoactivity and respiratory phenotypes in females (Chen et al., 2001). To test the hypothesis that *Mecp2* expression is necessary for normal brain function, deletion of *Mecp2* in nestin expressing cells, encompassing neurons and glia, also displayed premature death and motor defects in Bird *Mecp2*^{/y} null deletion males and delayed motor defects in Bird *Mecp2*^{+/+} deficient heterozygous female mice (Guy et al., 2001). Simultaneous histological examination of Jaenisch *Mecp2*^{/y} null deletion mice revealed reduced neuronal size along with reduced brain weight,

suggestive of brain restricted defects underlying Rett like phenotypes (Chen et al., 2001). To better examine the requirement for *Mecp2* expression during brain development, the Jaenisch team also generated mice conditional for *Mecp2* deletion in nestin expressing neurons and glia. These mice exhibited the same phenotypes as seen in *Mecp2* germline deletion mice as did mice conditional for *Mecp2* deletion in CamK expressing, post-mitotic neurons that presented with delayed neurological defects in both *Mecp2*^{-y} null and *Mecp2*^{-/+} deficient female mice and premature death in male mice (Chen et al., 2001). Together, these Bird and Jaenisch germline and conditional *Mecp2* deletion mouse models led to the conclusion that defects in *Mecp2* expression of MeCP2 protein during development is necessary for normal central nervous system function and a normal lifespan. Interestingly, subsequent deletion of *Mecp2* in specific neuronal subtypes revealed specific network defects but not reduced life span (Fyffe et al., 2008; Samaco et al., 2009; Chao et al., 2010) except for male mice with *Mecp2* deletion in somatostatin and parvalbumin neurons (Ito-Ishida et al., 2015). However, *MECP2* exon deletions homologous to *Mecp2* exon deletion model mice are rarely found in Rett patients and thus represent somewhat artificial genetic constructs of the disease.

The requirement for MeCP2 activity during murine development was examined by leveraging a system in which a stop codon in *Mecp2* could be removed using a systemically administered, tamoxifen induced, cre-lox deletion. Using this tool, a landmark series of experiments performed by teams from the Bird and Jaenisch labs indicated that the expression of *Mecp2* in neurons during early adult development was able to rescue motor defects, hypo-activity and premature death in *Mecp2*^{-y} null male mice bearing Bird and Jaenisch *Mecp2* deletion alleles (Giacometti et al., 2007; Guy et al., 2007). To further test the hypothesis that *Mecp2* expression is necessary for neurologic function after development in a whole animal, a team led by Huda Zoghbi examined the effects of the Bird *Mecp2* deletion allele in adult mice. The *Mecp2*^{-y} null male mice acquired motor defects, learning defects, apraxia and lethality within 15 weeks after *Mecp2* gene deletion (McGraw et al., 2011). From these

collective *Mecp2* deletion mouse model studies, it is clear that MeCP2 is necessary for normal neuronal function and overall health throughout the lifespan. One caveat of these and other studies is that Bird *Mecp2*^{Δ/y} null deletion males were the primary focus of these studies as they have a robust disease endpoint (death) by 20 weeks of age and present with motor defects at 6 weeks of age (Guy et al., 2001). However, the vast majority of Rett patients are female and are heterozygous for *MECP2* mutations as spontaneous *MECP2* point mutations are passed from the paternal X chromosome in sperm (Trappe et al., 2001). Female *Mecp2*^{Δ/+} deficient heterozygous deletion mice are more difficult to study as life span is normal and the motor, altered anxiety and respiratory symptoms common to *Mecp2*^{Δ/y} null deletion males arise after 6 months of age and are often subtle in presentation (Chen et al., 2001; Guy et al., 2001). It should be noted that maternally inherited *MECP2* mutations are extremely rare, tend to be gene duplications and present as a different disease in male patients (Del Gaudio et al., 2006).

Since the generation of the initial *Mecp2* deletion mouse models, multiple *Mecp2* gene “knock in” models that are based on actual *MECP2* mutations identified in Rett patients have been developed. The model that may have the greatest relevance to Rett is the T158A model developed by Zhou (Goffin et al., 2012) as mutation of threonine 158 is the most common *MECP2* missense mutation according to the RettBASE database (mecp2.chw.edu.au) and affects the ability of MeCP2 to bind to DNA. The non-sense *MECP2* mutation R168X correlates with severe disease symptoms in Rett. This mutation when recapitulated in *Mecp2* resulted in motor, cognitive and anxiety defects in *Mecp2*^{R168X/y} males and females, although the time of onset was delayed in females (Schaevitz et al., 2013). In fact, male T158A mice (*Mecp2*^{T158A/y}) recapitulated the motor and learning defects as well as premature mortality observed in male *Mecp2* null mice (Goffin et al., 2012) while female *Mecp2*^{T158A/+} exhibit breathing abnormalities similar to those observed in Rett females (Bissonnette et al., 2014). In 2014 a mouse model based on a human *MECP2* exon 1 mutation showed motor defects similar to Bird *Mecp2*

null mice along with altered anxiety and stereotypic behavior in *Mecp2-e1^{-y}* males (Yasui et al., 2014). Additional *Mecp2* knock in mouse models based on Rett *MECP2* mutations have been developed and exhibit a range of motor and behavioral phenotypes as reviewed in Schmidt and Cardoso (Schmidt et al., 2020). Rat *Mecp2* knock in models may have advantages over mouse models as a *Mecp2* truncation model exhibits neurologic regression in *Mecp2^{+/+}* females and thus more accurately models Rett syndrome than some existing mouse models (Veeraragavan et al., 2015).

1.4.2 Mutations in *MECP2/Mecp2* Affect Virtually All Organs and Tissues

Although the most severe disease phenotypes were observed in *Mecp2^{-y}* null deletion male mice with deletion engineered in brain (Chen et al., 2001; Guy et al., 2001), defects in *MECP2* expression in humans has the potential to affect almost all organs and cell types, as transcripts are detected at significant levels in virtually all adult human tissues (**Figure 1**). At the systems biology level, male mice engineered to have only peripheral deletion of *Mecp2* outside of the nervous system were hypoactive, had reduced exercise capacity and bone defects but survived beyond 50 weeks after birth (Ross et al., 2016). This study, led by Stuart Cobb, established that *Mecp2* expression in multiple organs and cell types likely accounts for these broad phenotypes thereby underlining the importance of MeCP2 function outside of the nervous system.

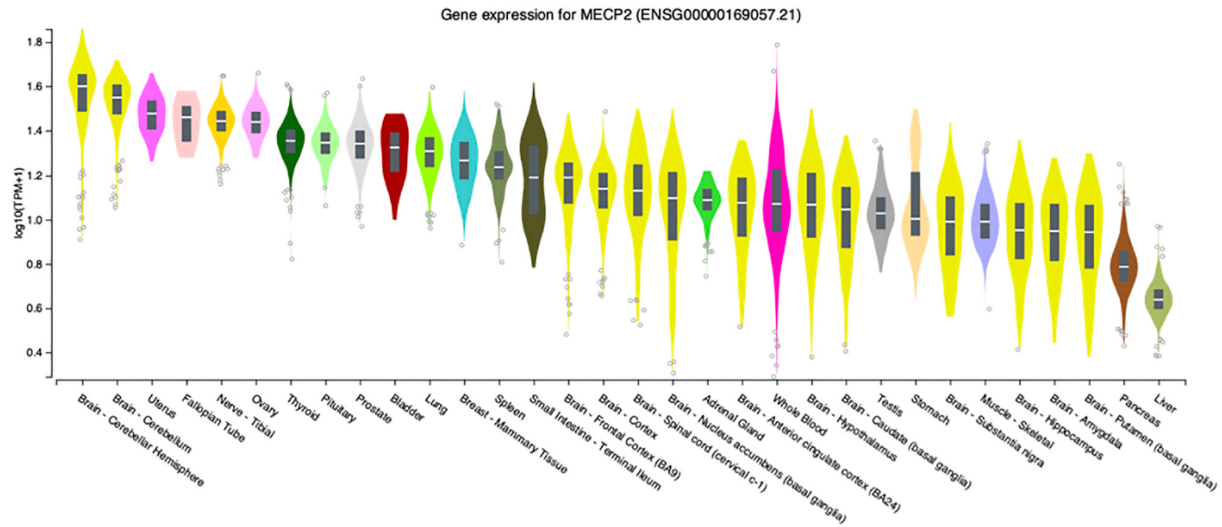


Figure 1: *MECP2* transcript levels in select human tissues. violin plots from the Genome-Tissue expression (GTEx) project are shown for *MECP2* with log10 values ranked from high to low levels.

These limited studies indicate that while *MeCP2* is necessary for normal neurologic function in mice, disease is still present in other tissues. Similarly, loss or diminished *MeCP2* activity contributes to Rett syndrome pathology outside of the nervous system. For example, an early report in 1994 described extended QTc intervals in Rett patients (Sekul et al., 1994). Analysis of *Mecp2* deletion mice revealed similar cardiac abnormalities as the patients as well as ventricular tachycardia (McCauley et al., 2011). To further investigate this finding, conditional deletion of *Mecp2* from cholinergic neurons was performed by a team led by Jeff Neul. They found that loss of *Mecp2* from cholinergic, parasympathetic neurons recapitulated the previous cardiac findings in Bird *Mecp2* null mice and symptoms in Rett patients (Herrera et al., 2016). Interestingly, while *MECP2* transcripts have relatively low abundance in the human liver (**Figure 1**) an ENU mutagenesis screen in Bird *Mecp2*^{+/y} null male mice revealed that *MeCP2* function is required for normal lipid metabolism as these animals have dysregulated cholesterol synthesis in brain and elevated cholesterol in liver and serum (Buchovecky et al., 2013; Kyle et al., 2016; Kyle et al., 2018). *MECP2* is also expressed at high level in human intestine (**Figure 1**) and 66% of Rett patients report gastrointestinal pain (Symons et al., 2013). A 2016 study

found that *Mecp2* null mice had gut hypomotility and reduced nitric oxide synthase expression in enteric neurons (Wahba et al., 2015). To examine the hypothesis that microbial alterations underlie intestinal defects, the overall gut microbiome diversity was examined and appeared to be reduced in Rett patients (Strati et al., 2016; Borghi et al., 2017). As the diets were similar between the Rett and control subjects, these results suggest that altered *MECP2* levels in the digestive tract may alter the intestinal environment and thus microbe growth. In fact, it was recently shown that mice with loss of *Mecp2* expression solely in the intestine have severe colonic epithelial defects (Millar-Büchner et al., 2016). As there is direct signaling from the gut to the brain, extension of these gut/microbiome studies should be extended to Bird *Mecp2*^{+/+} deficient female deletion mice and would thus provide critical insights for female Rett patients. While studies on the role of MeCP2 function in tissues outside of the brain, a recent study found evidence that MeCP2 represses LINE-1 activity in brain, but not in peripheral tissues such as heart and eye (Zhao et al., 2019).

1.4.3 MeCP2 Is a DNA Binding Protein

MeCP2 was first described in the lab of Adrian Bird as a DNA binding protein with a high affinity for CpG methylated DNA (Meehan et al., 1992) via the methyl DNA binding domain (MBD) (Free et al., 2001). Although this report is often overlooked, MeCP2 was later found to be homologous to the nuclear matrix binding protein ARBP, cloned from chicken (Weitzel et al., 1997). This finding is intriguing as the nuclear matrix structures the nucleus in eukaryotic cells and organizes chromatin for replication, DNA repair and transcription among other functions (Wasag and Lenartowski, 2016). These in vitro binding results from the Stratling lab suggested that ARBP/MeCP2 was involved in chromatin loop organization and heterochromatin structure (Weitzel et al., 1997). Basic biochemical analyses of MeCP2 by Jeff Hansen and colleagues revealed that MeCP2 is a highly disordered protein outside of the MBD domain (Adams et al., 2007). Significantly, another study by the Hansen and Woodcock team demonstrated that a common Rett *MECP2* mutation, R106W in the MBD altered

MeCP2 interaction with nucleosomal DNA (Nikitina et al., 2007). Consistent with these results, MeCP2 lacking the MBD domain was found to bind with low affinity to chromatin in vivo (Stuss et al., 2013). The Hansen group later determined that MeCP2 contains at least two distinct DNA binding domains, the MBD and sequences in the carboxyl terminus (Ghosh et al., 2010).

While Woodcock and colleagues had identified DNA binding regions in MeCP2 outside of the MBD, these sites were not well characterized at the time (Ghosh et al., 2010). Later however, these non-MBD binding sites were identified as AT hook regions by the Zoghbi lab in 2013 (Baker et al., 2013). The description of AT hook domains in MeCP2 fits well with a 2005 description of high affinity binding sites containing AT runs adjacent to methylated CpG sites (Klose et al., 2005). To summarize these diverse findings, a very recent paper nicely illustrates the in vivo activity of MeCP2 in live cells. In this study Nat Heintz and colleagues leveraged MeCP2 fluorescent tagging and real-time visual mobility analysis to describe the sum effect of stable MBD domain interactions and transient AT hook interactions that slow the diffusion of MeCP2 from DNA in living cells (Piccolo et al., 2019).

Since the early characterization of MeCP2 as a methylated CpG binding protein (Meehan et al., 1992), the range of MeCP2 binding activity has been extended. A discovery by the Heintz lab revealed that 5-hydroxymethylcytosine (5hmc) is highly abundant in mouse brain (Kriaucionis and Heintz, 2009). This finding was confirmed by Anjana Rao and colleagues who also identified TET1 as the enzyme responsible for the conversion of 5-methylcytosine (5mc) to 5hmc in brain (Tahiliani et al., 2009). Later it was determined by the Heintz lab that MeCP2 can bind to 5hmc with similar affinity as 5mc in actively transcribed genes (Mellén et al., 2012). However, in vitro results had previously shown that 5hmc inhibits binding of the MBD domain of MeCP2 to DNA (Valinluck et al., 2004). Although the activity of TET1 and related factors TET2 and TET3 can each lead to active CpG demethylation, potentially altering the transcriptional state of genes (He et al., 2011) the biologic implications of MeCP2 binding to 5hmc is still not clear.

While 5mc and 5hmc in CpG dinucleotides account for the majority of MeCP2 binding sites in the mammalian genome, MeCP2 can bind to other motifs. In 2014 the Hongjun Song lab revealed that surprisingly, 25% of CpA, CpC, and CpT (CH) dinucleotides are methylated in the brain compared with 75% of CpG sites (Guo et al., 2014). The Song lab also showed that MeCP2 was able to bind to methylated CpH in vitro (Guo et al., 2014). The following year, Harrison Gabel and other members of the Greenberg lab determined that MeCP2 binds to methyl CpA sites within gene bodies, although the effect of this binding is controversial as described in a subsequent section of this review (Gabel et al., 2015). More recently the binding of MeCP2 was correlated with the methylation status of CH sites in the brain (Lavery et al., 2020). These results conflict with data indicating that MeCP2 binding to DNA is mostly due to CpG methylation and that MeCP2 binds to promoters with low CpG density (Baubec et al., 2013). While it is clear that MeCP2 binds to dinucleotides with methylated cytosine, an unbiased statistical analysis of genome-wide MeCP2 ChIP-seq data sets indicated that the GC content of a particular genomic region is the most predictive of MeCP2 binding (Rube et al., 2016). In contrast to these results, recent in vitro and in vivo studies indicate that MeCP2 has minimal binding to non-methylated GT rich DNA (Connelly et al., 2020). To further define MeCP2 binding, Buchmuller et al. (2020) found that MeCP2 was able to bind to DNA in vitro with the asymmetric cytosine modifications C/mC, mC/mC, mC/hmC, and mC/fC, where C is cytosine, mC is 5 methyl cytosine, hmC is 5 hydroxymethylcytosine and fC is 5 formylcytosine. These recent findings appear to account for the ability of MeCP2 to bind virtually anywhere in the genome. It is important to note that the level of MeCP2 protein expression is critical for normal brain function as mice with 50% of normal expression (Kerr et al., 2008; Samaco et al., 2008) and twice normal expression (Collins et al., 2004) exhibit neurological defects. Recent studies suggest that altered neuronal MeCP2 levels correlate with heterochromatin changes and behavioral symptoms (Ito-Ishida et al., 2020).

1.4.4 MeCP2 Is Involved in Higher Order Chromatin Organization

Another fundamental aspect of MeCP2 function that is often overlooked is its ability to regulate large chromosomal domains. For example, some of the earliest in vitro structural studies revealed that MeCP2 could bind and condense nucleosomal arrays that incorporate both methylated and unmethylated DNA (Georgel et al., 2003). Interestingly, these studies did not indicate that MeCP2 was able to form dimers, suggesting that one MeCP2 molecule could bridge two independent arrays. A similar chromatin condensation activity was shown in myoblast cell lines by the Cardoso lab (Brero et al., 2005). Co-immunoprecipitation studies by the El-Osta lab found that the SWI/SNF chromatin remodeling factors Brahma and BAF57 can be recruited to DNA by MeCP2 (Harikrishnan et al., 2005). While these studies examined recombinant MeCP2 linking of artificial nucleosomal constructs, a study from the Kohwi-Shigematsu lab identified a functional chromatin looping activity in vivo. In this study Shin-Ichi Horike and colleagues found that MeCP2 mediated long range looping and thus transcriptional activity of the *Dlx5* gene (Horike et al., 2005). In a similar study it was found that long range interactions between the Prader-Willi imprinting center (PWS-IC) and *CHRNA7* modulated gene expression in neurons (Yasui et al., 2011). These studies in cell lines were also consistent with in vitro studies from the Hansen lab revealing that MeCP2 was able to interact with DNA outside of the MBD domain, thereby accounting for the ability of one MeCP2 molecule to bind to two sites simultaneously in chromatin looping (Nikitina et al., 2007). This 2007 study also found that MeCP2 molecules encoded by frequent Rett *MECP2* mutations were deficient in their ability to compact chromatin (Nikitina et al., 2007). Studies from the LaSalle lab extended this chromatin looping aspect of MeCP2 function further by showing that MeCP2 along with CTCF contributes to both inter and intra chromosomal interactions and gene regulation of 15q11-13 (Meguro-Horike et al., 2011; Yasui et al., 2011). Later studies from the Natalie Berube lab showed that MeCP2 recruits ATRX to imprinted genes to regulate CTCF chromatin looping interactions (Kernohan et al., 2014). Recent

atomic force microscopy (AFM) studies reveal that MeCP2 forms loops by bridging distant binding sites on continuous DNA strands (Liu et al., 2020).

1.4.5 MeCP2 Regulates RNA Splicing

While MeCP2 can affect gene transcription directly, one relatively unexplored activity of MeCP2 is its effect on RNA splicing. The first genome wide study of MeCP2 effects on splicing was performed in the Huda Zoghbi lab in 2005. These studies revealed significant splicing defects in *Mecp2*^{308/y} male mice expressing a truncated MeCP2 protein (Young et al., 2005). In vitro studies revealed that MeCP2 bound to RNA along with YB-1 resulting in gene construct splicing events in neuronal cell lines (Young et al., 2005). Following this report, few attempts were made to confirm the role of MeCP2 in RNA splicing regulation. It was not until 2013 that the Keji Zhao lab reported that *Mecp2* ablation correlated with alternatively spliced exon skipping (Maunakea et al., 2013). Research from the LaSalle lab in 2014 confirmed MeCP2 association with YB-1 shown by Young et al. (2005) as well as interaction with the RNA splicing factors MATR3, SFPQ, and SFRS1 (Yasui et al., 2014). The mechanism underlying MeCP2 regulation of splicing was further explored by the Rasko lab. They found that MeCP2 depletion reduces splicing factor recruitment to methylated DNA and increases intron retention events in blood cells (Wong et al., 2017). More recently Nurit Ballas and colleagues reported abnormal intron retention exon skipping in activation induced genes in *Mecp2* null neurons compared to wild-type neurons (Osenberg et al., 2018). The latest investigation of the role of MeCP2 in RNA splicing regulation using a machine learning approach, however, found subtle changes in RNA splicing in cells with varying levels of MeCP2 (Chhatbar et al., 2020). While the results may be not always be direct, or of great magnitude, MeCP2 binding to DNA has an effect on mRNA splicing events. However, one wonders how much current RNA-seq methodologies bias this sort of approach.

1.5 Unresolved Details of MeCP2 Function

1.5.1 MeCP2 Functions as a Transcriptional Repressor and Activator

The model of MeCP2 as a direct transcriptional repressor of single copy genes in cis derived from in vitro studies dating back to 1992. In 1997 this model was formally proposed in a report showing that affinity purified rat MeCP2 showed repressive activity of CpG methylated gene promoter constructs in vitro (Nan et al., 1997). Subsequent studies from the Alan Wolffe lab demonstrated that MeCP2 recruitment of histone deacetylase activity was correlated with transcriptional repression (Jones et al., 1998). An example of MeCP2 functioning as a transcriptional repressor is the regulation of viral elements that exist in the genome as multiple, highly CpG methylated elements. Specifically, LINE-1 elements were found to be transcriptionally repressed by MeCP2 binding (Lorincz et al., 2001; Yu et al., 2001). In 2010, a paradigm shifting report published by the Gage lab found that mammalian neuronal progenitor cells have a massive activation of LINE-1 transcription and transposition that is absent in cells from peripheral tissues (Muotri et al., 2010). More recently the Cardoso lab demonstrated in cell lines that MBD proteins including MeCP2 repress TET1 mediated 5hmc conversion and activation of LINE1 elements in human cells (Zhang et al., 2017). A recent analysis of LINE1 insertions revealed in that Rett tissues had fewer cells with exonic insertion, consistent with selection against neurons with multiple LINE1 insertion events and other cell types (Zhao et al., 2019). Thus, in vitro and in vivo evidence indicates that MeCP2 binds to CpG methylated DNA promoters and represses transcription of genes and viral elements along with co-factors such as HDACs.

The model of MeCP2 as a transcriptional repressor of single copy genes was first challenged by the genome wide correlation of MeCP2 binding with gene transcripts. In 2007, MeCP2 ChIPseq analysis combined with gene transcript analysis from the LaSalle lab found that MeCP2 bound to active promoters with low levels of DNA methylation (Yasui et al., 2007). This was closely followed by results published in 2008 from the Zoghbi lab showing that loss of MeCP2 correlated with both

reduced and elevated levels of gene transcripts in murine hypothalamus (Chahrour et al., 2008). Further analyses of select gene promoters upregulated by loss of MeCP2 activity revealed binding of the transcriptional activator CREB1 along with MeCP2 to expressed genes (Chahrour et al., 2008). In support of these findings, in 2017 it was shown that *MECP2* mutant and *MECP2* null human embryonic stem cell derived forebrain neurons had reduced CREB levels along with reduced dendritic complexity, neurite growth, and mitochondrial function (Bu et al., 2017). Most recently, results from the Harrison Gabel lab suggest that DNA methylation at enhancers may repress genes in trans (Clemens et al., 2020). Clearly these results show that MeCP2 binding, and recruitment of co-factors correlates with both activation and repression of transcription. However, the location of MeCP2 binding and co-factor interactions appear to be critical for how gene expression is affected.

1.5.2 MeCP2e1 and MeCP2e2 Protein Isoforms Have Distinct Functions

Prior to 2004, MeCP2 was believed to be the gene product of three exons comprising one protein isoform. In that year, a team from The Hospital for Sick Children in Toronto described a novel exon upstream of the known *MECP2* exons that could splice to exons 3 and 4 to produce a novel protein isoform with a distinctly different amino terminus than the originally described MeCP2 isoform (Mnatzakanian et al., 2004). The presence of this novel exon was also shown in mouse *Mecp2* by RT-PCR (Mnatzakanian et al., 2004). Around the same time, the homologous upstream exon in murine *Mecp2* was also reported by the lab of Adrian Bird (Kriaucionis and Bird, 2004). Therefore, the alternative splicing of *MECP2/Mecp2* exon 1 to exons 3 and 4 produces the MeCP2e1 isoform while translation of mRNA with all four exons leads to production of MeCP2e2 due to the use of an alternative translational start site (**Figure 2**). In humans and rodents, this produces MeCP2 isoforms which have almost identical amino acid sequences but with distinctly different amino termini (**Figure 2**). Furthermore, it was shown that mutations in *MECP2* exon 1 are present in Rett patients

(Mnatzakanian et al., 2004) and that *MECP2* exon 1 mutations disrupt translation of the MeCP2e1 but not MeCP2e2 isoforms (Gianakopoulos et al., 2012).

Evidence that MeCP2e1 is the isoform underlying Rett syndrome was provided by two key studies. In the first study, knockout of *Mecp2* exon 2, which only affects production of MeCP2e2, resulted in mice with normal development and function (Itoh et al., 2012). However, *Mecp2* exon 2 deletion did result in placental defects (Itoh et al., 2012). To follow up on these findings, the Berge Minassian lab screened a cohort of patients with idiopathic RTT. In one patient, they found an *MECP2* exon 1 A to T point mutation affecting the translational start site (Saunders et al., 2009). This exon 1 mutation was predicted to prevent translation of only the MeCP2e1 isoform but not the MeCP2e2 isoform, suggesting that it is the loss of MeCP2e1 activity that results in Rett syndrome. Therefore, the LaSalle lab replicated this mutation in a mouse model, which resulted in the complete loss of the MeCP2e1 protein isoform while retaining expression of MeCP2e2 (Yasui et al., 2014).

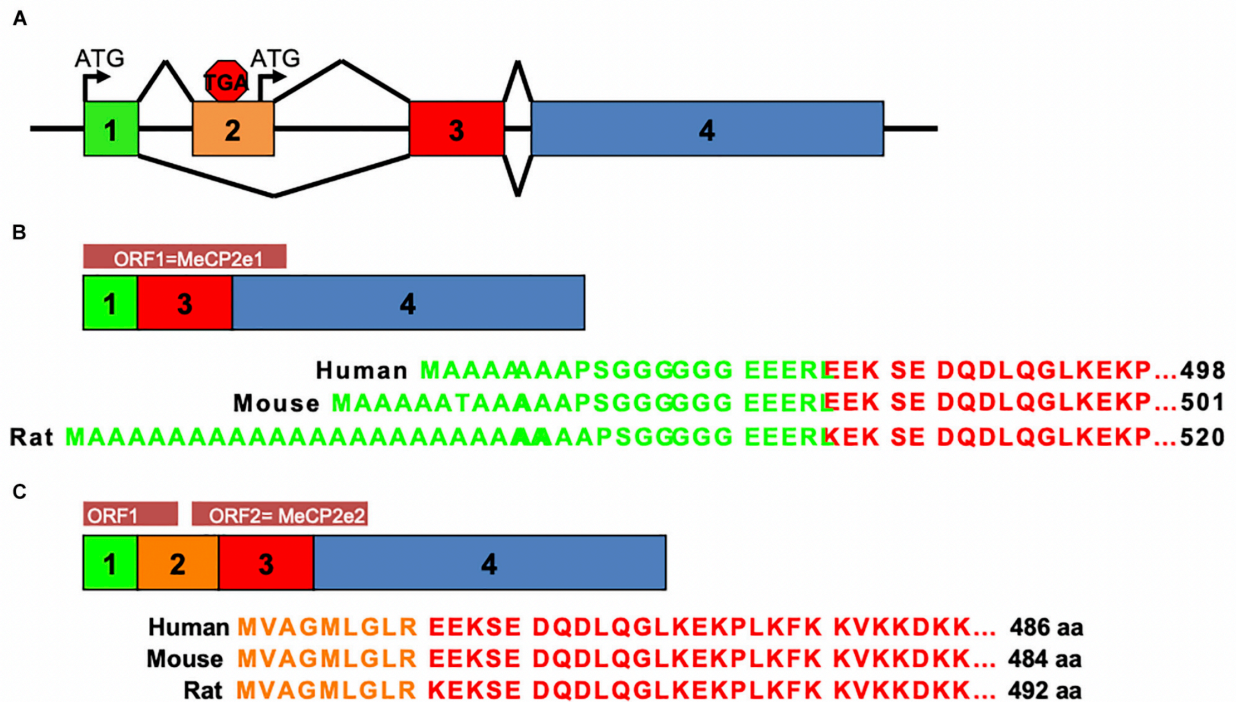


Figure 2: MeCP2 protein isoforms. MeCP2e1 and MeCP2e2 protein isoforms are produced from alternative splicing events. The MeCP2e1 isoform is translated from splicing of exon 1

to exons 3 and 4 with the translational start ATG in exon 1. The MeCP2e2 protein isoform is produced from a transcript from all four exons with a translational start ATG in exon 2. Interestingly, this transcript may encode a small peptide from an ORF in exons 1 and 2. (A) *MECP2/Mecp2* exon splicing. (B) MeCP2-e1 isoform. (C) MeCP2-e2 isoform.

Most importantly, the results from Yasui and colleagues found that *Mecp2e1^{-/-}* males had impaired motor function, exhibited apraxia-like limb clasping with altered anxiety behavior, and premature death (Yasui et al., 2014), similar to Bird *Mecp2* deletion male mice (Guy et al., 2001). These MeCP2e1 deficient mice accurately model Rett syndrome, as heterozygous *Mecp2e1^{-/+}* females exhibited significant motor impairment and altered body composition (Yasui et al., 2014; Vogel Ciernia et al., 2017). Interestingly, both *Mecp2e1^{-/-}* males and *Mecp2e1^{-/+}* females had elevated body fat accumulation correlating with reduced energy expenditure in early adulthood (Vogel Ciernia et al., 2018).

Additional evidence that MeCP2e1 has non-overlapping function with MeCP2e2 is the finding by the Rastegar lab that MeCP2e1 is expressed in the brain prior to MeCP2e2 at higher levels and with a more consistent distribution pattern between brain regions (Olson et al., 2014). Consistent with the mouse studies, the Ellis lab found that human MeCP2e1 deficient iPS-derived neurons recapitulate synaptic current, soma size, membrane potential, and maturation defects common to Rett syndrome, further establishing the necessity of MeCP2e1 in neuronal development (Djuric et al., 2015). MeCP2e1 and MeCP2e2 were also found by the Ausio lab to have differential DNA binding kinetics, distinct co-factor associations, and preferred different DNA binding motifs (Martínez De Paz et al., 2019). Perhaps the most compelling evidence that MeCP2e1 has an essential function that differs from MeCP2e2 is the finding that MeCP2e1 is the evolutionarily older isoform, as orthologs have been identified in vertebrates back to bony fish (Osteichthyes) and amphibians (Mnatzakanian et al., 2004).

1.5.3 Glial Cells Are Critical for Rett Syndrome Pathology

Another contentious issue in the MeCP2/Rett field is the contribution of non-neuronal cells to the disease process, despite considerable evidence that *MECP2* transcripts are present at high levels

in virtually all tissues (**Figure 1**). Although cell sorting experiments show that MeCP2 is present at near octomer levels in neurons, MeCP2 is present in glia, albeit at much lower levels (Skene et al., 2010). Direct evidence for the role of non-neural cells in Rett pathology was provided by research performed by the Jin lab at UC Davis, where it was found that astrocytes from Bird *Mecp2*^{-/-} null male mice were abnormal and produced underdeveloped wild-type neurons compared with wild-type astrocytes in a co-culture system (Maezawa et al., 2009). A subsequent study by the Jin lab revealed that Bird *Mecp2*^{-/-} null male microglia damaged adjacent neurons in a non-cell autonomous manner by releasing glutamate (Maezawa and Jin, 2010). The relevance of cell autonomous vs. non-cell autonomous effects in the *Mecp2* mutant brain is discussed further in the next section.

In support of these studies, a paper by the Mandel lab described that re-expression of *Mecp2* solely in astrocytes was able to reduce motor defects, breathing irregularity as well as prolong life in Bird *Mecp2*^{-/-} null male mice (Lioy et al., 2011). Surprisingly, microglia, which comprise a tiny percentage of brain cells, were found to be important for Rett pathology when it was shown that replacement of *Mecp2* null microglia with wild-type microglia in Bird *Mecp2*^{-/-} null male brain improved life span, normalized breathing, and reduced motor impairment (Derecki et al., 2012). Although these results were not independently reproduced, a recent study by the Stevens lab established that microglial engulfment of synapses damages neural circuits in *Mecp2*^{-/-} null male mice (Schafer et al., 2016). To further address the specific contribution of astrocytes to Rett phenotypes, Qiang Chang from the University of Wisconsin Madison derived wild-type and *MECP2* mutant iPS astrocytes and examined their function in culture with neurons. They found that *MECP2* mutant astrocytes have excess calcium stores that correlated with elevated NMDA receptor expression in neighboring neurons and hyper neuronal network excitability (Dong et al., 2018). Therefore, the authors conclude that astrocytes act both cell-autonomously and non-cell-autonomously to mediate Rett-like pathology (Dong et al., 2018).

An extensive review on the function of MeCP2 in glia has recently become available (Kahanovitch et al., 2019).

1.5.4 *MECP2/Mecp2* Has Both Cell Autonomous and Non-cell Autonomous Effects

The vast majority of Rett patients are female *MECP2* heterozygotes and are thus cellular mosaics for wild-type and mutant *MECP2* expressing cells. It was first reported by the LaSalle lab that Bird *Mecp2* null neurons affected the expression of adjacent wild-type neurons in *Mecp2*^{-/+} heterozygous female brains (Braunschweig et al., 2004). This finding led to the “bad neighborhood” or non-cell autonomous effects hypothesis of *Mecp2* deficiency, where defects in *Mecp2* mutant cells affect the function of local wild-type cells in the brain or other tissues (Braunschweig et al., 2004). This “bad neighborhood” becomes important when examining disease processes in female but not male mice, where all cells are deficient in *Mecp2* expression. Subsequent studies by the Gail Mandel lab showed that MeCP2 deficient glia impair neuronal function in a non-cell autonomous manner in vitro (Ballas et al., 2009). As mentioned in the previous section, a similar in vitro culture system showed that *Mecp2*^{-/y} null astrocytes were able to impair the growth of wild-type neuronal dendrites through soluble factors (Maezawa et al., 2009). To study these effects in vivo by transplanting either *Mecp2* mutant or wild-type neuroblasts into wild-type brain, Kishi and Macklis found that MeCP2 functions largely cell autonomously in the development of the cortex but that non-cell autonomous effects on neuronal function exist (Kishi and Macklis, 2010). In the most definitive study to date, the Zhou lab examined *Mecp2* mutant tagged and wild-type cells from female brains and found that non-cell autonomous effects were greater in magnitude than cell autonomous effects and tended to affect cell-to-cell signaling and phosphorylation (Johnson et al., 2017). These results were consistent with subsequent single-cell RNA-sequencing (sc-RNAseq) analysis of neurons from *Mecp2*^{-/+} deficient

female brains, suggesting that MeCP2 can affect gene expression in a non-cell autonomous manner in different neuronal cell types (Renthal et al., 2018).

1.5.5 MeCP2 Preferentially Regulates Long Genes Over Short Genes

The hypothesis that MeCP2 preferentially regulates long genes was first proposed by Sacha Nelson in 2014 (Sugino et al., 2014). This report was supported by data published from the Greenberg lab in 2015 (Gabel et al., 2015). This hypothesis was further tested in an important study mentioned previously by the Zhou lab, which found a trend toward long genes being more severely affected by MeCP2 deficiency than short genes (Johnson et al., 2017). However, these findings were refuted by a re-analysis of the data and transcriptional analysis using alternative methods, which together suggested that traditional PCR methods bias the RNA-seq results toward long genes (Raman et al., 2018). The controversy over the regulation of long genes is discussed further in a review by Connolly and Zhou (2019).

1.5.6 At What Developmental Stage Do MeCP2 Defects Impair Development

A lack of a developmental time course of how MeCP2 defects are manifested impedes Rett research. Although there is an abundance of data about MeCP2 defects in adult mice, the earliest stages of development have not been well characterized, aside from a 2003 study which found MeCP2 positive cells in E14 rat cortex (Jung et al., 2003). To address this gap, the Landsberger lab analyzed neurons from embryonic Bird *Mecp2*^{-/-} null male cortex (Bedogni et al., 2016). They found that Bird *Mecp2*^{-/-} null embryonic neurons have altered gene expression, altered morphology, reduced calcium flux, and mobility compared to wild-type neurons (Bedogni et al., 2016). Independent evidence suggests that there are deficiencies in *Mecp2* transcripts and MeCP2 protein in the mouse cortex as early as E14 (Zachariah et al., 2012). However, one limitation of these molecular studies of early MeCP2 expression in mice is that they employed conventional methods in bulk tissue such as qRT-PCR and Western blot and that the investigators studied Bird *Mecp2*^{-/-} null male mice, which do not

accurately model Rett syndrome in *MECP2* heterozygous females. To date, there has been only one study of MeCP2 embryonic function at single-cell resolution in disease-relevant Bird *Mecp2*^{-/+} deficient mouse brain by the Greenberg group. However, only adult Bird *Mecp2*^{-/+} deficient deletion female mice and adult human Rett brains were analyzed (Renthal et al., 2018). Yet it is known that Rett syndrome girls experience a postnatal developmental regression at about 6–18 months of age (Zoghbi, 2005). Thus, the molecular changes during this critical early time are still not well defined. While it is clear that MeCP2 function is critical for normal neurological function, the molecular phenotypes in distinct brain cell types have not been investigated over the full-time course of development in mouse models. The Rett field is lacking an extensive, longitudinal study in mosaic *Mecp2*^{-/+} deficient mice investigating the mechanisms and molecular pathways that are impacted by perturbed *Mecp2*/MeCP2 expression throughout disease progression.

1.5.7 MeCP2 Is an RNA Binding Protein

Coding RNAs and non-coding RNAs such as long non-coding RNAs, small nucleolar RNAs, and micro RNAs establish a diverse set of functions due to their direct interactions with RNA-binding proteins (RBPs). In one of the few studies on this topic, it was found that MeCP2 binds with high affinity to mRNA and siRNA outside of the MBD in vitro (Jeffery and Nakielnny, 2004). MeCP2 was also found to associate with long non-coding RNAs and imprinted genes in mouse brain extracts (Maxwell et al., 2013). To examine MeCP2 interaction with RNA in vivo, MeCP2 RNA-immunoprecipitation of small RNAs was performed in the lab of Assam El-Osta, who found that MeCP2 bound to specific RNAs including micro RNAs (Khan et al., 2017). These studies investigating MeCP2 interactions with RNA were performed in a targeted manner; therefore, an unbiased, genome-wide screen of RNA binding would provide a significant advance in understanding this MeCP2 activity.

1.6 Recent Developments in the MeCP2 Field

1.6.1 MeCP2 Enables Liquid Phase Separation Events in Nuclear

Compartmentalization

In the last 2 years, the biologic functions of MeCP2 have expanded appreciably. One concept that has emerged is the effect that liquid-liquid phase separations (LLPS) have on the formation of sub-nuclear compartments (Alberti et al., 2019). In their review, Alberti, Gladfelter, and Mittag explain that the nucleus of eukaryotic cells consists of membrane-less structures such as nucleoli and heterochromatin that form by LLPS through the activity of proteins and nucleic acids that condense into a dense phase (heterochromatin) and a loose phase depending on conditions such as molecular concentration, salt concentration, and pH (Alberti et al., 2019). Thus, the three near simultaneous reports that MeCP2 is involved with LLPS DNA phase separation in the cell nucleus were of great relevance to the field. In the first report, Wang and colleagues found that MeCP2 was able to condense chromatin constructs in vitro and that DNA methylation enhanced this effect (Wang et al., 2020). Furthermore, Wang et al. found that mutant MeCP2 proteins from Rett patients were defective in chromatin condensation and LLPS (Wang et al., 2020). Results from Fan and colleagues confirmed that wild-type MeCP2 was able to form condensates with DNA in vitro (Fan et al., 2020). Finally, research published by Richard Young and Rudolph Jaenisch at MIT also showed that MeCP2 condensed chromatin in vitro and MeCP2 mutant forms were defective in LLPS activity (Li et al., 2020). What set apart the Young and Jaenisch study from the other studies is the finding that the intrinsically disordered regions (IDR) of MeCP2 mediate the phase separation activity in the nucleus (Li et al., 2020). The concept of MeCP2 as a nuclear organizer via LLPS is reminiscent of early reports that a MeCP2 homolog, ARBP, functions as a component of the nuclear matrix as mentioned previously (Weitzel et al., 1997).

1.6.2 MeCP2 Functions with DNMT3A/Dnmt3a to Regulate Gene

Expression

MeCP2 and DNMT3A, a de novo DNA methyltransferase, have long been linked by the fact that the MeCP2 MBD preferentially binds to CpG methylated DNA (mCG). Recently, it was found that there is a direct molecular link between the two molecules as the MeCP2 TRD binds to the DNMT3A ADD domain. This interaction inhibits DNMT3A activity, and this inhibition is countered by DNMT3A interaction with H3K4 through the ADD domain in vitro (Rajavelu et al., 2018). Similar results were shown by a collaboration between the Joe Ecker and Huda Zoghbi labs that examined the in vivo effects of MeCP2 and DNMT3A interaction. It had been previously established by the Ecker lab that high levels of DNA methylation at CH (mCH) sites (H = Adenosine, Cytosine, and Thymidine) in neurons are due to the activity of DNMT3A (Lister et al., 2009). Therefore, Zoghbi and Ecker examined the hypothesis that as MeCP2 is the sole reader of methyl mCH and that DNMT3A is the sole writer of methyl CH, deletion of either factor in neurons would have similar effects on gene transcription. Surprisingly, loss of DNMT3A in neurons affected transcription of significantly more genes than the loss of MeCP2, highlighting the importance of mCH to neurologic gene regulation and function (Lavery et al., 2020). The relationship between mCH, mCG, and MeCP2 during development is explored further in a review by Lavery and Zoghbi (2019). These advances also underscore the concept that much of MeCP2 function is mediated through co-factor association.

1.6.3 MeCP2 May Function in DNA Repair Processes

The first evidence to suggest that MeCP2 could function in DNA repair processes was reported in 2005. Valinluck and colleagues found that MeCP2 binds with high affinity to DNA containing halogenated pyrimidines, a modification that can result from inflammation (Valinluck et al., 2005). More recently, it was found that neural stem cells from Bird *Mecp2*^{-/+} deficient female mice

are prone to early senescence and have elevated rates of cell death when exposed to DNA-damaging H₂O₂, UV light, and doxorubicin (Alessio et al., 2018). However, the most convincing evidence for the role of MeCP2 in DNA repair comes from an unbiased N-ethyl-N-nitrosourea (ENU) suppressor screen in Bird *Mecp2*^{-/-} null male mice (Enikanolaiye et al., 2020). Out of 2,498 males born to ENU-treated wild-type males mated with *Mecp2*^{-/+} females, 96 males had ameliorated neurologic symptoms from which 32 genes that suppressed the *Mecp2* null phenotype were identified. Of these 32 genes, Tet1, Birc6, and Spin1 function in the DNA damage response, while Rbbp8, Rad50, Fan1, Brca1, and Brca2 encode factors involved in DNA double-strand break repair (Enikanolaiye et al., 2020).

1.6.4 Potential Treatments Under Development for Rett Syndrome

The remarkable reversal of Rett-like phenotypes in Bird *Mecp2*^{-/-} null male mice provides a basis for clinical therapies that seek to restore expression of *MECP2* (Guy et al., 2007). To this aim, the potential of adeno-associated viral (AAV) vectors to deliver MeCP2 to deficient cells in a whole animal model was shown independently in 2013 by two groups. One group, led by Stuart Cobb, demonstrated that AAV9/*MECP2* delivered intravenously in presymptomatic Bird *Mecp2*^{-/-} null male mice extended survival modestly (Gadalla et al., 2013). Gail Mandel led a second study showing that intravenous AAV9/MeCP2e1 vector reduced motor defects in Bird *Mecp2*^{-/+} deficient female mice, which is particularly relevant to Rett syndrome as 95% of Rett patients are female (Garg et al., 2013). In 2017, following these early proofs of principle, advances in AAV9/*MECP2* gene therapy were reported by the combined efforts of Stuart Cobb and Steven J. Gray. The authors reported improved cell transduction with a reduced titer of a second-generation AAV9/hMeCP2 virus, reduced liver toxicity, and prolonged survival in Bird *Mecp2*^{-/-} null mice by delivery through the cerebrospinal fluid (CSF) or directly into the brain (Gadalla et al., 2017; Sinnett et al., 2017). These studies represent significant improvements in *MECP2* gene therapy tools but should be performed in Bird *Mecp2*^{-/+}

deficient heterozygous female deletion mice for greater relevance to Rett. Furthermore, the risks of treating human children with AAV vectors cannot be fully mitigated.

Alternatively, site-directed RNA editing is being developed as a potential Rett syndrome therapy. RNA editing strategies leverage an Adenosine Deaminase Acting on RNA (ADAR2) RNA editing factor with a guide RNA to repair *Mecp2*/*MECP2* point mutations in RNA. The first report on *Mecp2* RNA editing was published by the Gail Mandel lab in 2017, where an exogenous, modified ADAR2 and guide successfully edited 72% of G to A mutations in RNA from *Mecp2*R106Q mouse neurons (Sinnamon et al., 2017). In 2020, Sinnamon and colleagues successfully edited up to 50% of mRNA in neurons in developing brains of *Mecp2* mutant mice using a similar approach (Sinnamon et al., 2020). The potential of RNA editing as a Rett therapy is limited by the fact that, although endogenous RNA editing proteins such as ADAR1 exist naturally in the brain, current approaches express a modified RNA editing protein delivered by an AAV virus, and these ADAR2-like proteins primarily edit only G to A mutations (Sinnamon et al., 2017, 2020).

Despite the excitement for *MECP2* gene replacement and RNA editing strategies, other therapeutic methods are in development. One strategy explored by the Zoghbi lab in Rett model mice is deep brain stimulation. Electrode-based stimulation of symptomatic *Mecp2*^{-/+} female mice normalized learning and memory defects (Hao et al., 2015; Pohodich et al., 2018). Another promising area for potential Rett therapies is the repurposing of existing FDA-approved compounds. For example, in 2020, it was found that the diabetes drug metformin reduces mitochondrial defects and damage from oxidative stress in *Mecp2* 308^{-/+} truncation female mice (Zuliani et al., 2020). Finally, a recently developed compound for Alzheimer's disease reduced motor defects, learning deficits, and breathing abnormalities in *Mecp2*^{-/+} female mice with minimal side effects (Kaufmann et al., 2019). Until effective and safe *MECP2* gene replacement and/or RNA editing strategies are available, existing or novel chemical compounds could be used to reduce the most severe symptoms in Rett patients.

1.7 The Interaction Between Environmental Factors and MeCP2

RTT babies are born seemingly normal; however, in late infancy, they experience a regression in cognitive and motor functions (Maheshwari N et al. 2018). The variability in severity and timing of progression in RTT suggests that *MECP2* mutations may interact with environmental factors. While it is clear that MeCP2 function is very important for normal neurological function, it is unclear how molecular signatures of disease progression in distinct brain cell types interact with environmental toxicants such as persistent organic pollutants (POPs). These are a group of toxic chemicals that persist in the environment and accumulate in the food chain, posing risks to human health and the environment. Polychlorinated pollutants are a subclass of these chemicals, and some of the most notable include: Polychlorinated Biphenyls (PCBs), Dioxins (e.g., 2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD)), Furans (e.g., 2,3,4,7,8-Pentachlorodibenzofuran (PeCDF)), Chlordane (also known as Kepone), Hexachlorobenzene (HCB), DDT (Dichlorodiphenyltrichloroethane). PCBs are POPs that are identified as probable environmental risk factors for neurodevelopmental disorders (NDD) (Faroon O et al. 2016, Eubig PA et al. 2010). Evidence from animal studies provide evidence that PCBs are developmental neurotoxicants making them a major concern for human health (Berghuis SA et al. 2015). Although, there has been studies showing the effects of POPs in ASD, further studies need to be performed to understand the interactions between POPs and Rett syndrome specifically.

1.8 Discussion

Currently, there is debate as to whether patients with overlapping phenotypes have Rett syndrome or some other disease. For example, patients with mutations in *CDKL5* were considered to have a severe variant of Rett, however, this was later categorized as a separate disease (Fehr et al., 2013). Patients with *FOXP1* mutations are currently classified with Rett. However, as these rare

FOXP1 mutation patients present with disease from birth and have other symptoms that do not overlap with Rett, a proposal to classify these patients as a separate syndrome was recently submitted (Cutri-French et al., 2020). Therefore, with few exceptions, any discussion about Rett syndrome phenotypes is about how defective MECP2/Mecp2 expression presents in patients and animal models and manifests as a neurodevelopmental disease.

At this time, it is still unclear how MECP2 expression defects in non-neural tissues contribute to Rett phenotypes. The answer may be complex as illustrated by a report from the Landsberger lab showing that male Bird Mecp2 deletion mice have abnormal muscle tissue while deletion of the Bird Mecp2 allele only in muscle resulted in normal muscle tissue (Conti et al., 2015). The conclusion from this study is that non-cell-autonomous signaling to the muscle by defective neurons in the Bird Mecp2 deletion mice contributes to the muscle defect (Conti et al., 2015). So, there is the possibility that although MECP2 transcripts are abundant in all tissues and cell types, extracellular neuronal signaling could be responsible for some, if not all of the defects.

The recent identification of MeCP2 binding beyond methylated CpG sites has greatly expanded the range of potential functions of the protein. These advances have been driven by the discovery of additional nucleotide modifications in mammalian brain. Yet many questions remain. For example, does MeCP2 binding to 5-hmc (Kriaucionis and Heintz, 2009) in cerebellum have a significant neurologic function (Mellén et al., 2012)? Do MeCP2 isoforms have different 5-hmc binding activity? Some data suggests that MeCP2 protects 5mc from TET1 conversion to 5 hmc (Szulwach et al., 2011) but more study is needed to resolve this important question. The 2014 discovery of DNMT3A dependent CH methylation in mouse brain by the Song lab (Guo et al., 2014) and the description of MeCP2 binding to methyl CpA sites in brain in vivo (Gabel et al., 2015) has been suggested to be key to the pathogenesis of Rett (Lavery et al., 2020). Here again additional research is needed to clearly establish the effect of MeCP2 binding outside of methyl CpG sites.

Another evolving area of research is how MeCP2 functions as a transcriptional repressor and activator. The model of MeCP2 as a transcriptional repressor of single copy genes was developed over twenty years ago (Nan et al., 1997; Jones et al., 1998). Since that time, it was found that MeCP2 binds to partially CpG methylated promoters of active genes in neurons (Yasui et al., 2007) and can act as a transcriptional activator by recruitment of CREB in neurons (Chahrour et al., 2008). The latest evidence suggests that MeCP2 modulates gene transcription depending on co-factor interaction. Future gene expression analyses should focus on the differential activity of the MeCP2-e1 and MeCP2-e2 isoforms.

The role of glial cell defects in Rett syndrome pathology is still being described. However, the very fact that glial cells express MECP2/Mecp2 and are defective in Rett patients and Mecp2 null and deficient mice is relevant to neuronal disease phenotypes. As Rett females and Mecp2-null deficient heterozygous female mice are mosaic for normal and mutant cells, non-cell autonomous effects are clearly contributing to neuronal dysfunction (Johnson et al., 2017). Analysis of neurons from mosaic Bird Mecp2-null deficient female mouse brain and human Rett brain suggest that MeCP2 preferentially represses methylated long genes (Sugino et al., 2014), in a cell autonomous manner in neurons, although non-cell autonomous effects on short genes cannot be excluded (Renthal et al., 2018). As recent study concluded that MeCP2 bias toward long gene regulation may be an artifact of PCR based transcript detection (Raman et al., 2018), single cell analyses using alternatives to current RNA-seq methods in brain cell types will be necessary to resolve the role of non-cell autonomous effects in brains mosaic for wild-type and Mecp2 mutant neurons and glia.

Although it is widely assumed that Rett patients are phenotypically normal at birth this conclusion is based on limited behavioral observation. As there is little incentive and ethical limitations preclude studying apparently normal human infants, a full developmental time course of the disease is needed in Rett model Mecp2-null heterozygous mice that more accurately model Rett MECP2

mutations such as Mecp2-e1 mutant mice (Yasui et al., 2014). Surprisingly, apart from a study that revealed defects in Bird Mecp2-null male embryonic cells, little analysis has been done of early developmental time points (Bedogni et al., 2016). Clearly, a developmental time analyses of both Mecp2-null females and Mecp2-null males at embryonic, pre-disease, early disease, and late disease stages is needed to identify pathological mechanisms.

There are aspects of MeCP2 function that are oddly neglected. For example, surprisingly, few studies have directly examined the ability of MeCP2 to bind to RNA. The studies that do show a strong in vitro interaction of MeCP2 with mRNA and siRNA via the RG domain between the MBD and TRD (Jeffery and Nakielnny, 2004) and in vivo interaction with miRNAs (Khan et al., 2017). A logical progression of these studies would be to examine RNA binding ability of common Rett mutant MeCP2 proteins such as T158M in vitro and in vivo. The Rett field is lacking studies using more unbiased high-throughput methods such as targets of RNA-binding proteins identified by editing (TRIBE) (McMahon et al., 2016) which fuses MeCP2 to ADAR, an enzyme that modifies RNA where MeCP2 binds. An additional advantage of this approach is that it can be used for in vivo studies.

Finally, unexpected discoveries continue to revitalize the study of MeCP2. Recent cell biology studies reveal that compartmentalization of the nucleus by liquid-liquid phase separation forms critical regulatory compartments such as the nucleolus and heterochromatin (reviewed in Alberti et al., 2019). Now it appears that MeCP2 plays a key role in the formation of nuclear heterochromatin domains according to new in vitro studies (Fan et al., 2020; Wang et al., 2020) and that the MeCP2 intrinsically disorder region plays a key role in this process (Li et al., 2020). These results build on the extensive study of how MeCP2 organizes chromatin in neuronal cells and may provide key insights into Rett syndrome defects.

The most interesting advances in MeCP2 research concern Rett therapies, as restoration of MeCP2 in mouse brain led to disease reversal in a mouse model (Guy et al., 2007). Current MECP2

gene therapy for Rett patients has great potential and some risks. The latest gene therapy treatment of Bird Mecp2-null heterozygous mice with AAV9/MECP2 was able to normalize breathing but high viral doses produced severe liver toxicity and death in some animals (Matagne et al., 2020). While RNA editing strategies for repairing mutant MECP2 have great potential as a therapy, delivery of editing proteins to the brain is also as problematic as delivering functional MeCP2. Therefore, considerable time and effort will be needed to refine MECP2 gene replacement and RNA editing based therapies. Until then, alternative treatments are needed to treat Rett.

One potential therapy for Rett is deep brain stimulation (DSB) which has been approved by the FDA for epilepsy. Nonetheless, DSB also has its risks as electrodes are implanted into the brain. Existing drugs approved by the FDA for other diseases are being re-directed to treat Rett symptoms. Clinical trials of IGF-1 and IGF-1 analogs that have been approved for short stature are ongoing for Rett and offer some hope of reducing symptoms (Khwaja et al., 2014). Another compound, ANAVEX 2-73 which targets the sigma-1 receptor that affects learning, memory, and neuronal development and is in phase II clinical trials for Alzheimer's disease is also scheduled for Phase I trials in Rett girls (Kaufmann et al., 2019). In some cases, effective treatment for a specific disease have been employed without fully understanding the mechanisms of the disease process. While this may be the case with Rett syndrome, that should not preclude the continuing research on the evolving molecular function of MeCP2 in neurons and other cell types.

1.9 References

Adams, V. H., McBryant, S. J., Wade, P. A., Woodcock, C. L., & Hansen, J. C. (2007). Intrinsic disorder and autonomous domain function in the multifunctional nuclear protein, MeCP2. *Journal of Biological Chemistry*, 282, 15057–15064. <https://doi.org/10.1074/jbc.M700855200>

- Alberti, S., Gladfelter, A., & Mittag, T. (2019). Considerations and challenges in studying liquid-liquid phase separation and biomolecular condensates. *Cell*, 176, 419–434. <https://doi.org/10.1016/j.cell.2018.12.035>
- Alessio, N., Riccitiello, F., Squillaro, T., Capasso, S., Del Gaudio, S., Di Bernardo, G., et al. (2018). Neural stem cells from a mouse model of Rett syndrome are prone to senescence, show reduced capacity to cope with genotoxic stress, and are impaired in the differentiation process. *Experimental & Molecular Medicine*, 50(1). <https://doi.org/10.1038/s12276-017-0005-x>
- Amir, R. E., Van Den Veyver, I. B., Wan, M., Tran, C. Q., Francke, U., & Zoghbi, H. Y. (1999). Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nature Genetics*, 23, 185–188. <https://doi.org/10.1038/13810>
- Baker, S. A., Chen, L., Wilkins, A. D., Yu, P., Lichtarge, O., & Zoghbi, H. Y. (2013). An AT-hook domain in MeCP2 determines the clinical course of Rett syndrome and related disorders. *Cell*, 152, 984–996. <https://doi.org/10.1016/j.cell.2013.01.038>
- Ballas, N., Lioy, D. T., Grunseich, C., & Mandel, G. (2009). Non-cell autonomous influence of MeCP2-deficient glia on neuronal dendritic morphology. *Nature Neuroscience*, 12, 311–317. <https://doi.org/10.1038/nn.2275>
- Baubec, T., Ivánek, R., Lienert, F., & Schübeler, D. (2013). Methylation-dependent and -independent genomic targeting principles of the MBD protein family. *Cell*, 153, 480–492. <https://doi.org/10.1016/j.cell.2013.03.011>
- Bedogni, F., Gigli, C. C., Pozzi, D., Rossi, R. L., Scaramuzza, L., Rossetti, G., et al. (2016). Defects during Mecp2 null embryonic cortex development precede the onset of overt neurological symptoms. *Cerebral Cortex*, 26, 2517–2529. <https://doi.org/10.1093/cercor/bhv078>

- Bissonnette, J. M., Schaevitz, L. R., Knopp, S. J., & Zhou, Z. (2014). Respiratory phenotypes are distinctly affected in mice with common Rett syndrome mutations MECP2 T158A and R168X. *Neuroscience*, 267, 166–176. <https://doi.org/10.1016/j.neuroscience.2014.02.043>
- Borghi, E., Borgo, F., Severgnini, M., Savini, M. N., Casiraghi, M. C., & Vignoli, A. (2017). Rett syndrome: A focus on gut microbiota. *International Journal of Molecular Sciences*, 18(2). <https://doi.org/10.3390/ijms18020344>
- Braunschweig, D., Simcox, T., Samaco, R. C., & LaSalle, J. M. (2004). X-chromosome inactivation ratios affect wild-type MeCP2 expression within mosaic Rett syndrome and *Mecp2*^{-/+} mouse brain. *Human Molecular Genetics*, 13, 1275–1286. <https://doi.org/10.1093/hmg/ddh142>
- Brero, A., Easwaran, H. P., Nowak, D., Grunewald, I., Cremer, T., Leonhardt, H., et al. (2005). Methyl CpG-binding proteins induce large-scale chromatin reorganization during terminal differentiation. *Journal of Cell Biology*, 169, 733–743. <https://doi.org/10.1083/jcb.200502062>
- Bu, Q., Wang, A., Hamzah, H., Waldman, A., Jiang, K., Dong, Q., et al. (2017). CREB signaling is involved in Rett syndrome pathogenesis. *Journal of Neuroscience*, 37, 3671–3685. <https://doi.org/10.1523/JNEUROSCI.3735-16.2017>
- Buchmuller, B. C., Kosel, B., & Summerer, D. (2020). Complete profiling of methyl-CpG-binding domains for combinations of cytosine modifications at CpG dinucleotides reveals differential read-out in normal and Rett-associated states. *Scientific Reports*, 10, 4053. <https://doi.org/10.1038/s41598-020-61030-1>
- Buchovecky, C. M., Turley, S. D., Brown, H. M., Kyle, S. M., McDonald, J. G., Liu, B., et al. (2013). A suppressor screen in *Mecp2* mutant mice implicates cholesterol metabolism in Rett syndrome. *Nature Genetics*, 45, 1013–1020. <https://doi.org/10.1038/ng.2714>

- Chahrour, M., Sung, Y. J., Shaw, C., Zhou, X., Wong, S. T. C., Qin, J., et al. (2008). MeCP2, a key contributor to neurological disease, activates and represses transcription. *Science*, 320, 1224–1229. <https://doi.org/10.1126/science.1153252>
- Chao, H. T., Chen, H., Samaco, R. C., Xue, M., Chahrour, M., Yoo, J., et al. (2010). Dysfunction in GABA signalling mediates autism-like stereotypies and Rett syndrome phenotypes. *Nature*, 468, 263–269. <https://doi.org/10.1038/nature09582>
- Chen, R. Z., Akbarian, S., Tudor, M., & Jaenisch, R. (2001). Deficiency of methyl-CpG binding protein-2 in CNS neurons results in a Rett-like phenotype in mice. *Nature Genetics*, 27, 327–331. <https://doi.org/10.1038/85906>
- Chhatbar, K., Cholewa-Waclaw, J., Shah, R., Bird, A., & Sanguinetti, G. (2020). Quantitative analysis questions the role of MeCP2 as a global regulator of alternative splicing. *PLoS Genetics*, 16(6), e1009087. <https://doi.org/10.1371/journal.pgen.1009087>
- Clemens, A. W., Wu, D. Y., Moore, J. R., Christian, D. L., Zhao, G., & Gabel, H. W. (2020). MeCP2 represses enhancers through chromosome topology-associated DNA methylation. *Molecular Cell*, 77(2), 279–293.e8. doi: 10.1016/j.molcel.2019.10.033
- Collins, A. L., Levenson, J. M., Vilaythong, A. P., Richman, A., Armstrong, A. L., Noebels, J. L., et al. (2004). Mild overexpression of MeCP2 causes a progressive neurological disorder in mice. *Human Molecular Genetics*, 13(24), 2679–2689. doi: 10.1093/hmg/ddh282
- Connelly, J. C., Cholewa, J., Webb, S., Steccanella, V., Waclaw, B., & Bird, A. (2020). Absence of MeCP2 binding to non-methylated GT sequences in vivo. *Nucleic Acids Research*, 48(7), 3542–3552. doi: 10.1093/nar/gkaa102
- Connolly, D. R., & Zhou, Z. (2019). Genomic insights into MeCP2 function: a role for the maintenance of chromatin architecture. *Current Opinion in Neurobiology*, 59, 174–179. doi: 10.1016/j.conb.2019.07.002

- Conti, V., Gandaglia, A., Galli, F., Tirone, M., Bellini, E., Campana, L., et al. (2015). MeCP2 affects skeletal muscle growth and morphology through non-cell-autonomous mechanisms. *PLoS One*, 10(7), e0130183. doi: 10.1371/journal.pone.0130183
- Cutri-French, C., Armstrong, D., Saby, J., Gorman, C., Lane, J., Fu, C., et al. (2020). Comparison of core features in four developmental encephalopathies in the Rett natural history study. *Annals of Neurology*, 88(3), 396–406. doi: 10.1002/ana.25797
- Del Gaudio, D., Fang, P., Scaglia, F., Ward, P. A., Craigen, W. J., Glaze, D. G., et al. (2006). Increased MECP2 gene copy number as the result of genomic duplication in neurodevelopmentally delayed males. *Genetics in Medicine*, 8(12), 784–792. doi: 10.1097/01.gim.0000250502.28516.3c
- Derecki, N. C., Cronk, J. C., Lu, Z., Xu, E., Abbott, S. B. G., Guyenet, P. G., et al. (2012). Wild-type microglia arrest pathology in a mouse model of Rett syndrome. *Nature*, 484(7392), 105–109. doi: 10.1038/nature10907
- Djuric, U., Cheung, A. Y. L., Zhang, W., Mok, R. S., Lai, W., Pickna, A., et al. (2015). MECP2e1 isoform mutation affects the form and function of neurons derived from Rett syndrome patient iPS cells. *Neurobiology of Disease*, 76, 37–45. doi: 10.1016/j.nbd.2015.01.001
- Dong, Q., Liu, Q., Li, R., Wang, A., Bu, Q., Wang, K. H., et al. (2018). Mechanism and consequence of abnormal calcium homeostasis in Rett syndrome astrocytes. *eLife*, 7, e33417. doi: 10.7554/eLife.33417
- Enikanolaiye, A., Ruston, J., Zeng, R., Taylor, C., Schrock, M., Buchovecky, C. M., et al. (2020). Suppressor mutations in Mecp2-null mice implicate the DNA damage response in Rett syndrome pathology. *Genome Research*, 30(5), 540–552. doi: 10.1101/gr.258400.119
- Fan, C., Zhang, H., Fu, L., Li, Y., Du, Y., Qiu, Z., et al. (2020). Rett mutations attenuate phase separation of MeCP2. *Cell Discovery*, 6, 38. doi: 10.1038/s41421-020-0172-0

- Fehr, S., Wilson, M., Downs, J., Williams, S., Murgia, A., Sartori, S., et al. (2013). The CDKL5 disorder is an independent clinical entity associated with early-onset encephalopathy. *European Journal of Human Genetics*, 21(3), 266–273. doi: 10.1038/ejhg.2012.156
- Free, A., Wakefield, R. I. D., Smith, B. O., Dryden, D. T. F., Barlow, P. N., & Bird, A. P. (2001). DNA Recognition by the Methyl-CpG binding domain of MeCP2. *Journal of Biological Chemistry*, 276(6), 3353–3360. doi: 10.1074/jbc.M007224200
- Fyffe, S. L., Neul, J. L., Samaco, R. C., Chao, H. T., Ben-Shachar, S., Moretti, P., et al. (2008). Deletion of *Mecp2* in *Sim1*-expressing neurons reveals a critical role for MeCP2 in feeding behavior, aggression, and the response to stress. *Neuron*, 59(6), 947–958. doi: 10.1016/j.neuron.2008.07.030
- Gabel, H. W., Kinde, B., Stroud, H., Gilbert, C. S., Harmin, D. A., Kastan, N. R., et al. (2015). Disruption of DNA-methylation-dependent long gene repression in Rett syndrome. *Nature*, 522(7555), 89–93. doi: 10.1038/nature14319
- Gadalla, K. K., Bailey, M. E., Spike, R. C., Ross, P. D., Woodard, K. T., Kalburgi, S. N., et al. (2013). Improved survival and reduced phenotypic severity following AAV9/MECP2 gene transfer to neonatal and juvenile male *Mecp2* knockout mice. *Molecular Therapy*, 21(1), 18–30. doi: 10.1038/mt.2012.200
- Gadalla, K. K. E., Vudhironarit, T., Hector, R. D., Sinnett, S., Bahey, N. G., Bailey, M. E. S., et al. (2017). Development of a Novel AAV Gene therapy cassette with improved safety features and efficacy in a mouse model of Rett syndrome. *Molecular Therapy: Methods & Clinical Development*, 5, 180–190. doi: 10.1016/j.omtm.2017.04.007
- Garg, S. K., Lioy, D. T., Cheval, H., McGann, J. C., Bissonnette, J. M., Murtha, M. J., et al. (2013). Systemic delivery of MeCP2 rescues behavioral and cellular deficits in female mouse models

- of Rett syndrome. *Journal of Neuroscience*, 33(33), 13612–13620. doi: 10.1523/JNEUROSCI.1854-13.2013
- Georgel, P. T., Horowitz-Scherer, R. A., Adkins, N., Woodcock, C. L., Wade, P. A., & Hansen, J. C. (2003). Chromatin compaction by human MeCP2. Assembly of novel secondary chromatin structures in the absence of DNA methylation. *Journal of Biological Chemistry*, 278(34), 32181–32188. doi: 10.1074/jbc.M305308200
- Ghosh, R. P., Nikitina, T., Horowitz-Scherer, R. A., Gierasch, L. M., Uversky, V. N., Hite, K., et al. (2010). Unique physical properties and interactions of the domains of methylated DNA binding protein 2. *Biochemistry*, 49(21), 4395–4410. doi: 10.1021/bi9019753
- Giacometti, E., Luikenhuis, S., Beard, C., & Jaenisch, R. (2007). Partial rescue of MeCP2 deficiency by postnatal activation of MeCP2. *Proceedings of the National Academy of Sciences U.S.A.*, 104(5), 1931–1936. doi: 10.1073/pnas.0610593104
- Gianakopoulos, P. J., Zhang, Y., Pencea, N., Orlic-Milacic, M., Mittal, K., Windpassinger, C., et al. (2012). Mutations in MECP2 exon 1 in classical Rett patients disrupt MECP2_e1 transcription, but not transcription of MECP2_e2. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 159B, 210–216. doi: 10.1002/ajmg.b.32015
- Goffin, D., Allen, M., Zhang, L., Amorim, M., Wang, I. T. J., Reyes, A. R. S., et al. (2012). Rett syndrome mutation MeCP2 T158A disrupts DNA binding, protein stability and ERP responses. *Nature Neuroscience*, 15, 274–283. doi: 10.1038/nn.2997
- Guo, J. U., Su, Y., Shin, J. H., Shin, J., Li, H., Xie, B., et al. (2014). Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain. *Nature Neuroscience*, 17, 215–222. doi: 10.1038/nn.3607
- Guy, J., Gan, J., Selfridge, J., Cobb, S., and Bird, A. (2007). Reversal of neurological defects in a mouse model of Rett syndrome. *Science*, 315, 1143–1147. doi: 10.1126/science.1138389

- Guy, J., Hendrich, B., Holmes, M., Martin, J. E., and Bird, A. (2001). A mouse Mecp2-null mutation causes neurological symptoms that mimic Rett syndrome. *Nature Genetics*, 27, 322–326. doi: 10.1038/85899
- Hao, S., Tang, B., Wu, Z., Ure, K., Sun, Y., Tao, H., et al. (2015). Forniceal deep brain stimulation rescues hippocampal memory in Rett syndrome mice. *Nature*, 526, 430–434. doi: 10.1038/nature15694
- Harikrishnan, K. N., Chow, M. Z., Baker, E. K., Pal, S., Bassal, S., Brasacchio, D., et al. (2005). Brahma links the SWI/SNF chromatin-remodeling complex with MeCP2-dependent transcriptional silencing. *Nature Genetics*, 37, 254–264. doi: 10.1038/ng1516
- He, Y. F., Li, B. Z., Li, Z., Liu, P., Wang, Y., Tang, Q., et al. (2011). Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science*, 333, 1303–1307. doi: 10.1126/science.1210944
- Herrera, J. A., Ward, C. S., Wehrens, X. H. T., and Neul, J. L. (2016). Methyl-CpG binding-protein 2 function in cholinergic neurons mediates cardiac arrhythmogenesis. *Human Molecular Genetics*, 25, 4983–4995. doi: 10.1093/hmg/ddw326
- Horike, S. I., Cai, S., Miyano, M., Cheng, J. F., and Kohwi-Shigematsu, T. (2005). Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome. *Nature Genetics*, 37, 31–40. doi: 10.1038/ng1491
- Itoh, M., Tahimic, C. G. T., Ide, S., Otsuki, A., Sasaoka, T., Noguchi, S., et al. (2012). Methyl CpG-binding protein isoform MeCP2-e2 is dispensable for Rett syndrome phenotypes but essential for embryo viability and placenta development. *Journal of Biological Chemistry*, 287, 13859–13867. doi: 10.1074/jbc.M111.309864

- Ito-Ishida, A., Baker, S. A., Sillitoe, R. V., Sun, Y., Zhou, J., Ono, Y., et al. (2020). MeCP2 levels regulate the 3D structure of heterochromatic foci in mouse neurons. *Journal of Neuroscience*, 40, 8746–8766. doi: 10.1523/JNEUROSCI.1281-19.2020
- Ito-Ishida, A., Ure, K., Chen, H., Swann, J. W., and Zoghbi, H. Y. (2015). Loss of MeCP2 in Parvalbumin- and Somatostatin-expressing neurons in mice leads to distinct Rett syndrome-like phenotypes. *Neuron*, 88, 651–658. doi: 10.1016/j.neuron.2015.10.029
- Jeffery, L., and Nakielnny, S. (2004). Components of the DNA methylation system of chromatin control are RNA-binding proteins. *Journal of Biological Chemistry*, 279, 49479–49487. doi: 10.1074/jbc.M409070200
- Johnson, B. S., Zhao, Y. T., Fasolino, M., Lamonica, J. M., Kim, Y. J., Georgakilas, G., et al. (2017). Biotin tagging of MeCP2 in mice reveals contextual insights into the Rett syndrome transcriptome. *Nature Medicine*, 23, 1203–1214. doi: 10.1038/nm.4406
- Jones, P. L., Veenstra, G. J. C., Wade, P. A., Vermaak, D., Kass, S. U., Landsberger, N., et al. (1998). Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nature Genetics*, 19, 187–191. doi: 10.1038/561
- Jung, B. P., Jugloff, D. G. M., Zhang, G., Logan, R., Brown, S., and Eubanks, J. H. (2003). The expression of methyl CpG binding factor MeCP2 correlates with cellular differentiation in the developing rat brain and in cultured cells. *J. Neurobiol.* 55, 86–96. doi: 10.1002/neu.10201
- Kahanovitch, U., Patterson, K. C., Hernandez, R., and Olsen, M. L. (2019). Glial dysfunction in MeCP2 deficiency models: implications for Rett syndrome. *International Journal of Molecular Sciences*, 20, 3813. doi: 10.3390/ijms20153813
- Kaufmann, W. E., Sprouse, J., Rebowe, N., Hanania, T., Klammer, D., and Missling, C. U. (2019). ANAVEX®2-73 (blarcamesine), a Sigma-1 receptor agonist, ameliorates neurologic

- impairments in a mouse model of Rett syndrome. *Pharmacology Biochemistry and Behavior*, 187, 172796. doi: 10.1016/j.pbb.2019.172796
- Kernohan, K. D., Vernimmen, D., Gloor, G. B., and Bérubé, N. G. (2014). Analysis of neonatal brain lacking ATRX or MeCP2 reveals changes in nucleosome density, CTCF binding and chromatin looping. *Nucleic Acids Research*, 42, 8356–8368. doi: 10.1093/nar/gku564
- Kerr, B., Alvarez-Saavedra, M., Sáez, M. A., Saona, A., and Young, J. I. (2008). Defective body-weight regulation, motor control and abnormal social interactions in *Mecp2* hypomorphic mice. *Human Molecular Genetics*, 17, 1707–1717. doi: 10.1093/hmg/ddn061
- Khan, A. W., Ziemann, M., Rafehi, H., Maxwell, S., Ciccotosto, G. D., and El-Osta, A. (2017). MeCP2 interacts with chromosomal microRNAs in brain. *Epigenetics*, 12, 1028–1037. doi: 10.1080/15592294.2017.1391429
- Khwaja, O. S., Ho, E., Barnes, K. V., O’Leary, H. M., Pereira, L. M., Finkelstein, Y., et al. (2014). Safety, pharmacokinetics, and preliminary assessment of efficacy of mecasermin (recombinant human IGF-1) for the treatment of Rett syndrome. *Proceedings of the National Academy of Sciences U.S.A.*, 111, 4596–4601. doi: 10.1073/pnas.1311141111
- Kishi, N., and Macklis, J. D. (2010). MeCP2 functions largely cell-autonomously, but also non-cell-autonomously, in neuronal maturation and dendritic arborization of cortical pyramidal neurons. *Experimental Neurology*, 222, 51–58. doi: 10.1016/j.expneurol.2009.12.007
- Klose, R. J., Sarraf, S. A., Schmiedeberg, L., McDermott, S. M., Stancheva, I., and Bird, A. P. (2005). DNA binding selectivity of MeCP2 due to a requirement for A/T sequences adjacent to methyl-CpG. *Molecular Cell*, 19, 667–678. doi: 10.1016/j.molcel.2005.07.021
- Kriaucionis, S., and Bird, A. (2004). The major form of MeCP2 has a novel N-terminus generated by alternative splicing. *Nucleic Acids Research*, 32, 1818–1823. doi: 10.1093/nar/gkh349

- Kriaucionis, S., and Heintz, N. (2009). The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science*, 324, 929–930. doi: 10.1126/science.1169786
- Kyle, S. M., Saha, P. K., Brown, H. M., Chan, L. C., and Justice, M. J. (2016). MeCP2 coordinates liver lipid metabolism with the NCoR1/HDAC3 corepressor complex. *Human Molecular Genetics*, 25, 3029–3041. doi: 10.1093/hmg/ddw156
- Kyle, S. M., Vashi, N., and Justice, M. J. (2018). Rett syndrome: a neurological disorder with metabolic components. *Open Biology*, 8, 170216. doi: 10.1098/rsob.170216
- Lavery, L. A., Ure, K., Wan, Y. W., Luo, C., Trostle, A. J., Wang, W., et al. (2020). Losing DNMT3A dependent methylation in inhibitory neurons impairs neural function by a mechanism impacting Rett syndrome. *eLife*, 9, 52981. doi: 10.7554/eLife.52981
- Lavery, L. A., and Zoghbi, H. Y. (2019). The distinct methylation landscape of maturing neurons and its role in Rett syndrome pathogenesis. *Current Opinion in Neurobiology*, 59, 180–188. doi: 10.1016/j.conb.2019.08.001
- Lewis, J. D., Meehan, R. R., Henzel, W. J., Maurer-Fogy, I., Jeppesen, P., Klein, F., et al. (1992). Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell*, 69, 905–914. doi: 10.1016/0092-8674(92)90610-O
- Li, C. H., Coffey, E. L., Dall'Agnese, A., Hannett, N. M., Tang, X., Henninger, J. E., et al. (2020). MeCP2 links heterochromatin condensates and neurodevelopmental disease. *Nature*, 586, 440–444. doi: 10.1038/s41586-020-2574-4
- Lioy, D. T., Garg, S. K., Monaghan, C. E., Raber, J., Foust, K. D., Kaspar, B. K., et al. (2011). A role for glia in the progression of Rett syndrome. *Nature*, 475, 497–500. doi: 10.1038/nature10214
- Lister, R., Pelizzola, M., Downen, R. H., Hawkins, R. D., Hon, G., Tonti-Filippini, J., et al. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*, 462, 315–322. doi: 10.1038/nature08514

- Liu, M., Movahed, S., Dangi, S., Pan, H., Kaur, P., Bilinovich, S. M., et al. (2020). DNA looping by two 5-methylcytosine-binding proteins quantified using nanofluidic devices. *Epigenet. Chrom.* 13:18. doi: 10.1186/s13072-020-00339-7
- Lorincz, M. C., Schübeler, D., and Groudine, M. (2001). Methylation-mediated proviral silencing is associated with MeCP2 recruitment and localized histone H3 deacetylation. *Molecular and Cellular Biology*, 21, 7913–7922. doi: 10.1128/mcb.21.23.7913-7922.2001
- Maczawa, I., and Jin, L. W. (2010). Rett syndrome microglia damage dendrites and synapses by the elevated release of glutamate. *Journal of Neuroscience*, 30, 5346–5356. doi: 10.1523/JNEUROSCI.5966-09.2010
- Maczawa, I., Swanberg, S., Harvey, D., LaSalle, J. M., and Jin, L. W. (2009). Rett syndrome astrocytes are abnormal and spread MeCP2 deficiency through gap junctions. *Journal of Neuroscience*, 29, 5051–5061. doi: 10.1523/JNEUROSCI.0324-09.2009
- Martínez De Paz, A., Khajavi, L., Martin, H., Claveria-Gimeno, R., Tom Dieck, S., Cheema, M. S., et al. (2019). MeCP2-E1 isoform is a dynamically expressed, weakly DNA-bound protein with different protein and DNA interactions compared to MeCP2-E2. *Epigenetics & Chromatin*, 12, 63. doi: 10.1186/s13072-019-0298-1
- Matagne, V., Borloz, E., Ehinger, Y., Saidi, L., Villard, L., and Roux, J.-C. (2020). Severe off-target effects following intravenous delivery of AAV9-MECP2 in a female mouse model of Rett syndrome. *Neurobiology of Disease*, 149, 105235. doi: 10.1016/j.nbd.2020.105235
- Maunakea, A. K., Chepelev, I., Cui, K., and Zhao, K. (2013). Intragenic DNA methylation modulates alternative splicing by recruiting MeCP2 to promote exon recognition. *Cell Research*, 23, 1256–1269. doi: 10.1038/cr.2013.110
- Maxwell, S. S., Pelka, G. J., Tam, P. P., and El-Osta, A. (2013). Chromatin context and ncRNA highlight targets of MeCP2 in brain. *RNA Biology*, 10, 1741–1757. doi: 10.4161/rna.26921

- McCauley, M. D., Wang, T., Mike, E., Herrera, J., Beavers, D. L., Huang, T. W., et al. (2011). Rett syndrome: pathogenesis of lethal cardiac arrhythmias in *Mecp2* mutant mice: implication for therapy in Rett syndrome. *Science Translational Medicine*, 3, 2982. doi: 10.1126/scitranslmed.3002982
- McGraw, C. M., Samaco, R. C., and Zoghbi, H. Y. (2011). Adult neural function requires MeCP2. *Science*, 333, 186. doi: 10.1126/science.1206593
- McMahon, A. C., Rahman, R., Jin, H., Shen, J. L., Fieldsend, A., Luo, W., et al. (2016). TRIBE: hijacking an RNA-editing enzyme to identify cell-specific targets of RNA-binding proteins. *Cell*, 165, 742–753. doi: 10.1016/j.cell.2016.03.007
- Meehan, R., Lewis, J. D., and Bird, A. P. (1992). Characterization of MECP2, a vertebrate DNA binding protein with affinity for methylated DNA. *Nucleic Acids Research*, 20, 5085–5092. doi: 10.1093/nar/20.19.5085
- Meguro-Horike, M., Yasui, D. H. D. H., Powell, W., Schroeder, D. I., Oshimura, M., LaSalle, J. M., et al. (2011). Neuron-specific impairment of inter-chromosomal pairing and transcription in a novel model of human 15q-duplication syndrome. *Human Molecular Genetics*, 20, 3798–3810. doi: 10.1093/hmg/ddr298
- Mellén, M., Ayata, P., Dewell, S., Kriaucionis, S., and Heintz, N. (2012). MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. *Cell*, 151, 1417–1430. doi: 10.1016/j.cell.2012.11.022
- Millar-Büchner, P., Philp, A. R., Gutierrez, N., Villanueva, S., Kerr, B., and Flores, C. A. (2016). Severe changes in colon epithelium in the *Mecp2*-null mouse model of Rett syndrome. *Molecular and Cellular Pediatrics*, 3, 37. doi: 10.1186/s40348-016-0065-3

- Mnatzakanian, G. N., Lohi, H., Munteanu, I., Alfred, S. E., Yamada, T., MacLeod, P. J. M., et al. (2004). A previously unidentified MECP2 open reading frame defines a new protein isoform relevant to Rett syndrome. *Nature Genetics*, 36, 339–341. doi: 10.1038/ng1327
- Muotri, A. R., Marchetto, M. C. N., Coufal, N. G., Oefner, R., Yeo, G., Nakashima, K., et al. (2010). L1 retrotransposition in neurons is modulated by MeCP2. *Nature*, 468, 443–446. doi: 10.1038/nature09544
- Nan, X., Campoy, F. J., and Bird, A. (1997). MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell*, 88, 471–481. doi: 10.1016/S0092-8674(00)81887-5
- Nikitina, T., Ghosh, R. P., Horowitz-Scherer, R. A., Hansen, J. C., Grigoryev, S. A., and Woodcock, C. L. (2007). MeCP2-chromatin interactions include the formation of chromatosome-like structures and are altered in mutations causing Rett syndrome. *Journal of Biological Chemistry*, 282, 28237–28245. doi: 10.1074/jbc.M704304200
- Olson, C. O., Zachariah, R. M., Ezeonwuka, C. D., Liyanage, V. R. B., and Rastegar, M. (2014). Brain region-specific expression of MeCP2 isoforms correlates with DNA methylation within Mecp2 regulatory elements. *PLoS One*, 9, e0090645. doi: 10.1371/journal.pone.0090645
- Osenberg, S., Karten, A., Sun, J., Li, J., Charkowick, S., Felice, C. A., et al. (2018). Activity-dependent aberrations in gene expression and alternative splicing in a mouse model of Rett syndrome. *Proceedings of the National Academy of Sciences of the United States of America*, 115, E5363–E5372. doi: 10.1073/pnas.1722546115
- Piccolo, F. M., Liu, Z., Dong, P., Hsu, C. L., Stoyanova, E. I., Rao, A., et al. (2019). MeCP2 nuclear dynamics in live neurons results from low and high affinity chromatin interactions. *eLife*, 8, e051449. doi: 10.7554/eLife.51449

- Pohodich, A. E., Yalamanchili, H., Raman, A. T., Wan, Y. W., Gundry, M., Hao, S., et al. (2018). Forniceal deep brain stimulation induces gene expression and splicing changes that promote neurogenesis and plasticity. *eLife*, 7, e034021. doi: 10.7554/eLife.34031
- Rajavelu, A., Lungu, C., Emperle, M., Dukatz, M., Bröhm, A., Broche, J., et al. (2018). Chromatin-dependent allosteric regulation of DNMT3A activity by MeCP2. *Nucleic Acids Research*, 46, 9044–9056. doi: 10.1093/nar/gky715
- Raman, A. T., Pohodich, A. E., Wan, Y. W., Yalamanchili, H. K., Lowry, W. E., Zoghbi, H. Y., et al. (2018). Apparent bias toward long gene misregulation in MeCP2 syndromes disappears after controlling for baseline variations. *Nature Communications*, 9, 3225. doi: 10.1038/s41467-018-05627-1
- Reichow, B., George-Puskar, A., Lutz, T., Smith, I. C., and Volkmar, F. R. (2015). Brief report: systematic review of Rett syndrome in males. *Journal of Autism and Developmental Disorders*, 45, 3377–3383. doi: 10.1007/s10803-015-2519-1
- Renthal, W., Boxer, L. D., Hrvatin, S., Li, E., Silberfeld, A., Nagy, M. A., et al. (2018). Characterization of human mosaic Rett syndrome brain tissue by single-nucleus RNA sequencing. *Nature Neuroscience*, 21, 1670–1679. doi: 10.1038/s41593-018-0270-6
- Ross, P. D., Guy, J., Selfridge, J., Kamal, B., Bahey, N., Tanner, K. E., et al. (2016). Exclusive expression of MeCP2 in the nervous system distinguishes between brain and peripheral Rett syndrome-like phenotypes. *Human Molecular Genetics*, 25, 4389–4404. doi: 10.1093/hmg/ddw269
- Rube, H. T., Lee, W., Hejna, M., Chen, H., Yasui, D. H., Hess, J. F., et al. (2016). Sequence features accurately predict genome-wide MeCP2 binding in vivo. *Nature Communications*, 7, 11025. doi: 10.1038/ncomms11025

- Samaco, R. C., Fryer, J. D., Ren, J., Fyffe, S., Chao, H. T., Sun, Y., et al. (2008). A partial loss of function allele of Methyl-CpG-binding protein 2 predicts a human neurodevelopmental syndrome. *Human Molecular Genetics*, 17, 1718–1727. doi: 10.1093/hmg/ddn062
- Samaco, R. C., Mandel-Brehm, C., Chao, H. T., Ward, C. S., Fyffe-Maricich, S. L., Ren, J., et al. (2009). Loss of MeCP2 in aminergic neurons causes cell-autonomous defects in neurotransmitter synthesis and specific behavioral abnormalities. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 21966–21971. doi: 10.1073/pnas.0912257106
- Saunders, C. J., Minassian, B. E., Chow, E. W. C., Zhao, W., and Vincent, J. B. (2009). Novel exon 1 mutations in MECP2 implicate isoform MeCP2-e1 in classical Rett syndrome. *American Journal of Medical Genetics Part A*, 149, 1019–1023. doi: 10.1002/ajmg.a.32776
- Schaevez, L. R., Gómez, N. B., Zhen, D. P., and Berger-Sweeney, J. E. (2013). MeCP2 R168X male and female mutant mice exhibit Rett-like behavioral deficits. *Genes, Brain and Behavior*, 12, 732–740. doi: 10.1111/gbb.12070
- Schafer, D. P., Heller, C. T., Gunner, G., Heller, M., Gordon, C., Hammond, T., et al. (2016). Microglia contribute to circuit defects in *Mecp2* null mice independent of microglia-specific loss of *Mecp2* expression. *eLife*, 5, e015224. doi: 10.7554/eLife.15224
- Schmidt, A., Zhang, H., and Cardoso, M. C. (2020). MeCP2 and chromatin compartmentalization. *Cells*, 9, 878. doi: 10.3390/cells9040878
- Sekul, E. A., Moak, J. P., Schultz, R. J., Glaze, D. G., Dunn, J. K., and Percy, A. K. (1994). Electrocardiographic findings in Rett syndrome: an explanation for sudden death? *Journal of Pediatrics*, 125, 80–82. doi: 10.1016/S0022-3476(94)70128-8
- Sinnamon, J. R., Kim, S. Y., Corson, G. M., Song, Z., Nakai, H., Adelman, J. P., et al. (2017). Site-directed RNA repair of endogenous *Mecp2* RNA in neurons. *Proceedings of the National*

- Academy of Sciences of the United States of America, 114, E9395–E9402. doi: 10.1073/pnas.1715320114
- Sinnamon, J. R., Kim, S. Y., Fisk, J. R., Song, Z., Nakai, H., Jeng, S., et al. (2020). In vivo repair of a protein underlying a neurological disorder by programmable RNA editing. *Cell Reports*, 32, 107878. doi: 10.1016/j.celrep.2020.107878
- Sinnett, S. E., Hector, R. D., Gadalla, K. K. E., Heindel, C., Chen, D., Zaric, V., et al. (2017). Improved MECP2 gene therapy extends the survival of MeCP2-null mice without apparent toxicity after intracisternal delivery. *Mol. Therapy Methods Clin. Dev.* 5, 106–115. doi: 10.1016/j.omtm.2017.04.006
- Skene, P. J., Illingworth, R. S., Webb, S., Kerr, A. R. W., James, K. D., Turner, D. J., et al. (2010). Neuronal MeCP2 is expressed at near histone-octamer levels and globally alters the chromatin state. *Molecular Cell*, 37, 457–468. doi: 10.1016/j.molcel.2010.01.030
- Strati, F., Cavalieri, D., Albanese, D., De Felice, C., Donati, C., Hayek, J., et al. (2016). Altered gut microbiota in Rett syndrome. *Microbiome*, 4, 41. doi: 10.1186/s40168-016-0185-y
- Stuss, D. P., Cheema, M., Ng, M. K., Martinez De Paz, A., Williamson, B., Missiaen, K., et al. (2013). Impaired in vivo binding of MeCP2 to chromatin in the absence of its DNA methyl-binding domain. *Nucleic Acids Research*, 41, 4888–4900. doi: 10.1093/nar/gkt213
- Sugino, K., Hempel, C. M., Okaty, B. W., Arnson, H. A., Kato, S., Dani, V. S., et al. (2014). Cell-type-specific repression by methyl-CpG-binding protein 2 is biased toward long genes. *Journal of Neuroscience*, 34, 12877–12883. doi: 10.1523/JNEUROSCI.2674-14.2014
- Symons, F. J., Byiers, B., Tervo, R. C., and Beisang, A. (2013). Parent-reported pain in Rett syndrome. *Clinical Journal of Pain*, 29, 744–746. doi: 10.1097/AJP.0b013e318274b6bd

- Szulwach, K. E., Li, X., Li, Y., Song, C. X., Wu, H., Dai, Q., et al. (2011). 5-hmC-mediated epigenetic dynamics during postnatal neurodevelopment and aging. *Nature Neuroscience*, 14, 1607–1616. doi: 10.1038/nn.2959
- Tahiliani, M., Koh, K. P., Shen, Y., Pastor, W. A., Bandukwala, H., Brudno, Y., et al. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*, 324, 930–935. doi: 10.1126/science.1170116
- Trappe, R., Laccone, F., Cobilanschi, J., Meins, M., Huppke, P., Hanefeld, F., et al. (2001). MECP2 mutations in sporadic cases of Rett syndrome are almost exclusively of paternal origin. *American Journal of Human Genetics*, 68, 1093–1101. doi: 10.1086/320109
- Valinluck, V., Liu, P., Kang, J. I., Burdzy, A., and Sowers, L. C. (2005). 5-Halogenated pyrimidine lesions within a CpG sequence context mimic 5-methylcytosine by enhancing the binding of the methyl-CpG-binding domain of methyl-CpG-binding protein 2 (MeCP2). *Nucleic Acids Research*, 33, 3057–3064. doi: 10.1093/nar/gki612
- Valinluck, V., Tsai, H. H., Rogstad, D. K., Burdzy, A., Bird, A., and Sowers, L. C. (2004). Oxidative damage to methyl-CpG sequences inhibits the binding of the methyl-CpG binding domain (MBD) of methyl-CpG binding protein 2 (MeCP2). *Nucleic Acids Research*, 32, 4100–4108. doi: 10.1093/nar/gkh739
- Veeraragavan, S., Wan, Y.-W., Connolly, D. R., Hamilton, S. M., Ward, C. S., Soriano, S., et al. (2015). Loss of MeCP2 in the rat models regression, impaired sociability and transcriptional deficits of Rett syndrome. *Human Molecular Genetics*, 25, ddw178. doi: 10.1093/hmg/ddw178
- Vogel Ciernia, A., Pride, M. C., Durbin-Johnson, B., Noronha, A., Chang, A., Yasui, D. H., et al. (2017). Early motor phenotype detection in a female mouse model of Rett syndrome is improved by cross-fostering. *Human Molecular Genetics*, 26, 1839–1854. doi: 10.1093/hmg/ddx087

- Vogel Ciernia, A., Yasui, D. H., Pride, M. C., Durbin-Johnson, B., Noronha, A. B., Chang, A., et al. (2018). MeCP2 isoform e1 mutant mice recapitulate motor and metabolic phenotypes of Rett syndrome. *Human Molecular Genetics*, 27, 4077–4093. doi: 10.1093/hmg/ddy301
- Wahba, G., Schock, S. C., Claridge, E., Bettolli, M., Grynspan, D., Humphreys, P., et al. (2015). MeCP2 in the enteric nervous system. *Neurogastroenterology and Motility*, 27, 1156–1161. doi: 10.1111/nmo.12605
- Wang, L., Hu, M., Zuo, M. Q., Zhao, J., Wu, D., Huang, L., et al. (2020). Rett syndrome-causing mutations compromise MeCP2-mediated liquid-liquid phase separation of chromatin. *Cell Research*, 30, 393–407. doi: 10.1038/s41422-020-0288-7
- Wasag, P., and Lenartowski, R. (2016). Nuclear matrix - structure, function and pathogenesis. *Postępy Higieny i Medycyny Doświadczalnej*, 70, 1206–1219.
- Weitzel, J. M., Buhrmester, H., and Strätling, W. H. (1997). Chicken MAR-binding protein ARBP is homologous to rat methyl-CpG-binding protein MeCP2. *Molecular and Cellular Biology*, 17, 5656–5666. doi: 10.1128/mcb.17.9.5656
- Wong, J. J. L., Gao, D., Nguyen, T. V., Kwok, C. T., Van Geldermalsen, M., Middleton, R., et al. (2017). Intron retention is regulated by altered MeCP2-mediated splicing factor recruitment. *Nature Communications*, 8, 134. doi: 10.1038/ncomms15134
- Yasui, D. H., Gonzales, M. L., Aflatooni, J. O., Crary, F. K., Hu, D. J., Gavino, B. J., et al. (2014). Mice with an isoform-ablating *Mecp2* exon 1 mutation recapitulate the neurologic deficits of Rett syndrome. *Human Molecular Genetics*, 23, 2447–2458. doi: 10.1093/hmg/ddt640
- Yasui, D. H., Peddada, S., Bieda, M. C., Vallero, R. O., Hogart, A., Nagarajan, R. P., et al. (2007). Integrated epigenomic analyses of neuronal MeCP2 reveal a role for long-range interaction

- with active genes. *Proceedings of the National Academy of Sciences U.S.A.*, 104, 9416–9421. doi: 10.1073/pnas.0707442104
- Yasui, D. H., Scoles, H. A., Horike, S.-I., Meguro-Horike, M., Dunaway, K. W., Schroeder, D. I., et al. (2011). 15q11.2-13.3 chromatin analysis reveals epigenetic regulation of CHRNA7 with deficiencies in Rett and autism brain. *Human Molecular Genetics*, 20, 4311–4323. doi: 10.1093/hmg/ddr357
- Young, J. I., Hong, E. P., Castle, J. C., Crespo-Barreto, J., Bowman, A. B., Rose, M. F., et al. (2005). Regulation of RNA splicing by the methylation-dependent transcriptional repressor methyl-CpG binding protein 2. *Proceedings of the National Academy of Sciences U.S.A.*, 102, 17551–17558. doi: 10.1073/pnas.0507856102
- Yu, F., Zingler, N., Schumann, G., and Strätling, W. H. (2001). Methyl-CpG-binding protein 2 represses LINE-1 expression and retrotransposition but not Alu transcription. *Nucleic Acids Research*, 29, 4493–4501. doi: 10.1093/nar/29.21.4493
- Zachariah, R. M., Olson, C. O., Ezeonwuka, C., and Rastegar, M. (2012). Novel MeCP2 isoform-specific antibody reveals the endogenous MeCP2E1 expression in murine brain, primary neurons and Astrocytes. *PLoS One*, 7, e0049763. doi: 10.1371/journal.pone.0049763
- Zhang, P., Ludwig, A. K., Hastert, F. D., Rausch, C., Lehmkuhl, A., Hellmann, I., et al. (2017). L1 retrotransposition is activated by Ten-eleven-translocation protein 1 and repressed by methyl-CpG binding proteins. *Nucleus*, 8, 548–562. doi: 10.1080/19491034.2017.1393316
- Zhao, B., Wu, Q., Ye, A. Y., Guo, J., Zheng, X., Yang, X., et al. (2019). Somatic LINE-1 retrotransposition in cortical neurons and non-brain tissues of Rett patients and healthy individuals. *PLoS Genetics*, 15, e1008043. doi: 10.1371/journal.pgen.1008043
- Zoghbi, H. Y. (2005). MeCP2 dysfunction in humans and mice. *Journal of Child Neurology*, 20, 736–740. doi: 10.1177/08830738050200090701

- Zuliani, I., Urbinati, C., Valenti, D., Quattrini, M. C., Medici, V., Cosentino, L., et al. (2020). The anti-diabetic drug metformin rescues aberrant mitochondrial activity and restrains oxidative stress in a female mouse model of Rett syndrome. *Journal of Clinical Medicine*, 9, 1669. doi: 10.3390/jcm9061669
- Maheshwari N, Christodoulou J, Percy A (2018) Rett syndrome. *Anesthesiol und Intensivmed*. doi:10.19224/ai2018.S117.
- Faroon O, Ruiz P (2016) Polychlorinated biphenyls: New evidence from the last decade. *Toxicol Ind Health*. doi:10.1177/0748233715587849.
- Eubig PA, Aguiar A, Schantz SL (2010) Lead and PCBs as risk factors for attention deficit/hyperactivity disorder. *Environ Health Perspect*. doi:10.1289/ehp.0901852.
- Berghuis SA, Bos AF, Sauer PJJ, Roze E (2015) Developmental neurotoxicity of persistent organic pollutants: an update on childhood outcome. *Arch Toxicol*. doi:10.1007/s00204-015-1463-3.

Chapter 2

Sex-specific single cell-level transcriptomic signatures of Rett syndrome disease progression

2.1 Preface

Sections 2.1.1-2.1.12 contains text that was co-written and accepted in Nature Communications Biology.

It is also made available on BioRxiv. DOI: 10.1101/2024.05.16.594595. This article has been modified to satisfy the formatting requirements of graduate studies and this dissertation.

2.1.1 Authorship

Osman Sharifi^{1,3,4}, Viktoria Haghani^{1,3,4}, Kari E. Neier^{1,3,4}, Keith J. Fraga^{2,3}, Ian Korf^{2,3}, Sophia M. Hakam^{1,3,4}, Gerald Quon^{2,3}, Nelson Johansen^{2,3}, Dag H. Yasui^{1,3,4#}, Janine M. LaSalle^{1,3,4##}

¹Medical Microbiology and Immunology, School of Medicine, ²Cellular and Molecular Biology, College of Biological Sciences, ³Genome Center, ⁴MIND Institute, University of California, Davis, CA 95616

*corresponding author, jmlasalle@ucdavis.edu

#equal contribution

ORCID OS 0000-0002-4630-2587, VH000-0002-3700-4027, KEN 0000-0002-9700-4045, KF 0000-0002-8930-6510, IFK 0000-0002-3480-2031, SH 0009-0006-1465-6571 JQ 0000-0002-1716-0153, NJ 0000-0002-4436-969X, DHY0000-0003-1178-988X, JML 0000-0002-3480-2031

2.2 Abstract

Dominant X-linked diseases are uncommon due to female X chromosome inactivation (XCI). While random XCI usually protects females against X-linked mutations, Rett syndrome (RTT) is a female neurodevelopmental disorder caused by heterozygous *MECP2* mutation. After 6-18 months of typical neurodevelopment, RTT girls undergo poorly understood regression. We performed longitudinal snRNA-seq on cerebral cortex in a construct-relevant *Mecp2e1* mutant mouse model of RTT, revealing transcriptional effects of cell type, mosaicism, and sex on progressive disease phenotypes. Across cell types, we observed sex differences in the number of differentially expressed genes (DEGs) with 6x more DEGs in mutant females than males. Unlike males, female DEGs emerged prior to symptoms, were enriched for homeostatic gene pathways in distinct cell types over time and correlated with disease phenotypes and human RTT cortical cell transcriptomes. Non-cell-autonomous effects were prominent and dynamic across disease progression of *Mecp2e1* mutant females, indicating wild-type-expressing cells normalizing transcriptional homeostasis. These results improve understanding of RTT progression and treatment.

2.3 Introduction

Rett syndrome is a neurodevelopmental disorder primarily affecting females and is characterized by a range of symptoms such as loss of speech, motor abnormalities, and developmental regression at about 6-18 months of age¹. Rett syndrome most often occurs through spontaneous missense or truncation mutations, thereby predominantly affecting XX females². Males with *MECP2* mutations that survive infancy are rare but can be observed clinical phenotypes ranging from severe neonatal encephalopathy to cognitive impairment³. *MECP2* encodes the DNA binding protein, methyl CpG binding protein 2 (MeCP2), which is a critical regulator of neuronal gene expression in the brain. Among the two alternatively spliced *MECP2* transcripts, only the MeCP2e1 isoform contributes to

RTT disease phenotypes⁴. However, most mouse studies of RTT utilize the exon 3-4 knockout model in *Mecp2*^{+/y} males, which is a null model effective for studying MeCP2 function, but not a construct- or sex-relevant model for human RTT⁵. RTT females are heterozygous for *MECP2* (*MECP2*^{+/+}) mutations and are therefore mosaic for both *MECP2* wild-type and mutant cells in brain. Prior studies in Rett syndrome have suggested potential non-cell-autonomous effects of MeCP2 deficiency on wild-type expressing cells in brain, but these effects have been poorly characterized at a cellular and molecular level⁶⁻⁸. RTT is characterized by a seemingly typical development in infancy, followed by progressive stages of regression in developmental milestones beginning around 6-18 months of age and lasting through early adulthood¹. We have previously demonstrated that the *Mecp2e1* deficient mouse model of RTT, modeled after a human mutation, recapitulates the RTT-relevant extended period of disease symptom progression^{4,9,10}. However, it is not known when and in which cell types the molecular changes responsible for disease progression occur in *MECP2* mutant females versus males.

To explore the effects of cellular mosaicism, sex, and cell type on the progression of disease in Rett syndrome, we employed single nuclei RNA-seq (sn-RNA-seq 5') analysis in the cerebral cortex of the *Mecp2e1* mutant mouse model. We examined the influence of sex, cell type, cellular mosaicism, and disease stage, correlated with progressive disease phenotypes using a systems-level perspective. These results demonstrate that MeCP2 deficiency in females results show an inherently different disease progression at the cellular and molecular level compared to males, involving non-cell-autonomous transcriptional changes to homeostatic gene pathways that correlate with disease phenotypes and stages.

2.4 Results

2.4.1 Experimental design to test longitudinal, cellular, and sex-specific

transcriptional dysregulation in a symptomatically progressive mouse model of Rett syndrome

To identify sex, cell type, and disease stage specific transcriptional differences in *Mecp2e1* deficient mouse cortex, single nuclei RNA sequencing (sn-RNA seq 5') analysis was performed to include the engineered mutation at the 5' translational start site of the *Mecp2e1* isoform⁴. Three longitudinal post-natal time points were chosen to correspond to pre-symptomatic (PND 30), disease onset (PND 60) and late disease stages (PND 120 for *Mecp2e1*^{-/-} males, PND 150 for *Mecp2e1*^{-/+} females) compared to sex-matched wild-type (WT) littermates^{9,10} (**Figure 1a**). A total of 28 cortex samples were used, including two per timepoint and genotype for males (PND30, 60, 120) and two per genotype for females at PND30 and 60), but four replicates per genotype were used in females at P150 in order to study mosaicism (**Table 1**).

Cortical nuclei were assigned to 14 different cell types based on 3,000 cell marker genes from the Allen brain atlas cortex transcriptomics data¹¹. 93,798 cells from both sexes, four genotypes and three time points were all clustered unsupervised (**Figure 1b**). Four excitatory neuron cell types were identified, corresponding with cortical layers 2 to 6 (L2-6), as well as five inhibitory cell types (Pvalb, Vip, Sst, Sncg, Lamp5) and five non-neuronal cell types (pericytes, endothelial, oligodendrocytes, astrocytes, non-neuronal including microglia). Unbiased marker genes for all 14 cell types were identified, supporting the distinction of our candidate cell types (**Figure 1c**).

Five separate hypotheses were tested, comparing cells across different genotypes and expression phenotypes (mutant vs wild-type-expressing cells within females). In addition to comparing cells from *Mecp2e1*^{-/-} to *Mecp2e1*^{+/-} (experiment 1) and *Mecp2e1*^{-/+} to *Mecp2e1*^{+/+} (experiment 2), wild-

type *Mecp2e1* expressing cells from the *Mecp2e1*^{-/+} females were compared to the wild-type expressing cells from the *Mecp2e1*^{+/+} (experiment 3) and mutant *Mecp2e1* expressing cells from the *Mecp2e1*^{-/+} females were compared to either wild-type expressing cells from the *Mecp2e1*^{+/+} (experiment 4) or wild-type expressing cells within *Mecp2e1*^{-/+} females (experiment 5) to test for cell non-autonomous effects (**Figure 1d**).

2.4.2 Sexually dimorphic trajectories of transcriptional dysregulation across cortical cell types

To accurately characterize alterations in gene transcript abundance, four computational methods for identifying differentially expressed genes (DEGs) from single nucleus RNA sequencing (snRNA-seq) data were evaluated with single cell data sets (Limma-VoomCC, Limma, EdgeR, and DESeq2) with partial overlap (**Supplemental Figure 1**). An overlap of DEGs between EdgeR and Limma-VoomCC showed that Limma-VoomCC identifies mostly the same DEGs, but EdgeR identifies many more DEGs that likely include many false positives, as highlighted by Squair et al¹². (**Supplemental Figure 1b**). Ultimately, Limma-Voom Consensus Correlation (Limma-VoomCC) was selected for DEG analysis based on the ability to reveal high expressing DEGs amongst diverse gene transcripts expressed¹³. Further, Limma-VoomCC controlled for the inter-correlations of cells from the same animals^{14,15}. Overall, Limma-VoomCC analyses of all cell types in experiments 1 and 2 revealed a total of 1440 significant DEGs after adjusting for false discovery (**Table 1**). In males from experiment 1, 165 or 85% showed higher and 30 or 15% showed lower transcript levels in *Mecp2e1* mutant cortical cells compared to wild-type controls across the three time points, with fold changes ranging from a low of -1.99 for *Sst* to a high of +2.31 for *Cst3*. In females from experiment 2, 282 or 22% showed higher and 959 or 77% showed lower transcript levels in *Mecp2e1* mutant cortical cells compared to wild-type controls across the three time points, with fold changes ranging from a low of -2.69 for *Snhg11* to a high of +3.47 for *Ay036118* (**Supplemental Table 1**). DEsingle was also used

as a complementary approach to identify lower confidence DEGs for transcripts expressed at low levels (**Supplemental Figure 2**). To ensure that DEGs detected were not due to changes in cell types, we examined cell proportions which did not show changes over time (**Supplemental Figure 3**). Cell clustering based on cell type, time point, sex and *Mecp2e1* genotype did not show evidence of batch effects (**Supplemental Figure 4a-d**). Further, an analysis of the top high and low expressing genes showed that brain samples from replicate mice were consistent (**Supplemental Figure 5**).

Analysis of DEGs by Limma-VoomCC, DEsingle, and EdgeR revealed that cell type transcriptional changes associated with *Mecp2e1* deficiency were markedly different by sex and disease stage in multiple cortical cell types and these differences were consistent across bioinformatic tools (**Figure 2, Supplemental Figure 2**). At the pre-symptomatic stage, *Mecp2e1^{-/-}* male P30 from experiment 1 cortical cells had only 9 DEGs compared with wild-type, including 3 DEGs in L2/3 neurons, 4 DEGs in L4 neurons (including immediate early genes *Arc* and *Junb*), and 1 DEG (*AC149090.1*) in Lamp5 and Vip neurons (**Figure 2a**). In contrast, *Mecp2e1^{-/+}* female from experiment 2 single cortical cells showed the strongest transcriptional dysregulation, for a total of 1215 DEGs at P30 (Limma-VoomCC). Interestingly, *Mecp2e1^{-/+}* female Pvalb DEGs at P30 had a significant (*p-value* ≤ 0.00075) enrichment of imprinted genes, including *Meg3*, *Xist*, *Gnas*, *Kcnq1ot1*, *Np1l5*, *Ntm*, *Peg3* and *Snrpn* (**Figure 2b**), a result that was not observed in males.

At the disease onset P60 timepoint, 73 DEGs were identified in *Mecp2e1^{-/-}* males, predominated by 56 DEGs in L2/3 neurons, but also including 7 DEGs in astrocytes, 4 DEGs in L4 and 6 in Scng neurons and 1 DEG in L6 and Sst neurons (**Figure 2c**). *Mecp2e1^{-/+}* female cortical cells had 47 DEGs, with 46 in L6 excitatory neurons and 1 DEG in Sncg inhibitory neurons (**Figure 2d**). Further, *Mecp2e1^{-/+}* female DEGs at P60 included *AY036118* (+3.47-fold change), *Ptprd*, *Edil3*, *Ptgsd*, *Plp1*, *Atp6v0b*, *Kcn11ot1*, *Gria2*, *Nrxn1*, *Arpp21*, *Snhg11*. *Mecp2e1^{-/+}* females had 3 DEGs with 2 in VIP inhibitory neurons and 1 DEG in Pvalb inhibitory neurons.

By the late disease P150 time point, only VIP interneurons contained DEGs in *Mecp2e1^{-/+}* cortical cell types, including long non-coding RNAs *Snhg11* (p value = 0.0043) and *Meg3* (p -value = 0.017). Remarkably, *Mecp2e1^{-/+}* female cortical cells were most transcriptionally dysregulated prior to the onset of symptoms, as the number of DEGs decreased in number as disease symptoms progressed (**Figure 2d**). Overall, *Mecp2e1^{-y}* male DEGs increased in number with disease progression, but *Mecp2e1^{-y}* male cortical cell types had only 199 DEGs across all three time points, which was only 16.3% of the total *Mecp2e1^{-/+}* female DEGs (**Figure 2c, d**).

To identify enriched functional pathways connecting RTT transcriptional progression, Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis was performed using DEGs (Limma-VoomCC, p -value ≤ 0.05) from each cell type. KEGG pathways that were persistent over P30, P60, and P120 in *Mecp2e1^{-y}* male cortical cells or P30, P60 and P150 in *Mecp2e1^{-/+}* female cortical cells are shown (**Figure 2e-f, Supplemental Table 2**). Distinctly different pathway dysregulation was observed between *Mecp2e1^{-/+}* females and *Mecp2e1^{-y}* male cortical cells by two key metrics. First, *Mecp2e1^{-/+}* cortical cell DEGs were enriched for 18 different pathways consistently across time points, compared to only two in *Mecp2e1^{-y}* males, of which only gastric acid secretion overlaps with *Mecp2e1^{-/+}* pathways. Second, specifically in pre-symptomatic P30 *Mecp2e1^{-/+}* females, Pvalb and Sst neurons shared 14 enriched pathways including synaptic vesicle cycle, retrograde endocannabinoid signaling, oxytocin signaling, morphine addiction, long-term potentiation, insulin secretion, glutamatergic synapse, gastric acid secretion, dopaminergic synapse, circadian entrainment, cAMP signaling pathway, amphetamine addiction, Alzheimer's disease, and aldosterone synthesis and secretion (**Figure 2f**) compared to only the gastric acid secretion pathway at the disease onset P60 and late disease P120 time point in *Mecp2e1^{-y}* males (**Figure 2e**). Interestingly, by symptom onset at P60 in *Mecp2e1^{-/+}* females, 6 dysregulated KEGG pathways including nicotine addiction, long-term potentiation, insulin secretion, glutamatergic synapse, cAMP signaling, and amphetamine addiction (found in P30 Pvalb

and Sst neurons) were distinctly significantly enriched in L5 and L6 excitatory neurons (**Figure 2f**). At the late disease P150 timepoint, *Mecp2e1*^{-/+} female cortical astrocytes remarkably and distinctly were significantly enriched for 15 out of the 18 total convergent KEGG pathways (**Figure 2f**). While some of the reduced KEGG pathway enrichment in *Mecp2e1* deficient males compared to females could be due to fewer DEGs observed overall and especially at the pre-symptomatic stage in *Mecp2e1*^{-/y} cortical cells, the significant enrichment of ubiquitin mediated proteolysis specifically at *Mecp2e1*^{-/y} P30 when DEGs were fewest (**Figure 2e**) suggests that DEG number is less important than the specificity of gene pathways dysregulated in the male *Mecp2e1*^{-y} mouse model. We also performed an enrichment analysis for DEGs based on gene length, but did not find evidence to support the previously reported repression of long genes in *Mecp2* deficient neurons of either sex¹⁶ (**Supplemental Figure 6**).

2.4.3 Co-expression networks of dysregulated genes within cortical cell types correlate with *Mecp2e1* genotype, time point, sex, body weight, and disease score

To complement the DEG analysis, we performed a systems-biology based approach, High-Definition Weighted Gene Co-expression Network Analysis (hdWGCNA) which is specifically designed for analysis of high dimensional data such as snRNA-seq¹⁷⁻¹⁹ (**Figure 3**). hdWGCNA groups genes that are co-expressed together into colored modules based on scale-free topology^{17,19} and was used to define nine distinct modules based on co-expression within a network built from transcriptomes of all detected genes from all cell types and experimental conditions. Genes in each module were compared to cell type marker genes to identify modules that uniquely correlate with phenotype (**Supplemental Figure 7**). In co-expression network analysis, we focus on the hub genes, those which are highly connected within each module. Therefore, we determine the eigengene-based connectivity, also known as kME, of each gene. The top 10 ranked co-expressed hub genes were

identified per module (**Figure 3a**) and expression enrichment for each cortical cell type was determined, which was distinct from cell type markers (**Figure 3b**). The blue hdWGCNA module corresponded to genes enriched in oligodendrocytes, while the magenta module genes were enriched in L5 and L6 neurons, with *Sez6* and *Nrp1* as hub genes. The brown module included *Grin2a*, *Grin2b*, and *Camk2a* and the black module included *Slit3* and *Gabrb3* enriched in excitatory neurons (L2-3, L4, L5, L6) showed similar cellular patterns of expression. The green module genes like *Grik1* and *Adarb2* were most highly expressed in inhibitory neurons. In contrast, the turquoise, pink, red and yellow modules are more cell type independent, being enriched in all neuronal subtypes (**Figure 3b**).

To explore the relationship between cortical co-expression gene networks and disease progression in Rett syndrome, the eigengene value of each sample within each module was correlated with the body weight, disease score, genotype, sex, and disease time point of each mouse. Eigengene values were calculated for all cortical cells, as well as each cell type individually, so that correlations with each variable of interest could be examined for each cell type (**Figure 3c, Supplemental Figure 7**). While the genes within each module partially overlapped with those that served as cell type markers, the genes within modules were independent from those that defined cell type specificity (**Supplemental Figure 7**). For all cell types, 6 out of the 9 gene modules significantly correlated with all phenotypes and experimental variables, and all modules showed at least one significant correlation (**Figure 3c, top row**). Yet, certain gene set modules such as green correlated with phenotypes in L4, L6 and Sst neurons while the turquoise and blue modules correlated most strongly with phenotypes in all neuronal subtypes (**Figure 3c**). Interestingly, magenta module genes only correlated with genotype in neurons. Astrocytes were distinct in that only blue and yellow modules correlated with both disease score and genotype. While most module-genotype correlations were positive (red), meaning that co-expressed genes in these modules were upregulated in mutant animals, the blue module uniquely was inversely correlated (blue), representing downregulated genes. Interestingly,

module-sex associations were frequent but sometimes were absent in specific cell types or time points with strong module-genotype correlations, such as the blue module in L6. Pvalb, and Sst neurons (**Figure 3c**). Further, genes from each module were tested for KEGG pathway enrichment to identify phenotype correlated dysregulated pathways. Many RTT disease progression relevant pathways such as glutamatergic synapse, GABAergic synapse, circadian rhythm and axon guidance were identified (**Supplemental Figure 8**). KEGG analysis for the turquoise module showed enrichment in neurological pathways such as Alzheimer disease and metabolic pathways such as choline metabolism (**Supplemental Figure 8**).

2.4.4 X chromosome expression mosaicism in female cortical cell populations reveals dynamic non-cell-autonomous transcriptional homeostasis

To examine non-cell-autonomous effects, we considered all *Mecp2* expressing cells within *Mecp2e1^{-/+}* mosaic female cortical cells. Within *Mecp2e1^{-/+}* female snRNA-seq data, we identified 1,146 *Mecp2* expressing cells, of which 607 could be genotyped as WT-expressing and 539 were expressing the *Mecp2e1* mutation (**Supplemental Figure 9**). These cells were clustered based on expression and thus twelve different cell types were identified (**Figure 4a**). To reduce the impact of lower cell counts on DEG calling following parsing, we further grouped the *Mecp2* expressing cells into two broad categories of GABAergic neurons and glutamatergic neurons (**Figure 4b**). A summary of the number of DEGs in each of these three broader cell type categories (glutamatergic, GABAergic, non-neuronal) for each of the five experimental comparisons (**Fig 1d**) is shown in **Table 1**. LimmaVoom was used for DEG calling of experiments 3, 4 and 5. As expected based on random XCI, both cell populations (*Mecp2*_MUT and *Mecp2*_WT), were randomly represented in all cell types (**Figure 4a, b, c**). Cells from the males were also parsed, showing 184 *Mecp2*_MUT in the *Mecp2e1^{-/y}* and 175 *Mecp2*_WT in the *Mecp2e1^{+/y}* (**Supplemental Figure 9**).

In order to examine non-cell-autonomous effects over the disease progression in a cell category specific manner, we followed the third experiment design (**Figure 1d**) and compared the WT expressing cells from the *Mecp2e1^{-/+}* mosaic females to the WT expressing cells from the *Mecp2e1^{+/+}* females (**Figure 4d, Table 1**). At P30, both glutamatergic and GABAergic WT-expressing cells from *Mecp2e1^{-/+}* showed a large number of significant downregulated genes (blue) but a low number of upregulated genes (red), despite these cell populations being WT-expressing. These differences in gene expression were likely due to non-cell-autonomous effects of the *Mecp2e1* mutation on nearby WT-expressing cells. Further evidence was obtained from the experimental comparison from experiment 5, where mutant-expressing glutamatergic and GABAergic neurons from female *Mecp2e1^{-/+}* were compared to WT-expressing cells from female *Mecp2e1^{-/+}*, resulting in only 10 DEGs, compared to 2,216 in experiment 3 (non-cell-autonomous WT vs WT) (**Table 1**). This non-cell-autonomous effect was dynamic over time, as glutamatergic cells showed mostly upregulated genes with only a few downregulated genes at P60, while GABAergic cells only showed upregulated genes (**Figure 4d**). Interestingly, at the late disease stage P150, the number of dysregulated genes were diminished and primarily back to being downregulated, indicating a dynamic process of non-cell-autonomous effects across disease progression. In order to test this hypothesis, we overlapped the significant (LimmaVoom adjusted p-value ≤ 0.05) DEGs from each cell type and time point (**Figure 4e, 4f**). In glutamatergic cells, the largest overlap (129 DEGs) was between P30 and P60 and only 3 DEGs in common to all three time points (**Figure 4e**). Similar results were seen in GABAergic neurons, where DEGs were predominantly unique to each time point (**Figure 4f**).

In order to look for functional pathway enrichments of non-cell-autonomous effects of *Mecp2e1^{-/+}* mosaicism, KEGG analysis on significant DEGs (p-value ≤ 0.05) from glutamatergic and GABAergic neurons was performed and significant (adjusted p-value ≤ 0.05) terms overlapped across time (**Figure 4g, 4i**). Non-cell-autonomous DEGs from glutamatergic cells were enriched for 8 terms

that were shared across all disease stages which include Parkinson, Alzheimer, and Huntington diseases, as well as homeostatic pathways of retrograde endocannabinoid signaling, ubiquitin mediated proteolysis, oxidative phosphorylation, and protein processing in endoplasmic reticulum, while the P60 time point was uniquely enriched for terms such as GABAergic synapse and SNARE interactions in vesicular transport (**Figure 4h**). Further, glutamatergic cells showed molecular dysregulation associated with MeCP2 activity such as mRNA surveillance pathway, cholinergic synapse, and AMPK signaling pathway (**Figure 4h**). In contrast, GABAergic cells shared axon guidance as an enriched pathway common across all time points (**Figure 4i**). Other interesting RTT related pathways included metabolism and energy related terms such as riboflavin metabolism, phosphonate and phosphonate metabolism, choline metabolism, and alanine, aspartate, and glutamate metabolism (**Figure 4j**).

In order to compare these non-cell-autonomous effects to cell-autonomous effects over the disease progression in a cell category specific manner, we followed the fourth experiment design (**Figure 1d**) and compared the MUT *Mecp2* expressing cells from the *Mecp2^{1/+}* mosaic females to the WT *Mecp2* expressing cells from the *Mecp2^{1/+}* females (**Supplemental Figure 10a**). Similar to the results of experiment 3, glutamatergic and GABAergic significant DEGs were predominantly time point specific (**Supplemental Figure 10b-c**). Glutamatergic cells had 12 significant KEGG pathways shared over time while GABAergic cells had 6 significant terms both containing RTT related pathways such as mRNA surveillance and circadian rhythm (**Supplemental Figure 10d-g**). In order to examine if the dysregulated KEGG pathways are shared between experiment 3 and experiment 4, a comprehensive overlap test was performed showing majority of the pathways are unique to each experiment and each time, with the glutamatergic P150 KEGG pathways from non-cell-autonomous DEGs outnumbering those of cell-autonomous (17 in exp 3 vs 1 in exp 4) (**Supplemental Figure 11**).

Lastly, we examined non-cell-autonomous effects by comparing MUT-expressing to WT-expressing cells within the mosaic *Mecp2^{1/+}* females, as described in experiment 5 (**Figure 1d**).

Overall, glutamatergic and GABAergic had only a few genes dysregulated which were mostly at P150 when analyzed separately (**Supplemental Figure 12a**). For higher statistical power in KEGG term enrichment, DEGs glutamatergic and GABAergic cells were each combined across time points, revealing dysregulated retrograde endocannabinoid signaling and other pathways (**Supplemental Figure 12b-c**). The top10 enriched KEGG pathways when both glutamatergic and GABAergic cells were combined across all time points included pathways involved in cell signaling and addiction (**Supplemental Figure 12d**). The differences between WT-expressing and MUT-expressing cells within mosaic females in experiment 5 were far less than the differences between WT-expressing cells in mosaic *Mecp2e1*^{-/+} females compared to WT cells in *Mecp2e1*^{+/+} females in experiment 4. Together, these analyses demonstrate that transcriptional dysregulation across disease progression in mosaic *Mecp2e1*^{-/+} females is dynamic and predominated by non-cell-autonomous effects on homeostatic gene pathways.

2.4.5 Human RTT cortical cell transcriptional dysregulation is recapitulated by the female but not the male RTT mouse model

To examine how closely *Mecp2e1*^{-/+} mice phenocopy Rett syndrome (RTT) at the cellular transcriptome level, we examined the relationship between altered transcript levels by cell type in *Mecp2e1* deficient and human *MECP2*^{-/+} cortices. Thus sn-RNA seq analysis was performed on eight *MECP2*^{-/+} (RTT) and eight age matched control female cortex samples from post-mortem human brains (**Figure 5a**). Nine neuronal and six non-neuronal cell type clusters could be assigned from these human cortices based on 3,000 gene markers from the Bakken Trygve et al. dataset²⁰ (**Figure 5b**). Cell type labeling based on scTransform containing elevated expression of at least three cell marker genes was validated (**Figure 5c**). DEG analysis via limmaVoomCC compared *MECP2*^{-/+} to *MECP2*^{+/+} cortical cells, resulting in cell type-specific dysregulated genes (**Figure 5d**). Importantly, of the top 20 upregulated DEGs identified by LimmaVoomCC at the adjusted *p*-value ≤ 0.05 level in

female *MECP2*^{-/+} cortical cells or significant DEGs (adjusted *p* value ≤ 0.05) in *Mecp2e1*^{-/+} female mouse cortices with 14 gene transcripts out of 20 upregulated (**Figure 5e**). Similarly, of the top 20 Rett cortical LimmaVoomCC DEGs that were significantly downregulated (adjusted *p* value ≤ 0.05), the homologous *Mecp2e1*^{-/+} female gene transcripts were also downregulated (**Figure 5e**). In contrast, there were very few overlapping DEGs between human RTT and *Mecp2e1*^{-/-} male cortical cell transcriptomes (**Figure 5f**). This demonstrates that *Mecp2e1*^{-/+} female mice are a better model for the dynamic transcriptomic dysregulation due to cellular complexities in Rett syndrome disease progression.

2.5 Methods

2.5.1 *Mecp2-e1* mouse breeding and phenotyping

Mecp2e1^{-/+} female dams were bred with C57BL6/J wild-type male mice (Jackson Labs strain 000664) to produce *Mecp2e1*^{-/+} *Mecp2e1*^{+/+}, *Mecp2*^{-/-} and *Mecp2*^{+/+} littermates for single-nuclei RNA seq. Mice were weaned at 21 days post-natal and genotyped using DNA from tail snips as described⁴. Mouse disease scoring was performed blinded based on a seven-point scale with points assigned for matted fur, unusually large abdomen, skin lesions, hypoactivity and responses to tail suspension⁴. All animal procedures were approved by the UC Davis Institutional Animal Use and Care Committee (IACUC).

2.5.2 Single nuclei isolation for mouse and human post-mortem cortex

Mecp2-e1 and control mice were sacrificed by carbon dioxide inhalation just prior to brain removal. Cerebral cortex was removed from each brain from the mice. About 10mg of cerebral cortex tissue was isolated from human post-mortem and control samples. Single nuclei were prepared from the left hemisphere cortex according to a previously established protocol Martelotto (<https://cdn.10xgenomics.com/image/upload/v1660261285/support->

documents/CG000124_Demonstrated_Protocol_Nuclei_isolation_RevF.pdf). Briefly, a 3.0 mm² section of cortex was removed from each mouse brain. Both mouse and human brain tissue were minced with a scalpel then homogenized in 0.5 mls of nuclei lysis buffer with RNase inhibitor (Roche, Indianapolis, ID) then transferred to a larger tube with an additional 1.0 ml of nuclei lysis buffer, mixed then incubated on ice for 5 minutes. Nuclei were filtered from the lysate using a 70 µM FlowMi cell strainer (Sp-Belart, Wayne, NJ). Nuclei were pelleted at 4°C for 5 minutes at 500xG, resuspended in 1.5 ml of nuclei wash buffer, incubated for 5 minutes. Nuclei were then pelleted again as above then washed twice in nuclei wash and resuspension buffer then filtered with a 35 µM FlowMi filter (Sp-Belart, Wayne, NJ) then resuspended in nuclei wash and resuspension buffer with 5 µg/ml DAPI and assayed on a Countess cell counter to determine concentration and nuclear integrity (Fisher Scientific, Waltham, MA). Nuclei were then sorted to remove debris and nuclear aggregates on a MoFlow Astrios cell sorter (Beckman-Coulter, Brea, CA). Approximately 150,000 nuclei per sample were sorted and stored on ice prior to sn-RNA seq 5' library generation.

2.5.3 Single nuclei-RNA sequencing

Single Cell 5' Library & Gel Bead Kits (10x Genomics, Pleasanton, CA) were used to prepare cDNA and generate bar coded and indexed snRNA-seq 5' libraries according to the manufacturers protocol. 10,000 nuclei per sample were targeted. snRNA-seq 5' libraries were balanced using a Kapa library quantification kit (Roche, Indianapolis, IN) and pooled to generate 150 base pair, paired end sequences from using a NovaSeq S4 sequencer (Illumina, San Diego, CA). Mouse cortices had about 75,000 reads per cell on average and 240,437,728 reads per sample on average. Human cortices had about 50,000 reads per cell on average and 300,000,000 reads per sample on average.

2.5.4 Pre-processing and quality control

Cellranger v.2.0.2 was used to align the mouse raw reads to mm10-1.2.0 reference genome and the human raw reads to GRCh38 human reference genome. Cell by gene count matrices were used to

create a Seurat object using Seurat_4.3.0.1 in R 4.2.2. Mouse samples were filtered with the criteria that cells should have less than 7% mitochondrial, greater than 200 and less than 5,625 genes and greater than 208 and less than 16,300 UMI respectively.

2.5.5 Cell type identification by dimensionality reduction

The expression counts were log transformed and normalized via Seurat 4.3.0.1. Information about the samples such as sex, genotype, time point, disease score, body weight and *Mecp2e1* expression allele were all added to the metadata. Single cell mouse and human cortex data from the Allen brain institute were used as a reference for cell type labeling both data sets separately^{11,38}. scTransform was used to align cell types and transfer labels over to the Rett data. Cell marker test was performed for validating the cell type labeling. Dot plots showing validation of the cell type markers were created via scCustomize 2.1.1 (10.5281/zenodo.5706430).

2.5.6 DEG analysis

A total of five different DEG analysis methods were used to evaluate the best method for comparing mutant samples to WT samples in a cell-type-specific manner. EdgeR, Limma, and DeSeq2 yielded inconsistent DEGs (**Supplemental Figure 1**). For experiments 1 and 2, low expressing genes were filtered out. Low expressing was defined by expression in less than 25% of cells of a given cell type. LimmaVoomCC was used on the remaining high expressing genes to determine differentially expressed genes while considering cells from the same mouse will have correlated expression. For the low expressing genes, DEsingle was used for DEG analysis on genes that are not as robustly expressed (expressed in <25% of cells of type). For experiments 3, 4, and 5, LimmaVoom was exclusively used to identify differentially expressed genes. For each of the DEG experiments, the number of cells were normalized by down sampling. Parameters for all DEG analysis are available in the GitHub repository.

2.5.7 KEGG analysis

DEGs with a p-value of ≤ 0.05 from each of the experiments were used as the input for KEGG analysis. This was performed using the R package `enrichR` 3.2.

The 20 upregulated and downregulated genes were determined by `LimmaVoomCC` adjusted p value < 0.05 . We also included gene ontology analysis using the same DEGs.

2.5.8 hdWGCNA analysis

Cells from both males and females in the processed Seurat object were used as the input for hdWGCNA analysis. We also included phenotype data such as disease score. The criteria for the fraction of cells that a gene needs to be expressed in order to be included was set at 5%. The network type used is signed with a softpower of 0.8. A total of 9 modules were produced and scores for each module was computed using UCell method. Standard pipeline for hdWGCNA 0.2.4 were followed and the parameters are available in the GitHub repository.

2.5.9 WT and mutant cell parsing in mosaic female mouse cortices

All *Mecp2* reads were extracted from the raw fastq files generated from each individual sample. `abBLAST` 3.0 and `BWA` 0.7.17 mem were used in conjunction to extract *Mecp2* reads (`alleler.py`). The reference used for alignment was 100 bp of the *Mecp2* gene; 50 bp upstream of the exon1 start codon and 50 bp downstream. With the aligned reads, the number of mutant (TTG) and wild type (ATG) start codons were counted using `alleler.py`. Each read also contains the cell barcode and UMI information which was used to add the mutant cell and wild-type cell information back to the Seurat object as metadata.

2.5.10 Overlap of human and mouse DEGs and KEGG pathways

`LimmaVoomCC` DEGs from both human cell types and mouse cell types were filtered at adjusted p-value ≤ 0.05 . Significant human DEGs were overlapped with female mouse and male

mouse respectively. GeneOverlap 1.38.0 package was used to perform a Fisher's exact test to determine the significance of overlapped genes. The same overlap approach was performed to determine the significant overlapping mouse and human KEGG pathways.

2.5.11 Statistics and Reproducibility

There were 28 mouse cortical samples were collected at four different timepoints corresponding to three different disease stages ($n = 28$). Four different *Mecp2e1* genotypes were considered that include both sexes producing *Mecp2e1*^{-/+} *Mecp2e1*^{+/+}, *Mecp2*^{/y} and *Mecp2*^{+/y} littermates. At the P30 and P60 timepoints, there were two biological replicates of the WT and mutant animals for each sex. At P120, there were two WT and mutant males replicates while at P150, there were four WT and mutant replicates. For DEG analysis, for each gene in each sample, the DEG analysis utilizes the summarized count which follows a negative binomial distribution with mean equal to the multiplication of library size and relative abundance (the gene expression levels), and the variance for each gene is a function of the mean. Significant DEGs are called based on adjusted p.value of ≤ 0.05 which utilizes the Benjamini-Hochberg method to adjust for false discovery. Significant DEGs were used as the input for the KEGG pathway analysis. KEGG analysis uses fishers exact test to determine if the overlap of input DEGs with the background genes are significant (adjusted p.value of ≤ 0.05). Overlap tests of human and mouse DEGs and KEGG pathways were both performed used GeneOverlap which also utilizes a fisher's exact test to determine if the overlap of two gene sets are significant. hdWGCNA used Pearson correlation test to make comparisons of gene expression with the phenotypes. Significant correlations were determined based on adjusted p.value of ≤ 0.05 .

2.6 Data Availability

Data is available through NCBI SRA, accession number PRJNA1142998 and PRJNA1157887

2.7 Code Availability

The codes generated during this study are available in the GitHub repository:

(<https://github.com/osmansharifi/snRNA-seq-pipeline>).

2.8 Discussion

This study advances our understanding of RTT, offering insights into sex-specific, cell type-dependent, and disease stage-associated transcriptional dysregulation resulting from the cellular complexities related to the X-linked dominant inheritance of *MECP2/Mecp2* mutation. This longitudinal analysis of single cortical cell transcriptomes during the gradual progression of disease symptoms in the *Mecp2e1^{-/+}* mouse model of RTT provided several new findings critical to understanding and treatment of human RTT. First, we demonstrated that the female *Mecp2e1^{-/+}* mice are inherently different, not simply less severe, in their transcriptional dysregulation compared to mutant *Mecp2^{0/y}* males that completely lack *Mecp2e1*. Second, we identified transcriptionally dysregulated gene pathways across cell types in female *Mecp2e1^{-/+}* cortices that were significantly associated with progression of multiple disease phenotypes over time. Third, we showed that non-cell-autonomous effects in mosaic female *Mecp2e1^{-/+}* mice are responsible for the homeostatic gene pathway dysregulations observed dynamically over time. Lastly, and most important for translational relevance, we showed that female mosaic *Mecp2e1* mutant mice better recapitulate the transcriptional dysregulation observed in human RTT cortical cells than *Mecp2* null males and may help explain the complexities of progressive and regressive stages of disease in RTT girls.

The earliest studies examining the effect of MeCP2 levels on transcription in brain relied on bulk RNA-seq comparing male *Mecp2* null to wild-type controls yielded few differentially expressed gene (DEG) transcripts^{21–24}. However, bulk analyses of transcript levels in *Mecp2* null compared to *Mecp2* duplication mouse brain revealed 2582 altered DEGs in hypothalamus²⁵, 1180 DEG transcripts

in cerebellum²⁶, and 1060 DEGs in amygdala²⁷. Interestingly, analysis of transcripts in individual brain cell types yielded non-overlapping lists of DEGs suggesting that bulk tissue DEG analysis suffers from a “dilution effect” potentially masking DEGs²⁸. While these initial studies comparing *Mecp2* null to wild-type and *Mecp2* duplication control male brains can reveal gene targets of MeCP2 *in vivo*, *Mecp2*^{+/+} female mice are the relevant model for understanding RTT, where brain cell autonomous and non-autonomous effects require analysis of individual cells and cell types.

For autosomal genes, heterozygous mutations are expected to show reduced phenotypic severity than the homozygous state, but for X-linked genes, there is the added complication of random XCI that creates epigenetic mosaicism within cell populations. We were able to utilize sn-RNA seq 5' in the *Mecp2e1* mouse model to parse by both cortical cell type and mutation to improve understanding of transcriptional dysregulation in RTT. Our results can both help confirm certain aspects of previous bulk transcriptomic studies and help explain some of the prior discrepancies between bulk transcriptomic studies in RTT mouse models. A study using bulk RNA-seq on 7-week-old *Mecp2* null mice showed 48 genes upregulated and 32 genes down-regulated in pathways such as circadian entrainment that are consistent with our single nucleus data, despite the lower overall number of DEGs identified²⁹. We identified circadian entrainment as a homeostatic gene pathway dynamically dysregulated in *Pvalb* and *Sst* at P30, L5 excitatory neurons at P60, and astrocytes at P150. To date, there has been one prior study conducted using single nucleus RNA-seq in both RTT brain and a mouse model. Renthal *et al* used *Mecp2* null cortex at a single time point (8 weeks for males, 12-20 weeks for females) and compared to human RTT cortex with *MECP2* 255X¹⁶. Our studies were consistent in finding both up- and down-regulated genes with MeCP2 deficiency across cell types and in finding evidence for non-cell-autonomous gene dysregulation, but inconsistent in demonstrating a significant effect of MeCP2 deficiency on repression of long genes. Differences between the study designs, including genetic mouse model, time points, statistical approaches for DEGs, and single cell

technology (iDrops versus 5' V2 technology) could explain the discrepancies. We specifically designed the current study to overcome some of the prior technical limitations, including improvement in signal to noise ratio³⁰, higher number of genes detected per nucleus, higher UMI per nucleus, and higher number of cells analyzed³¹. Further, we used 3,000 marker genes from the Allen brain atlas cortex single nucleus dataset to label cell types (compared to one marker gene per cell) and used five different statistical approaches to robustly identify differentially expressed genes.

To understand how transcriptional dysregulation in RTT cortex was related to symptom progression, our study uniquely utilized a longitudinal study design and systems biology approaches to correlate networks of dysregulated gene expression patterns with disease phenotypes over time. Remarkably, these disease-relevant gene networks were not specific to individual cell types, but instead were enriched in pathways also dysregulated in neurodegenerative disorders and addiction pathways that regulate brain homeostasis across cell types, including metabolism, circadian entrainment, and retrograde endocannabinoid signaling. Previous studies had shown a link between MeCP2 and addiction³² that involve *Arc* and *Junb* transcription consistent with our results in *Mecp2e1^{-ly}* cortical cells³³ and KEGG pathways enriched in *Mecp2e1^{-/+}* cortical cells. A prior study showing that selective deletion of *Mecp2* from excitatory neurons had no effect on excitatory transmission but reduced inhibitory synapse numbers and neurotransmission in the somatosensory and prefrontal cortex³⁴ is consistent with our results showing a spread of dysregulated gene pathways from excitatory to inhibitory neurons. A more recent study investigating neurons and astrocytes found KEGG pathways such as calcium signaling pathway and Rap1 signaling pathway were enriched in RTT, consistent with our findings³⁵. While non-cell autonomous effects have been previously noted in RTT mouse models our comprehensive analyses of cellular transcriptomes over disease progression implicates these effects as a central and defining feature of transcriptional dysregulation in RTT mosaic females. Sun et al. argue that the abnormal morphologies of neurons and astrocytes in human RTT are caused by non-cell-

autonomous effects driven by altered gene expression and enriched energy related KEGG pathways consistent with our findings from experiment 2³⁵. Defects in signaling pathways suggests RTT disease progression is not caused exclusively by autonomous transcriptional changes in individual cells, but rather due to a failure of wild-type *MECP2* expressing cells to compensate for mutant *MECP2* expressing cells.

Since RTT in humans almost exclusively affects females, our results have important implications for translational medicine. First, pre-clinical models for testing new therapies should be female and construct-relevant, ideally modeling actual human RTT mutations. While male *Mecp2* null models provide important basic insights into MeCP2 function, we clearly demonstrate that the *Mecp2e1* deficient males do not recapitulate the transcriptional dysregulation observed in RTT human cortical cells as well as their female mutant littermates. Furthermore, the non-cell-autonomous dynamic waves of dysregulation in WT-expressing cortical neurons may help explain why human RTT symptoms appear gradually and are staggered in a series of regressions followed by plateaus. Our results showing that transcriptional dysregulation appears pre-symptomatic in female *Mecp2e1* mutant across multiple cortical cell types suggest that diagnosis and treatment should ideally begin as early as possible, potentially by including *MECP2* mutations in newborn screening panels. To date the only drug in the market for RTT is Trofinetide which is based on IGF-1 a growth factor previously used for diseases such as Laron syndrome and liver cirrhosis³⁶³⁷. The overlap with other neurologic disease pathways including oxidative phosphorylation suggests that some existing drugs for neurodegenerative disorders could potentially be repurposed to counteract some of the RTT non-cell-autonomous transcriptional dysregulations in pathways regulating homeostasis. Conversely, the molecular pathogenesis of RTT may provide insights for understanding epigenetic regulation of transcriptional homeostasis of gene pathways relevant to common neurodegenerative and addiction disorders.

2.9 References

- Amir, R. E. et al. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat Genet* 23, 185–188 (1999).
- Zhu, X. et al. Analysis of the parental origin of de novo MECP2 mutations and X chromosome inactivation in 24 sporadic patients with Rett syndrome in China. *J Child Neurol* 25, 842–848 (2010).
- Neul, J. L. et al. The array of clinical phenotypes of males with mutations in Methyl-CpG binding protein 2. *American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics* 180, 55–67 (2019).
- Yasui, D. H. et al. Mice with an isoform-ablating Mecp2exon 1 mutation recapitulate the neurologic deficits of Rett syndrome. *Hum Mol Genet* 23, 2447–2458 (2014).
- Guy, J., Hendrich, B., Holmes, M., Martin, J. E. & Bird, A. A mouse Mecp2-null mutation causes neurological symptoms that mimic Rett syndrome. *Nat Genet* 27, 322–326 (2001).
- Braunschweig, D., Simcox, T., Samaco, R. C. & LaSalle, J. M. X-chromosome inactivation ratios affect wild-type MeCP2 expression within mosaic Rett syndrome and Mecp2-/+ mouse brain. *Hum Mol Genet* 13, 1275–1286 (2004).
- Ballas, N., Lioy, D. T., Grunseich, C. & Mandel, G. Non-cell autonomous influence of MeCP2-deficient glia on neuronal dendritic morphology. *Nat Neurosci* 12, 311–317 (2009).
- Belichenko, P. V. et al. Widespread Changes in Dendritic and Axonal Morphology in Mecp2-Mutant Mouse Models of Rett Syndrome: Evidence for Disruption of Neuronal Networks. *J Comp Neurol* 514, 240–258 (2009).
- Vogel Ciernia, A. et al. MeCP2 isoform e1 mutant mice recapitulate motor and metabolic phenotypes of Rett syndrome. *Hum Mol Genet* 27, 4077–4093 (2018).
- Neier, K. et al. Sex disparate gut microbiome and metabolome perturbations precede disease progression in a mouse model of Rett syndrome. doi:10.1038/s42003-021-02915-3.

- Tasic, B. et al. Shared and distinct transcriptomic cell types across neocortical areas. *Nature* (2018) doi:10.1038/s41586-018-0654-5.
- Squair, J. W. et al. Confronting false discoveries in single-cell differential expression. *Nat Commun* 12, (2021).
- Mou, T., Deng, W., Gu, F., Pawitan, Y. & Nghia Vu, T. Reproducibility of Methods to Detect Differentially Expressed Genes from Single-Cell RNA Sequencing. doi:10.3389/fgene.2019.01331.
- Soneson, C. & Robinson, M. D. Bias, robustness and scalability in single-cell differential expression analysis. *Nat Methods* 15, 255–261 (2018).
- Liu, Y. et al. iDESC: identifying differential expression in single-cell RNA sequencing data with multiple subjects. *BMC Bioinformatics* 24, 1–20 (2023).
- Renthal, W. et al. Characterization of human mosaic Rett syndrome brain tissue by single-nucleus RNA sequencing. *Nat Neurosci* 21, 1670–1679 (2018).
- Morabito, S. et al. Single-nucleus chromatin accessibility and transcriptomic characterization of Alzheimer’s disease. *Nat Genet* doi:10.1038/s41588-021-00894-z.
- Morabito, S., Reese, F., Rahimzadeh, N., Miyoshi, E. & Swarup, V. hdWGCNA identifies co-expression networks in high-dimensional transcriptomics data. doi:10.1016/j.crmeth.2023.100498.
- Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. (2008) doi:10.1186/1471-2105-9-559.
- Bakken, T. E. et al. Comparative cellular analysis of motor cortex in human, marmoset and mouse. doi:10.1038/s41586-021-03465-8.
- Tudor, M., Akbarian, S., Chen, R. Z. & Jaenisch, R. Transcriptional Profiling of a Mouse Model for Rett Syndrome Reveals Subtle Transcriptional Changes in the Brain. *PNAS*.

- Nuber, U. A. et al. Up-regulation of glucocorticoid-regulated genes in a mouse model of Rett syndrome. doi:10.1093/hmg/ddi229.
- Kriaucionis, S. et al. Gene Expression Analysis Exposes Mitochondrial Abnormalities in a Mouse Model of Rett Syndrome. *Mol Cell Biol* 26, 5033–5042 (2006).
- Urdinguio, R. G. et al. Mecp2-Null Mice Provide New Neuronal Targets for Rett Syndrome. doi:10.1371/journal.pone.0003669.
- Chahrour, M. et al. MeCP2, a key contributor to neurological disease, activates and represses transcription. *Science* 320, 1224–1229 (2008).
- Ben-Shachar, S., Chahrour, M., Thaller, C., Shaw, C. A. & Zoghbi, H. Y. Mouse models of MeCP2 disorders share gene expression changes in the cerebellum and hypothalamus. *Hum Mol Genet* 18, 2431–2442 (2009).
- Samaco, R. C. et al. Loss of MeCP2 in aminergic neurons causes cell-autonomous defects in neurotransmitter synthesis and specific behavioral abnormalities. *Proc Natl Acad Sci U S A* 106, 21966–21971 (2009).
- Sugino, K. et al. Cell-type-specific repression by methyl-CpG-binding protein 2 is biased toward long genes. *Journal of Neuroscience* 34, 12877–12883 (2014).
- Osenberg, S. et al. Activity-dependent aberrations in gene expression and alternative splicing in a mouse model of Rett syndrome. *Proc Natl Acad Sci U S A* 115, E5363–E5372 (2018).
- Zilionis, R. et al. Single-cell barcoding and sequencing using droplet microfluidics. *Nature Publishing Group* (2016) doi:10.1038/nprot.2016.154.
- Zhang, X. et al. Comparative Analysis of Droplet-Based Ultra-High-Throughput Single-Cell RNA-Seq Systems. *Mol Cell* 73, 130–142.e5 (2019).
- Deng, J. V. et al. MeCP2 in the nucleus accumbens contributes to neural and behavioral responses to psychostimulants. *Nature Publishing Group* (2010) doi:10.1038/nn.2614.

- Su, D., Cha, Y. M. & West, A. E. Mutation of MeCP2 alters transcriptional regulation of select immediate-early genes. *Epigenetics* 7, 146–154 (2012).
- Zhang, W., Peterson, M., Beyer, B., Frankel, W. N. & Zhang, Z. W. Loss of MeCP2 from forebrain excitatory neurons leads to cortical hyperexcitation and seizures. *Journal of Neuroscience* 34, 2754–2763 (2014).
- Sun, J. et al. Mutations in the transcriptional regulator MeCP2 severely impact key cellular and molecular signatures of human astrocytes during maturation. *Cell Rep* 42, 111942 (2023).
- Neul, J. L. et al. Trofinetide for the treatment of Rett syndrome: a randomized phase 3 study. *Nature Medicine* 29, 1468–1475 (2023).
- Puche, J. E. & Castilla-Cortázar, I. Human Conditions of Insulin-like Growth Factor-I (IGF-I) Deficiency. *Journal of Translational Medicine* vol. 10 <http://www.translational-medicine.com/content/10/1/224> (2012).
- Yao, Z. et al. A taxonomy of transcriptomic cell types across the isocortex and hippocampal formation. (2021) doi:10.1016/j.cell.2021.04.021.

2.10 Main Figures and Tables

2.10.1 Main Tables

Table 1. Summary of LimmaVoom DEG numbers resulting from the parsing of mutant- and WT-expressing cell type categories in all comparison experiments

Exp #	Cell type	#UP DEGs	#DOWN DEGs	Time point	Animal (n)	Total sig DEGs	Notes
1	Glutamatergic	7	0	P30	4	7	Comparing all cortical glutamatergic cells between mutant male <i>Mecp2e1-ly</i> to WT male <i>Mecp2e1+/y</i>
1	GABAergic	2	0	P30	4	2	Comparing all cortical GABAergic cells between mutant male <i>Mecp2e1-ly</i> to WT male <i>Mecp2e1+/y</i>
1	Non-neuronal	0	0	P30	4	0	Comparing all cortical non-neuronal cells between mutant male <i>Mecp2e1-ly</i> to WT male <i>Mecp2e1+/y</i>
1	Glutamatergic	0	27	P60	4	27	Comparing all cortical glutamatergic cells between mutant male <i>Mecp2e1-ly</i> to WT male <i>Mecp2e1+/y</i>

1	GABAergic	4	1	P60	4	5	Comparing all cortical GABAergic cells between mutant male <i>Mecp2e1-ly</i> to WT male <i>Mecp2e1+ly</i>
1	Non-neuronal	2	1	P60	4	3	Comparing all cortical non-neuronal cells between mutant male <i>Mecp2e1-ly</i> to WT male <i>Mecp2e1+ly</i>
1	Glutamatergic	21	0	P120	4	21	Comparing all cortical glutamatergic cells between mutant male <i>Mecp2e1-ly</i> to WT male <i>Mecp2e1+ly</i>
1	GABAergic	121	1	P120	4	126	Comparing all cortical GABAergic cells between mutant male <i>Mecp2e1-ly</i> to WT male <i>Mecp2e1+ly</i>
1	Non-neuronal	8	0	P120	4	8	Comparing all cortical non-neuronal cells between mutant male <i>Mecp2e1-ly</i> to WT male <i>Mecp2e1+ly</i>
1	Total	165	30			195	
2	Glutamatergic	31	913	P30	4	944	Comparing all cortical glutamatergic cells between mutant female <i>Mecp2e1-/+</i> to WT female <i>Mecp2e1+/+</i>
2	GABAergic	236	35	P30	4	271	Comparing all cortical GABAergic cells between mutant female <i>Mecp2e1-/+</i> to WT female <i>Mecp2e1+/+</i>
2	Non-neuronal	0	0	P30	4	0	Comparing all cortical non-neuronal cells between mutant female <i>Mecp2e1-/+</i> to WT female <i>Mecp2e1+/+</i>
2	Glutamatergic	13	11	P60	4	24	Comparing all cortical glutamatergic cells between mutant female <i>Mecp2e1-/+</i> to WT female <i>Mecp2e1+/+</i>
2	GABAergic	0	0	P60	4	0	Comparing all cortical GABAergic cells between mutant female <i>Mecp2e1-/+</i> to WT female <i>Mecp2e1+/+</i>
2	Non-neuronal	0	0	P60	4	0	Comparing all cortical non-neuronal cells between mutant female <i>Mecp2e1-/+</i> to WT female <i>Mecp2e1+/+</i>
2	Glutamatergic	0	0	P150	8	0	Comparing all cortical glutamatergic cells between mutant female <i>Mecp2e1-/+</i> to WT female <i>Mecp2e1+/+</i>
2	GABAergic	2	0	P150	8	2	Comparing all cortical GABAergic cells between mutant female <i>Mecp2e1-/+</i> to WT female <i>Mecp2e1+/+</i>
2	Non-neuronal	0	0	P150	8	0	Comparing all cortical non-neuronal cells between mutant female <i>Mecp2e1-/+</i> to WT female <i>Mecp2e1+/+</i>
2	Total	282	959			1241	
3	Glutamatergic	17	782	P30	4	799	Comparing WT-expressing glutamatergic cells from mutant female <i>Mecp2e1-/+</i> to WT-expressing cells from WT female <i>Mecp2e1+/+</i>
3	GABAergic	2	393	P30	4	395	Comparing WT-expressing GABAergic cells from mutant female <i>Mecp2e1-/+</i> to WT-expressing cells from WT female <i>Mecp2e1+/+</i>
3	Glutamatergic	835	5	P60	4	840	Comparing WT-expressing glutamatergic cells from mutant female <i>Mecp2e1-/+</i> to WT-expressing cells from WT female <i>Mecp2e1+/+</i>
3	GABAergic	5	0	P60	4	5	Comparing WT-expressing GABAergic cells from mutant female <i>Mecp2e1-/+</i> to WT-expressing cells from WT female <i>Mecp2e1+/+</i>

3	Glutamatergic	2	41	P150	8	43	Comparing WT-expressing glutamatergic cells from mutant female <i>Mecp2e1-/+</i> to WT-expressing cells from WT female <i>Mecp2e1+/+</i>
3	GABAergic	1	43	P150	8	44	Comparing WT-expressing GABAergic cells from mutant female <i>Mecp2e1-/+</i> to WT-expressing cells from WT female <i>Mecp2e1+/+</i>
3	Total	862	1264			2126	
4	Glutamatergic	3	574	P30	4	577	Comparing mutant-expressing glutamatergic cells from female <i>Mecp2e1-/+</i> to WT-expressing cells from female WT <i>Mecp2e1+/+</i>
4	GABAergic	9	984	P30	4	993	Comparing mutant-expressing GABAergic cells from female <i>Mecp2e1-/+</i> to WT-expressing cells from female WT <i>Mecp2e1+/+</i>
4	Glutamatergic	1251	8	P60	4	1259	Comparing mutant-expressing glutamatergic cells from female <i>Mecp2e1-/+</i> to WT-expressing cells from female WT <i>Mecp2e1+/+</i>
4	GABAergic	0	0	P60	4	0	Comparing mutant-expressing GABAergic cells from female <i>Mecp2e1-/+</i> to WT-expressing cells from female WT <i>Mecp2e1+/+</i>
4	Glutamatergic	4	8	P150	8	12	Comparing mutant-expressing glutamatergic cells from female <i>Mecp2e1-/+</i> to WT-expressing cells from female WT <i>Mecp2e1+/+</i>
4	GABAergic	0	4	P150	8	4	Comparing mutant-expressing GABAergic cells from female <i>Mecp2e1-/+</i> to WT-expressing cells from female WT <i>Mecp2e1+/+</i>
4	Total	1267	1578	0		2845	
5	Glutamatergic	0	0	P30	2	0	Comparing mutant-expressing glutamatergic cells from female <i>Mecp2e1-/+</i> to WT-expressing cells from female <i>Mecp2e1-/+</i>
5	GABAergic	0	0	P30	2	0	Comparing mutant-expressing GABAergic cells from female <i>Mecp2e1-/+</i> to WT-expressing cells from female <i>Mecp2e1-/+</i>
5	Glutamatergic	3	0	P60	2	3	Comparing mutant-expressing glutamatergic cells from female <i>Mecp2e1-/+</i> to WT-expressing cells from female <i>Mecp2e1-/+</i>
5	GABAergic	0	0	P60	2	0	Comparing mutant-expressing GABAergic cells from female <i>Mecp2e1-/+</i> to WT-expressing cells from female <i>Mecp2e1-/+</i>
5	Glutamatergic	6	0	P150	4	6	Comparing mutant-expressing glutamatergic cells from female <i>Mecp2e1-/+</i> to WT-expressing cells from female <i>Mecp2e1-/+</i>
5	GABAergic	1	0	P150	4	1	Comparing mutant-expressing GABAergic cells from female <i>Mecp2e1-/+</i> to WT-expressing cells from female <i>Mecp2e1-/+</i>
5	Total	10	0	0		10	

2.10.2 Main Figures

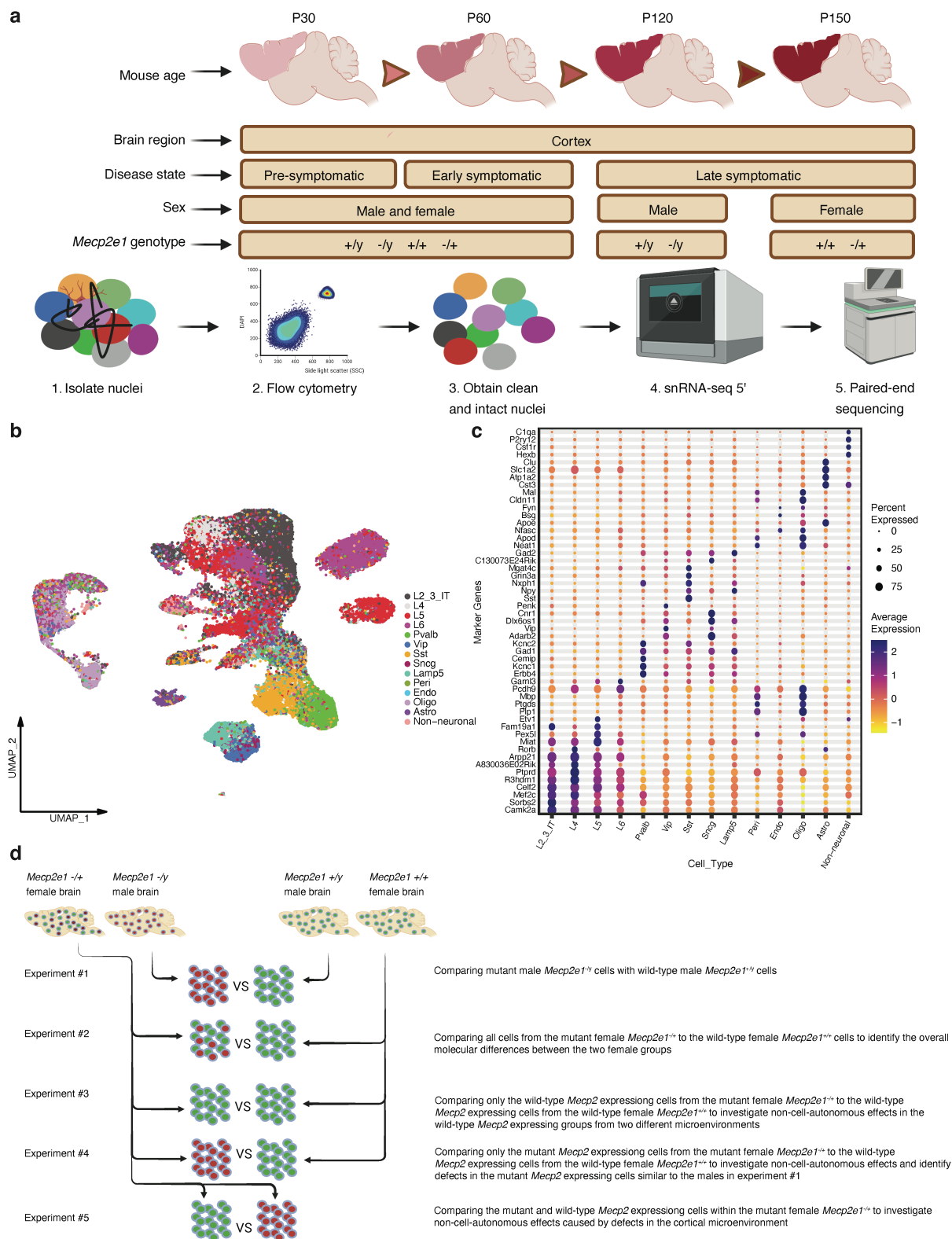


Figure 3: A scheme showing the overall mouse study design. **a.** Cortical samples were collected from postnatal mice at four different timepoints corresponding to three different disease stages (n = 28). Four different Mecp2e1 genotypes were considered that include both sexes. **b.** UMAP of the unsupervised clustering of cell types (n = 93,798 cells post QC) identified. Cell type labels were transferred from³⁹ Van, Yao et al. 2021. **c.** Top gene markers for each cell type are shown on y-axis. The color refers to the average expression of genes in a cell type and the percent expressed describes the percentage of cells within a cell type that express each gene marker. **d.** Design of computational experiments comparing mutant to WT cells from mice of both sexes. Experiments 3 to 5 are comparing subtypes of cells in females due to X chromosome inactivation to examine potential non-cell-autonomous effects of Mecp2e1 mutation.

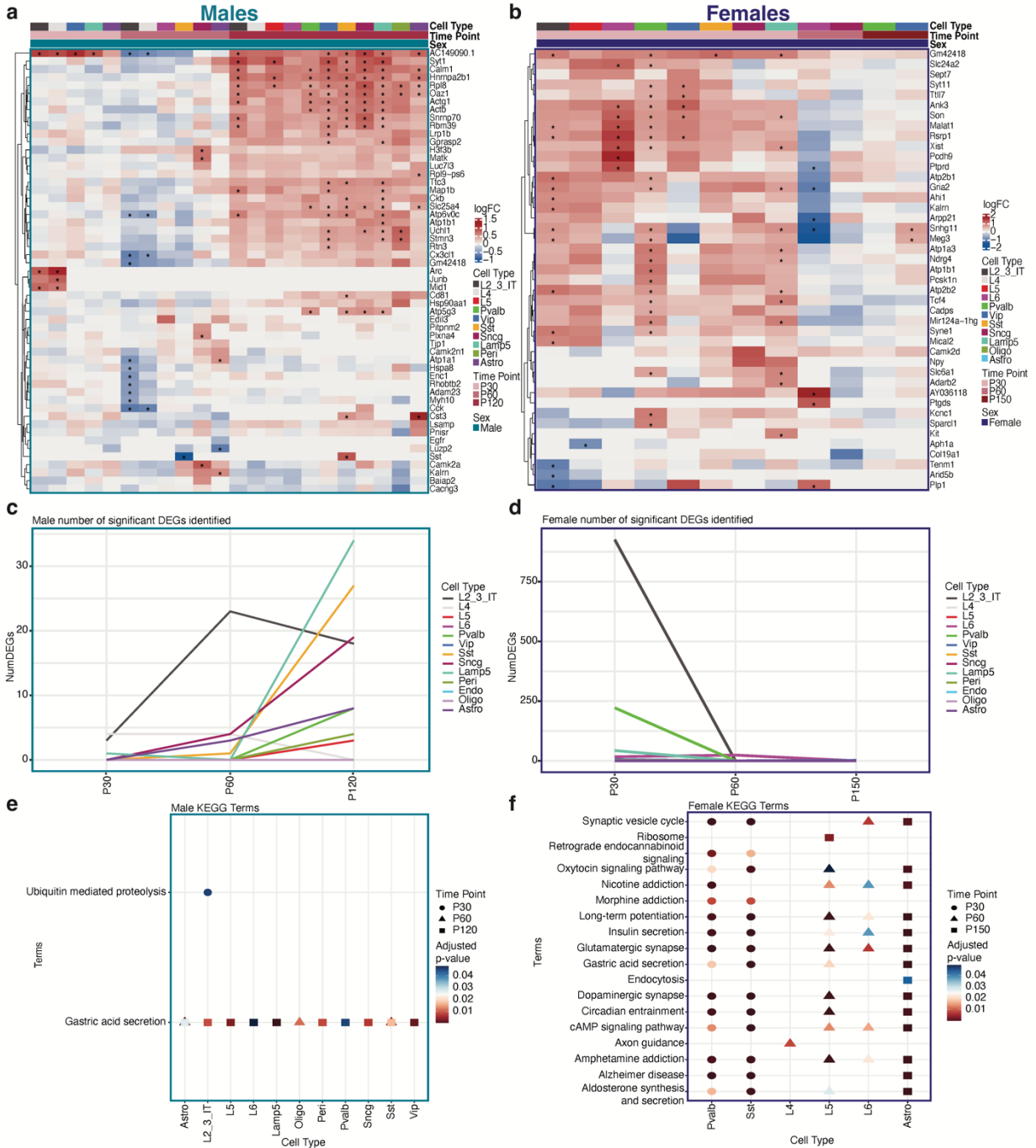


Figure 2. Sexually-dimorphic dynamic patterns of DEGs and KEGG pathway terms across time and cell type. **a.** Heatmap of top 5 differentially expressed genes (DEGs) based on the lowest adjusted p-value ≤ 0.05 comparing male *Mecp2e1*^{-/-} and *Mecp2e1*^{+/-} cortical cells across timepoints (experiment 1). **b.** Heatmap of top 5 DEGs comparing female *Mecp2e1*^{-/-} and *Mecp2e1*^{+/-} cortical cells across timepoints (experiment 2). *indicates adjusted p-value ≤ 0.05 (corrected via Benjamini and Hochberg method). **c, d.** Number of DEGs over time at adjusted p-value ≤ 0.05 for experiments 1 and 2, respectively. **e, f.** Dot plots showing the KEGG pathway terms for DEGs (adjusted p-value ≤ 0.1) from

each cell type, selected for terms that are persistent over time for experiments 1 and 2, respectively.

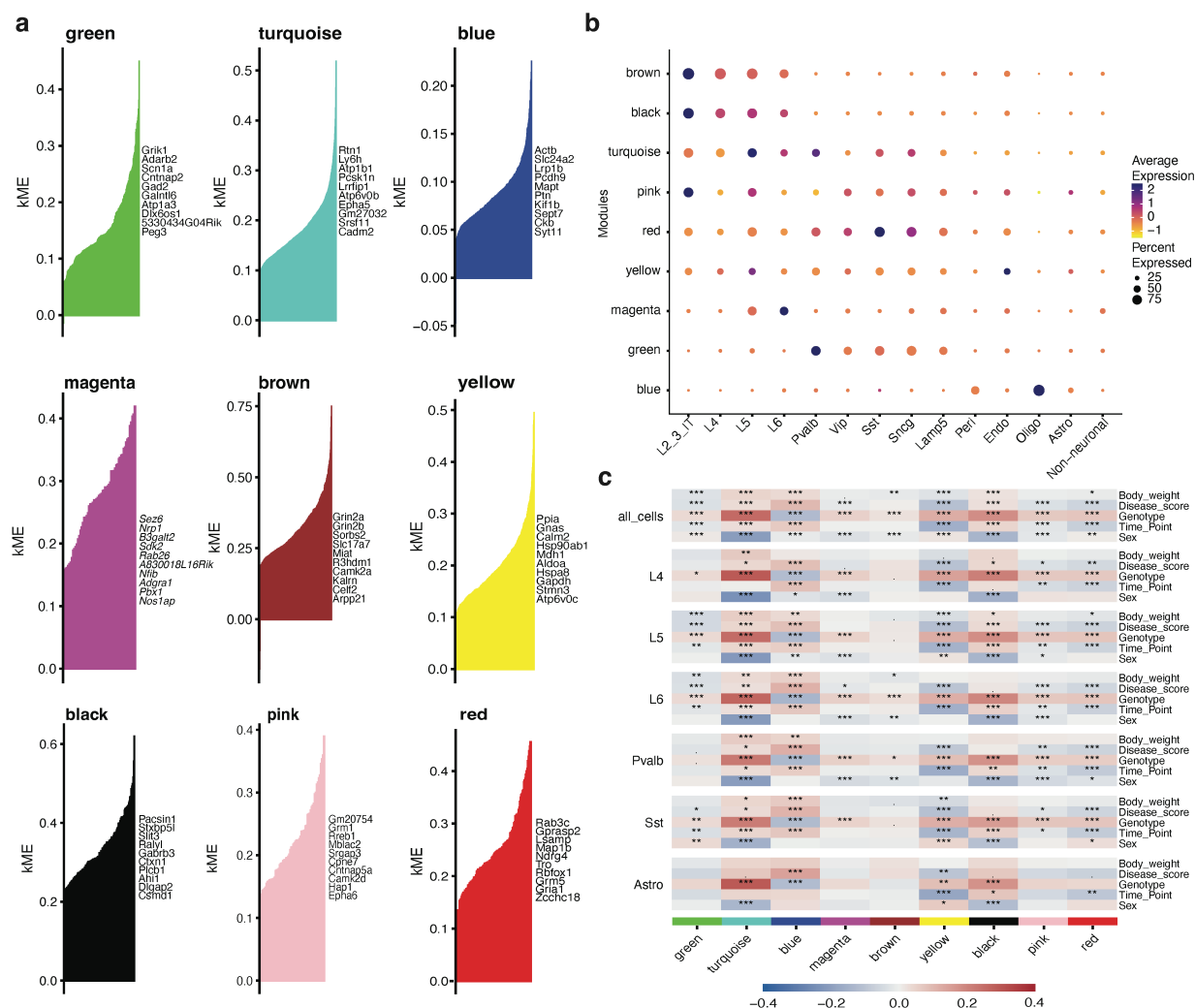


Figure 3. hdWGCNA identifies co-expression networks for each cell type in the mouse cortex that correlated with *Mecp2e1* genotype, disease phenotypes, and sex. a. Top 10 hub genes identified for each of the 9 modules generated by hdWGCNA on entire snRNA-seq dataset, identified by color. The x-axis are all the genes in each module and the y-axis is the corresponding kME value. **b.** Dot plot of the average gene expression of the top 10 hub genes in each module generated for each cortical cell type. **c.** A heat map of correlations between experimental phenotypes and variables (body weight, disease score, genotype, time point, sex) and averaged gene expression (eigenmode value) for each cell type (cell types not shown are in **Supplemental Figure 7a**). *, **, *** indicates FDR-corrected p-value ≤ 0.05 , 0.01, and 0.001, respectively. The color bar shows the Pearson correlation coefficient.

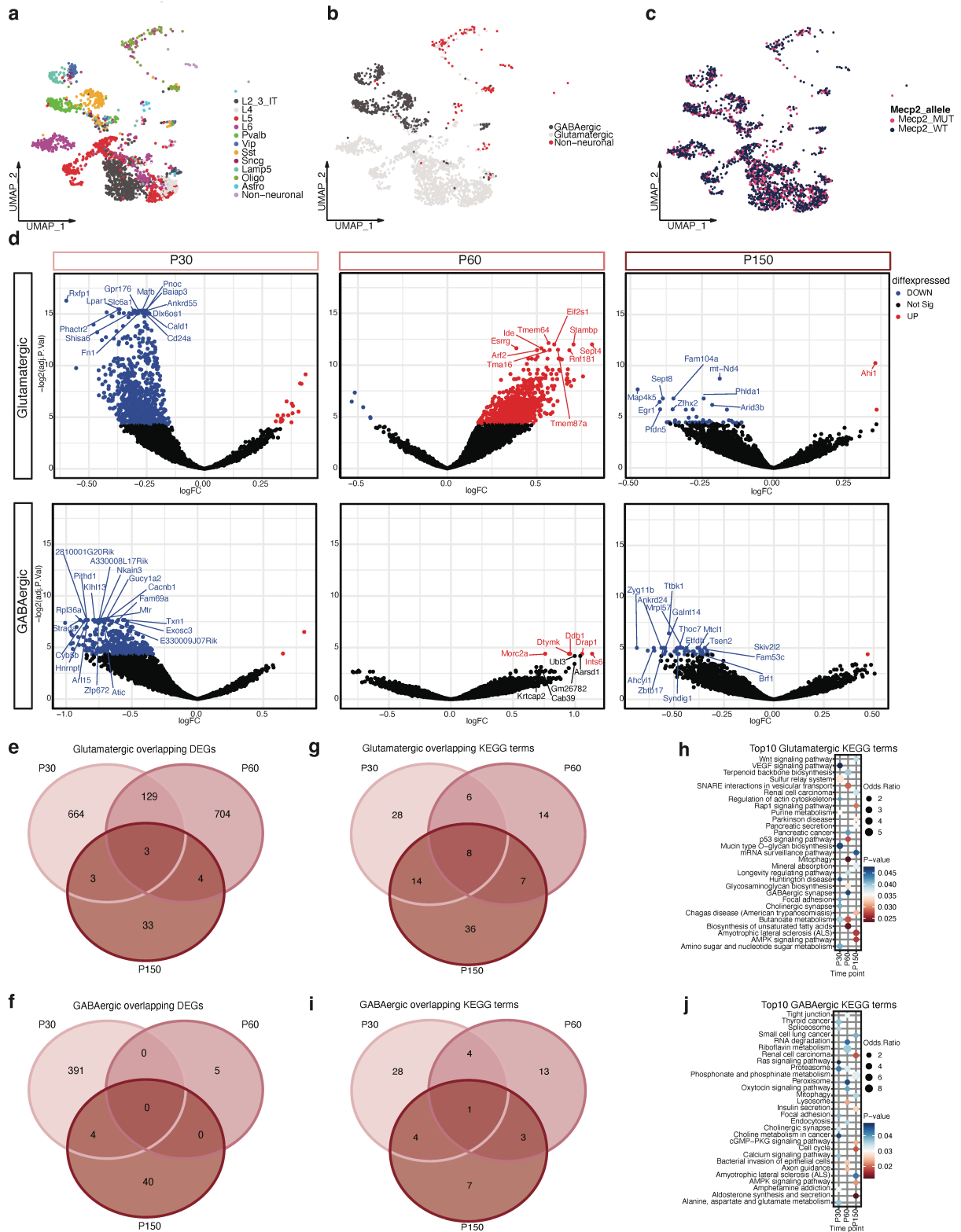


Figure 4. Dynamic non-cell-autonomous effects on differentially expressed genes and KEGG pathways over disease progression. As shown in Experiment #3 (Fig. 1d), we compared WT cells from *Mecp2e1^{-/-}* female with WT cells from *Mecp2e1^{+/-}* glutamatergic

and GABAergic neurons longitudinally. **a.** UMAP plot of cell types identified in the mosaic females **b.** UMAP plot of the female cortices showing the clustering of the broad cell type categories. **c.** UMAP plot of mosaic female cells parsed by *Mecp2* allele **d.** Volcano plots showing differentially expressed genes (DEGs) of the mouse cortical neurons contrasting WT cells from WT *Mecp2*^{+/+} females and WT cells from *Mecp2*^{-/-} mosaic females. **e,f.** Venn diagrams of overlapping glutamatergic and GABAergic DEGs respectively over time. **g,i.** Venn diagrams of significant KEGG terms of glutamatergic and GABAergic neurons over time. **H.** Top 10 KEGG terms of glutamatergic neurons over time. **j** Top 10 KEGG terms of GABAergic neurons over time.

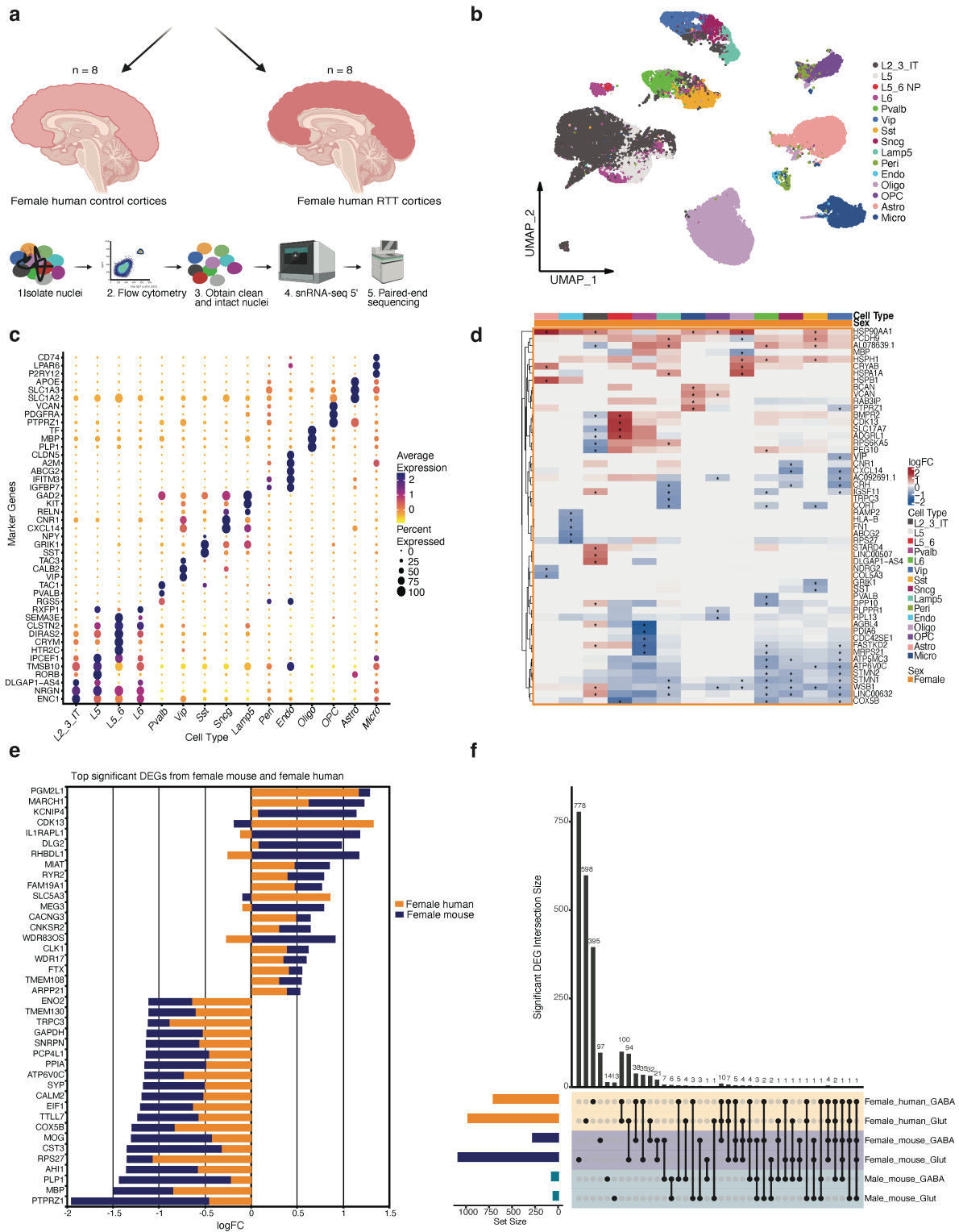


Figure 5. Human RTT cortical neurons share transcriptional dysregulation specifically with *Mecp2e1*^{+/−} mosaic female mice. **a.** A schematic of postmortem human RTT cortices and age/sex matched control cortices. **b.** UMAP of the unsupervised clustering of cell types identified in the human cortices (n = 39336 cells post QC). Cell type labels were transferred

from Bakken Trygve et al. 2021⁴⁰. **c.** Top gene markers for each cell type in the human cortex. **d.** Heatmap of top differentially expressed genes (DEGs) for human female cortices. *indicates adjusted p-value ≤ 0.05 . **e.** Bar graph showing overlapping of the top significant upregulated and downregulated genes by logFC in female mouse and female human. **f** Upset plot showing overlap of the significant DEGs from both GABAergic and glutamatergic neurons in female human, female mouse, and male mouse.

2.11 Supplemental Figures and Supplemental Tables

2.11.1 Supplemental Tables

Supplemental table 1 and 2 are too large to include in this dissertation. They can be found in BioRxiv and/or the published manuscript.

Supplemental Table 1. Table containing all significant up and down regulated DEGs

(LimmaVoomCC) from experiment 1 and 2. Table contains gene name, logFC, adjusted p-value, sex, cell type and time point information.

Supplemental Table 2. Table containing all significant KEGG pathways from experiment 1 and 2. Table contains Term, overlap, odds ratio, adjusted p-value, Genes contained in the pathway, sex, cell type and metadata information.

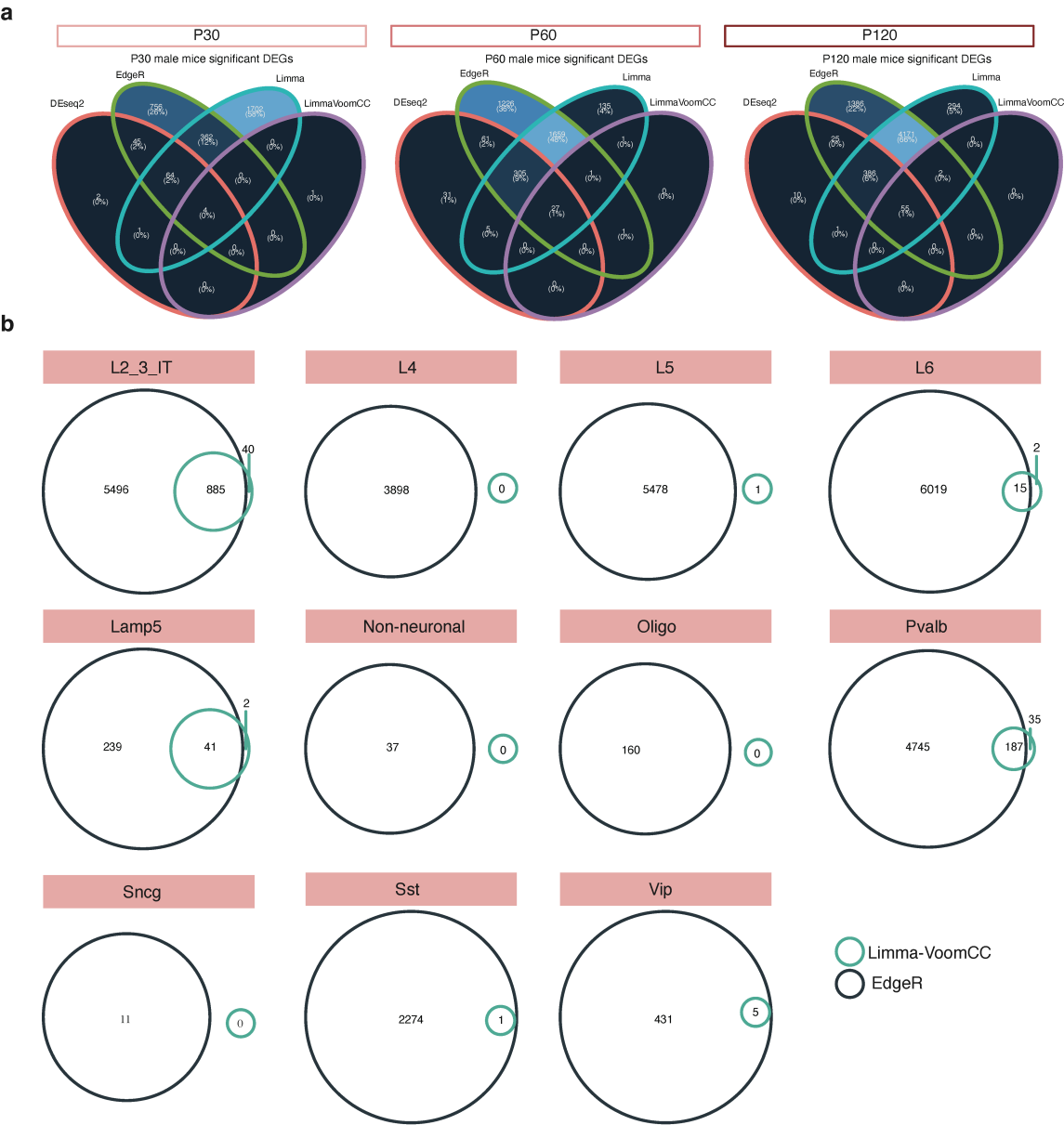
Supplemental Table 3. Human postmortem brain sample characteristics.

number	disease	sex	age in years	region	ethnicity	mutation
5125	CTRL	F	24	Cortex		NA
5161	CTRL	F	10	Cortex		NA
4591	CTRL	F	16	Cortex		NA
812	CTRL	F	18	Cortex	aa	NA
662	CTRL	F	12	Cortex		NA
1136	CTRL	F	33	Cortex		NA
1711	CTRL	F	27	Cortex	ca	NA
1846	CTRL	F	21	Cortex		NA
		AVERAGE	20.125			
RTT 4687	RTT	F	8	Cortex		MECP2 p.R255X
RTT 5075	RTT	F	20	Cortex		none detected

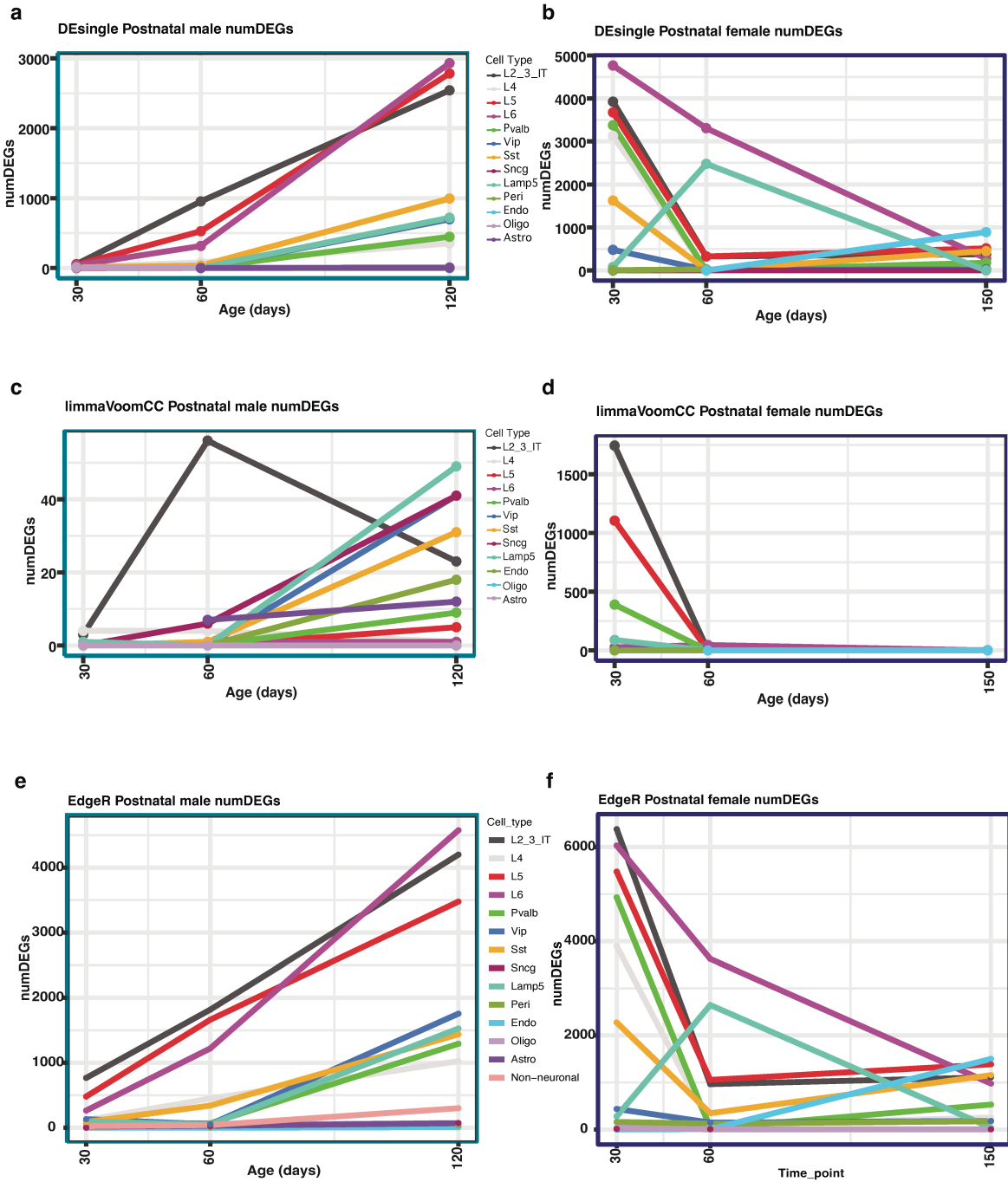
RTT 1420	RTT	F	21	Cortex		none detected
RTT 1815	RTT	F	18	Cortex	ca	MECP2 c.378-2 A>G
RTT 1748	RTT	F	22	Cortex		none detected
RTT 3381	RTT	F	22	Cortex		none detected
RTT 4852	RTT	F	19	Cortex	ca	MECP2 c.G451
RTT B5020	RTT	F	24	Cortex		MECP2 p.R255X
		AVERAGE	19.25			

2.12.1 Supplemental Figures

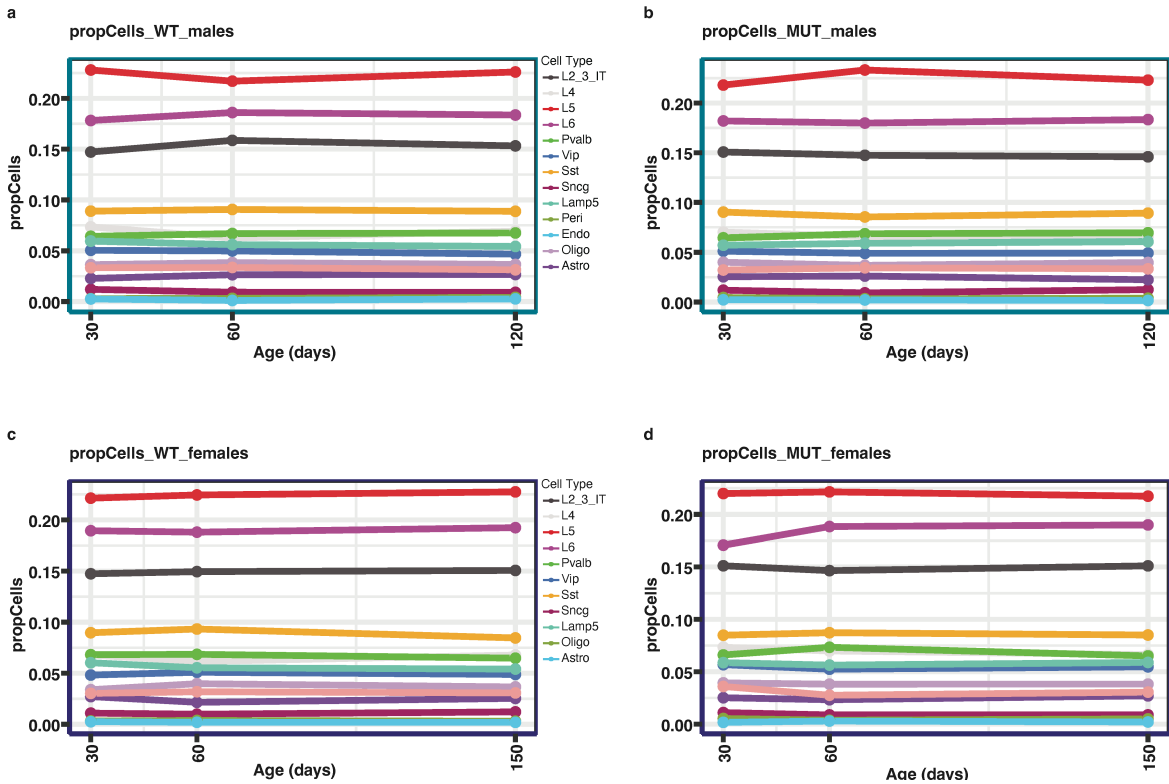
Supplemental fig.1: Identification of significant DEGs using Limma, EdgeR and DESeq2 produced few common genes from male sn-RNA seq data



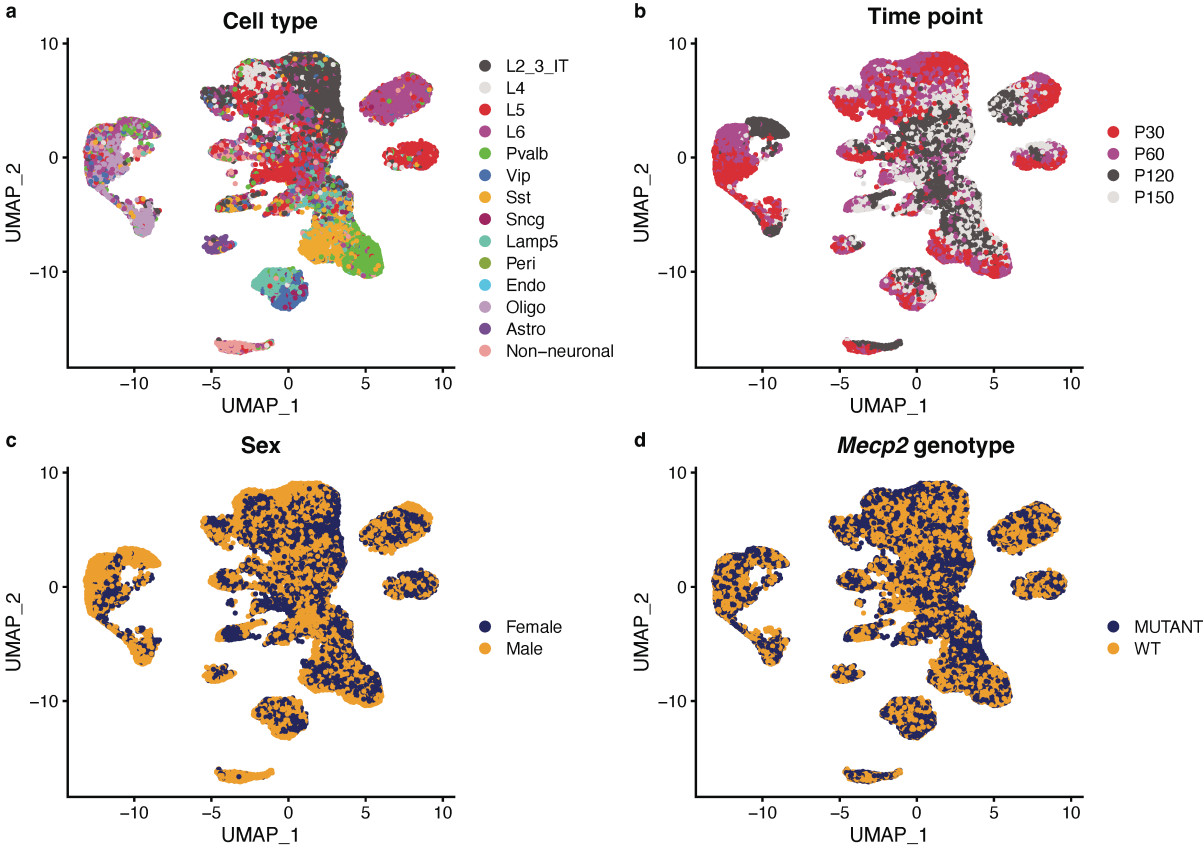
Supplemental fig.2: Line graph of the number of post-natal DEGs identified using DESingle from *Mecp2e1^{-/-}* female and *Mecp2e1^{-y}* cortices



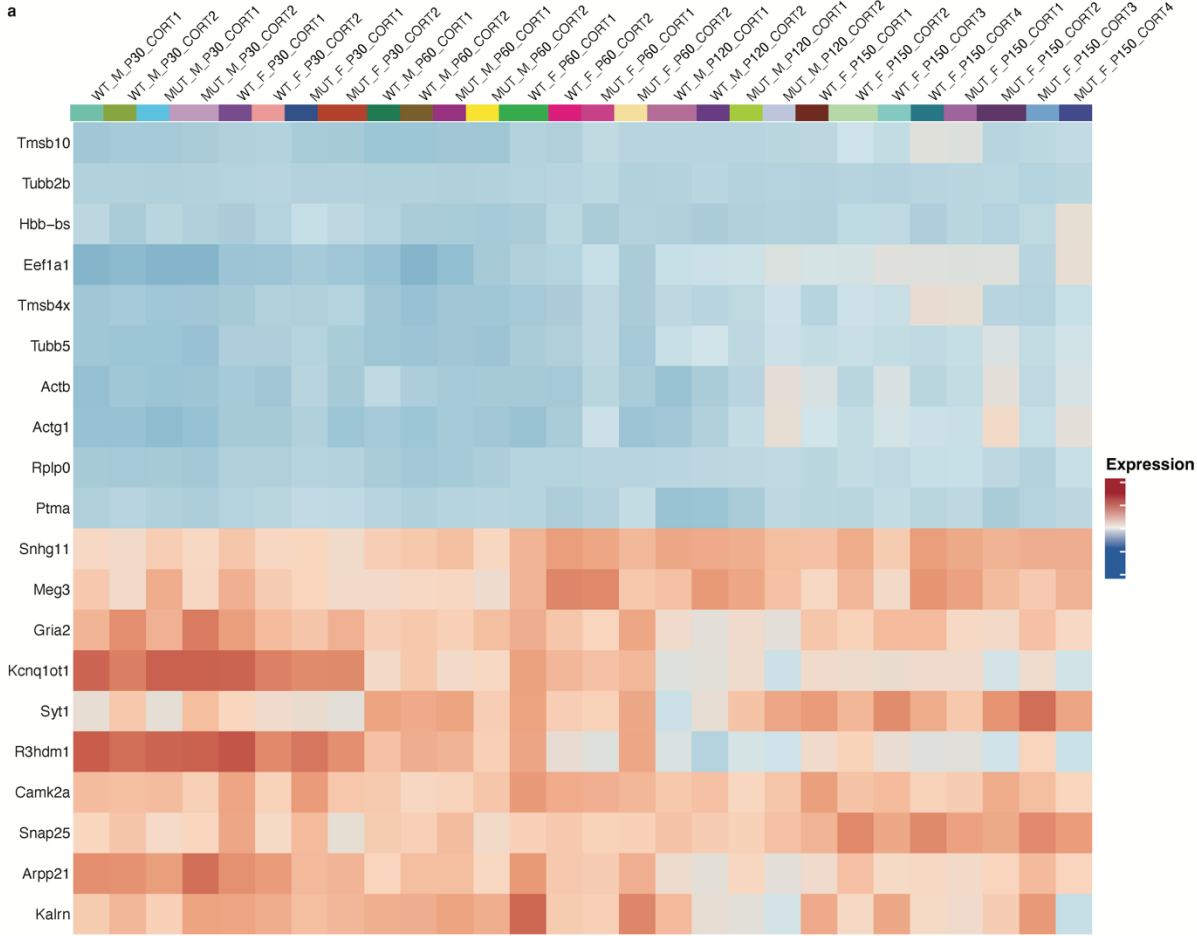
Supplemental fig.3: *Mecp2e1* mutant cortices have normal proportions of cortical cell types



Supplemental fig.4: Cell clustering does not differ between sexes and *Mecp2* genotypes in postnatal cortex

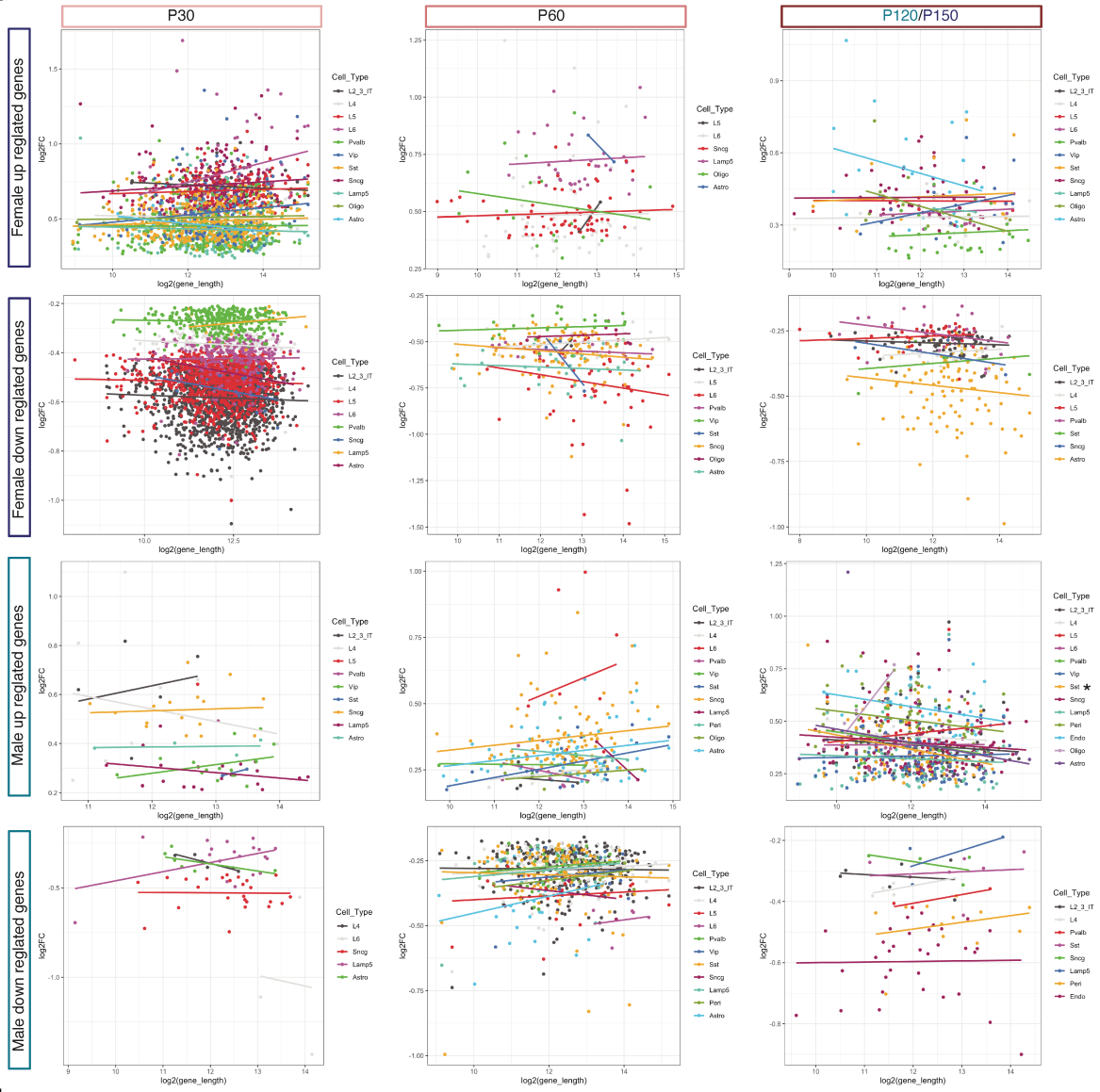


Supplemental fig.5: Replicate mouse brain samples have similar patterns of cell type-specific genes expressed in common

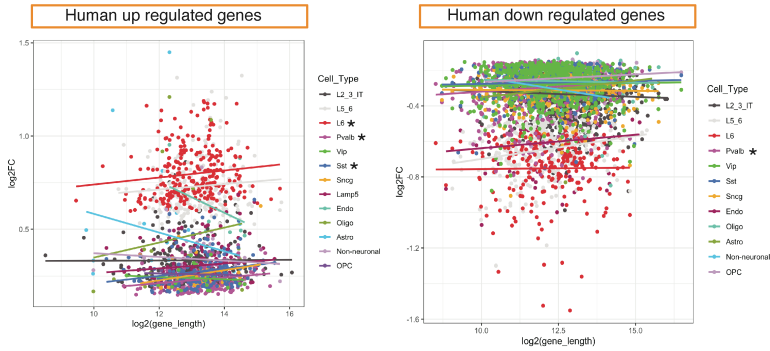


Supplemental fig.6: Differentially expressed genes are not correlated with gene length

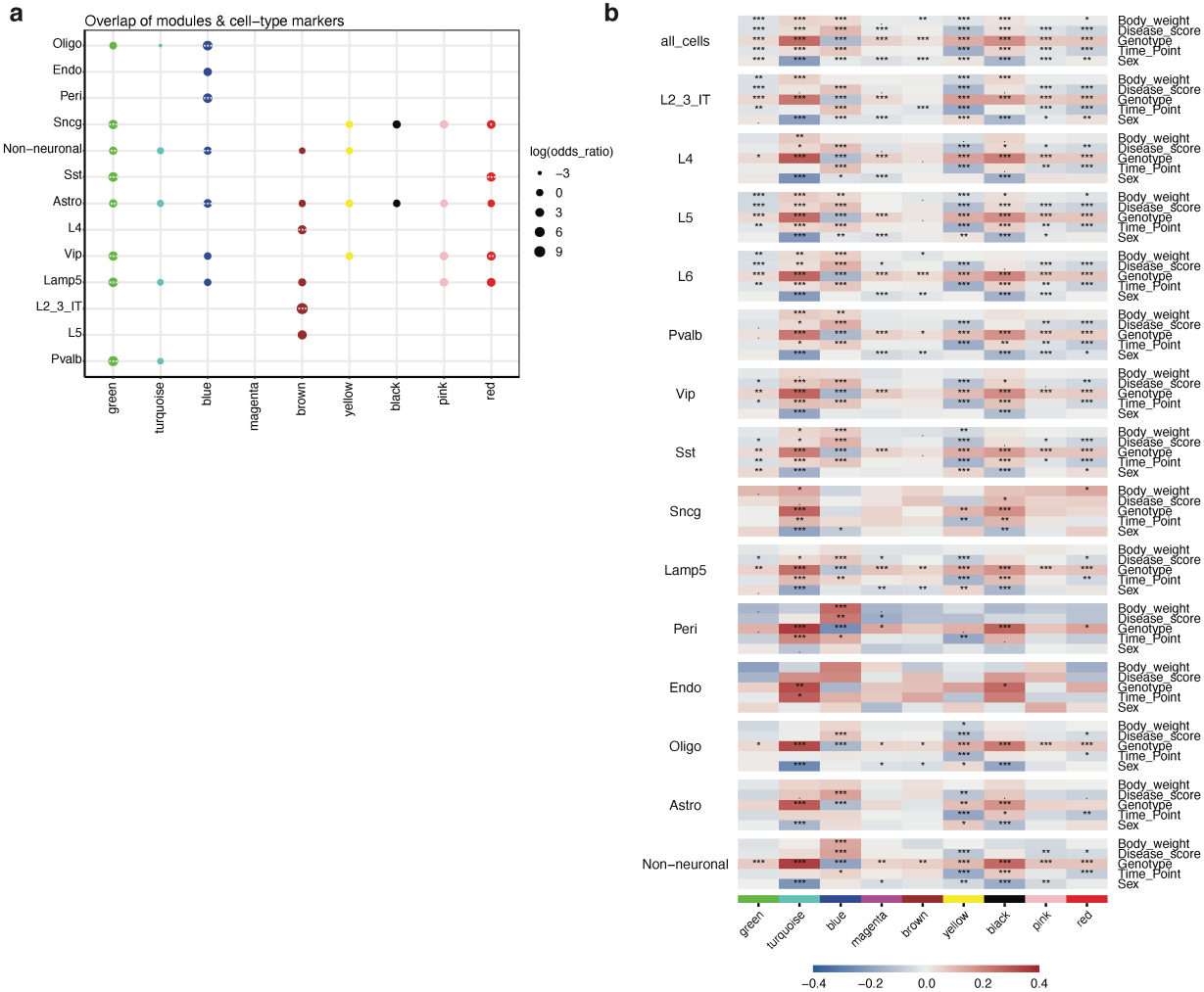
a



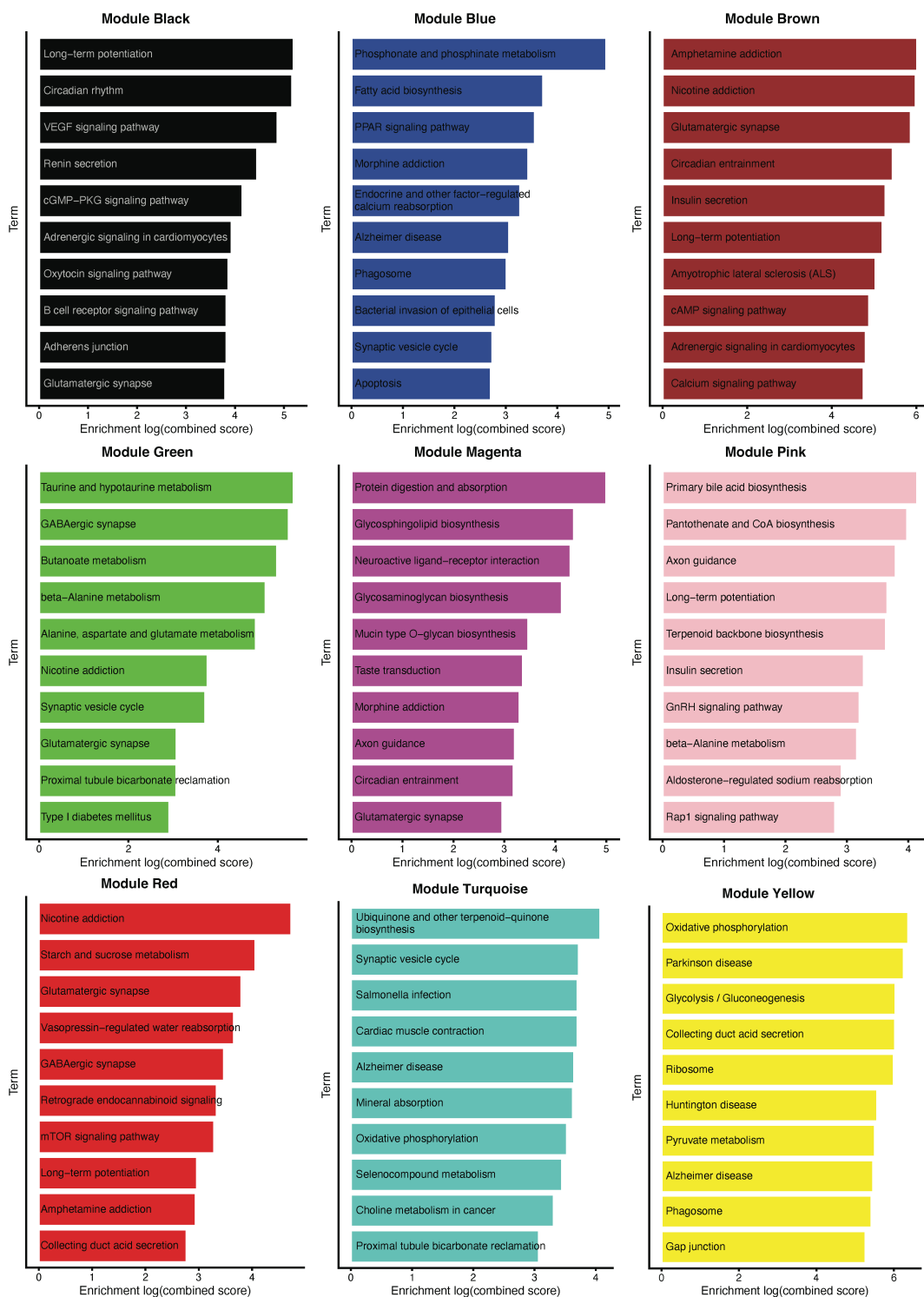
b



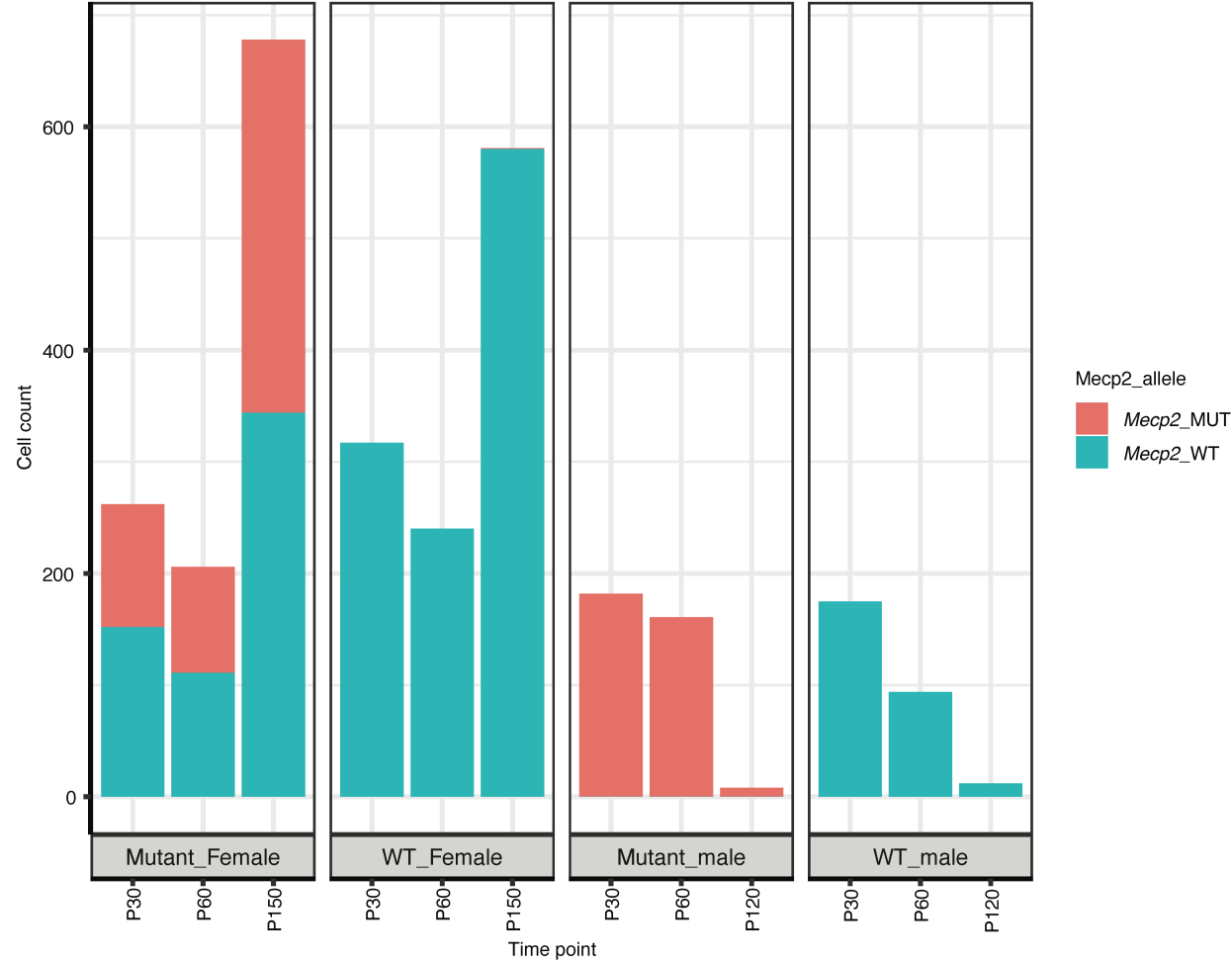
Supplemental Fig.7: Complete co-expression networks for each cell type in the mouse cortex and RTT relevant traits



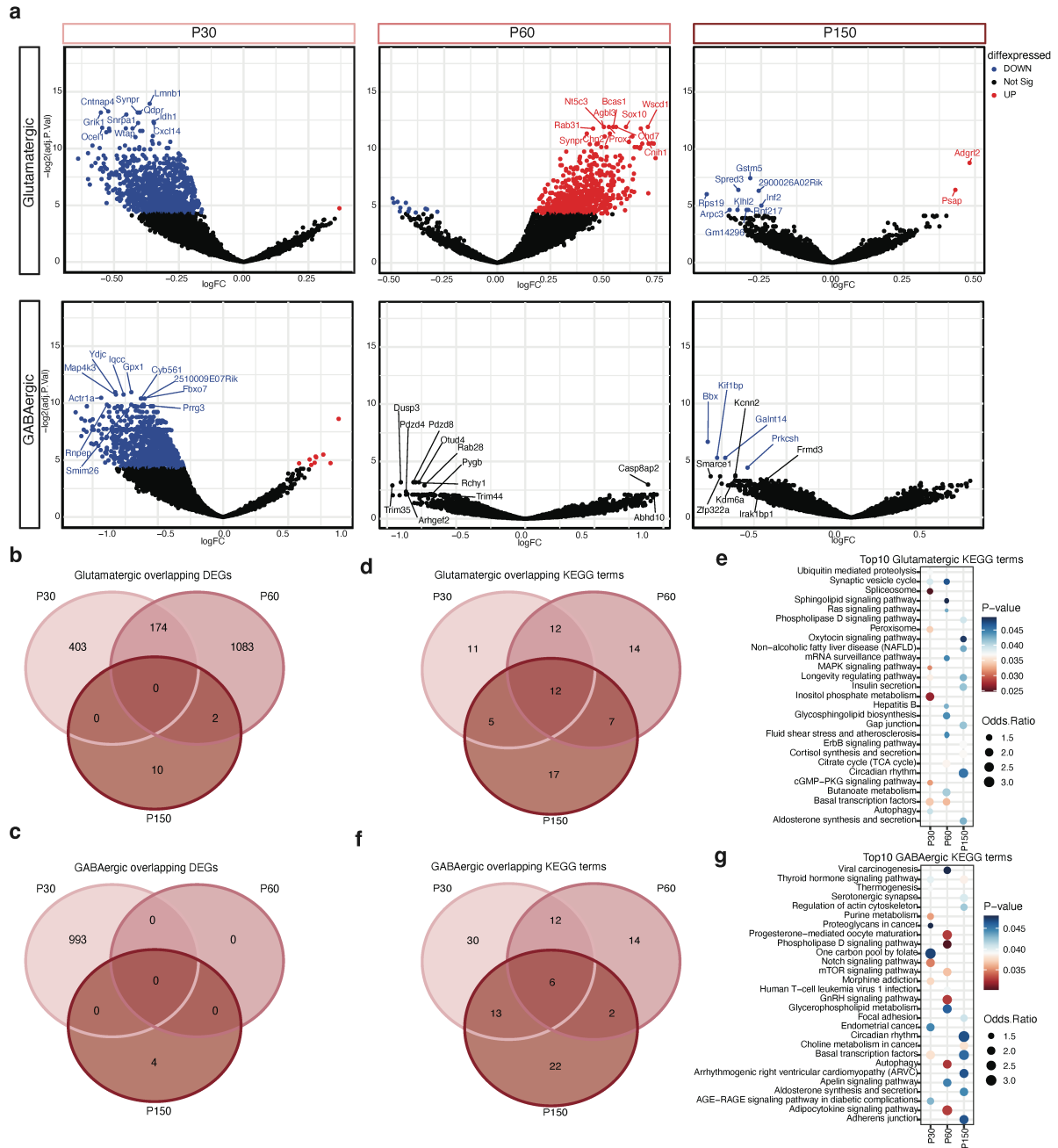
Supplemental fig.8: Top 10 hdWGCNA module KEGG terms



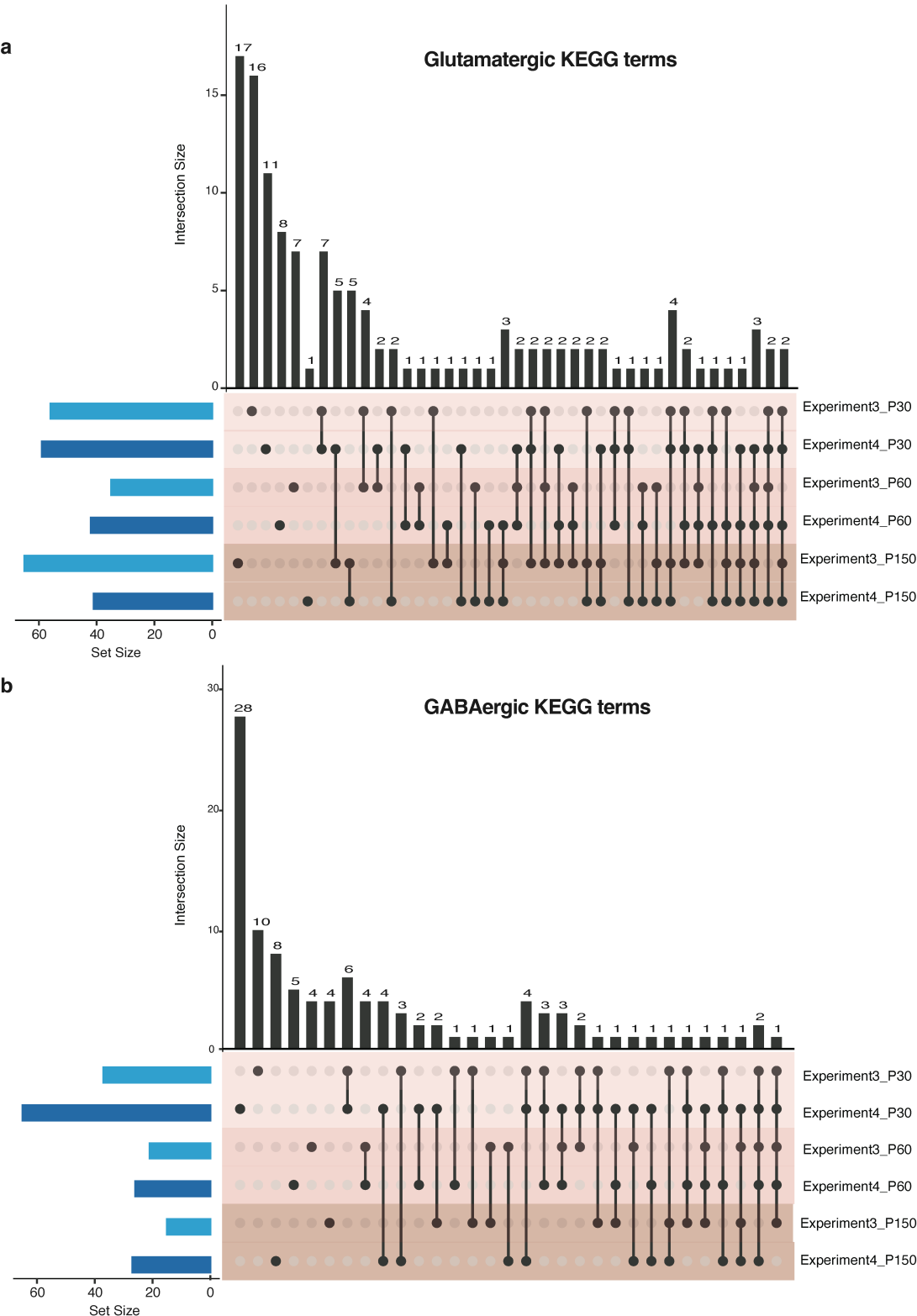
Supplemental fig.9: Bar graphs showing the ratio of WT *Mecp2* and Mutant *Mecp2* expressing cells



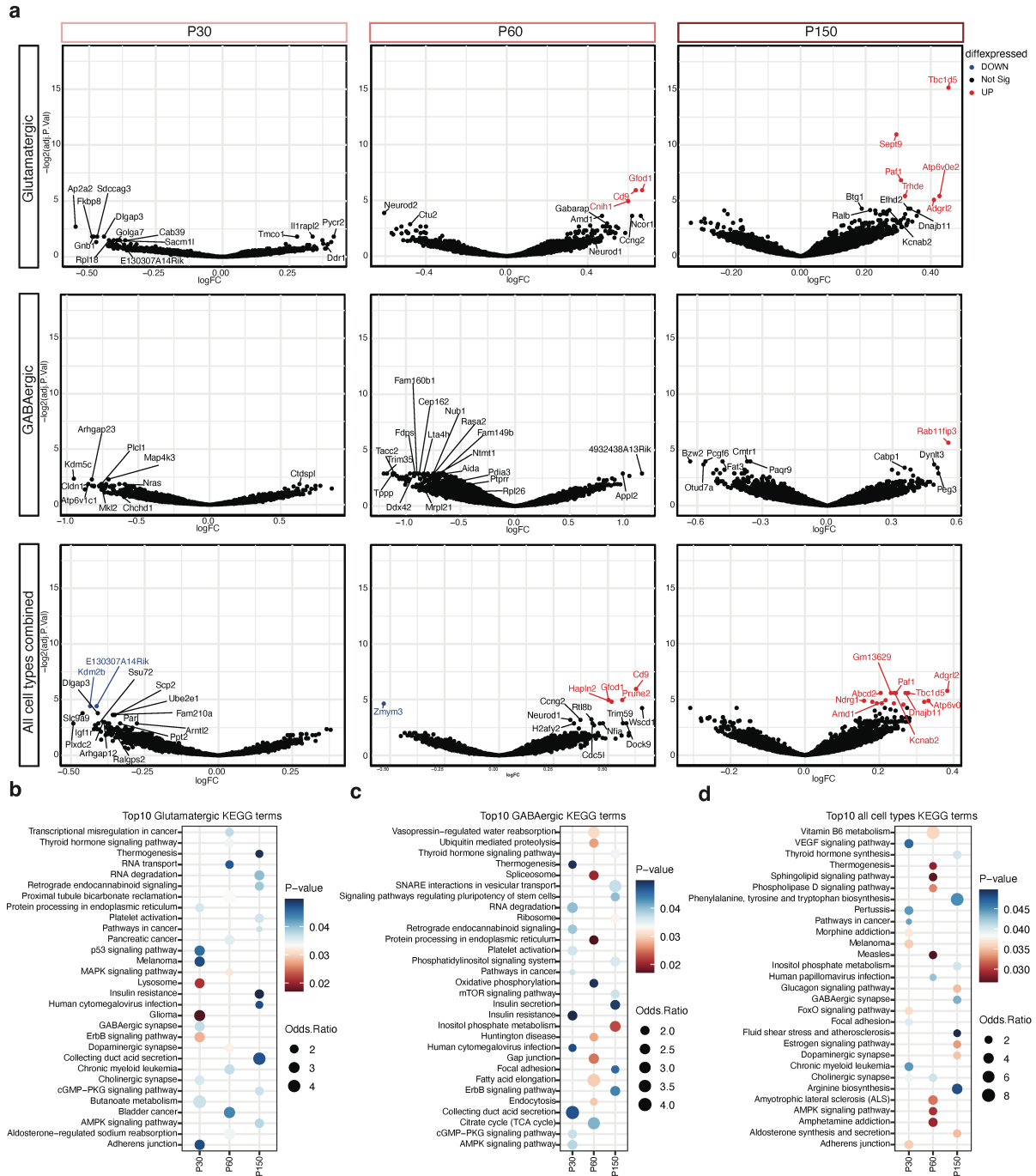
Supplemental fig.10: Differentially expressed genes comparing MUT cells from mosaic females and WT cells from WT females over time in Glutamatergic and GABAergic neurons. As shown in Experiment #4 (Fig. 1d), we compared MUT cells from *Mecp2e1*^{-/-} female with WT cells from *Mecp2e1*^{+/+} glutamatergic and GABAergic neurons longitudinally



Supplemental fig.11: Overlap of KEGG terms from experiment 3 and experiment 4 over time



Supplemental fig.12: Differentially expressed genes and KEGG pathways comparing MUT and WT cells within *Mecp2e1*^{-/-} female cortices over time in glutamatergic and GABAergic neurons



Chapter 3

Effects of PCB Mixtures on Gene Expression in Distinct Cell Types of the Rett Syndrome Mouse Model Cortex

3.1 Preface

This chapter was written by Osman Sharifi with contribution for the project from the co-authors listed below.

3.1.1 Authorship

Osman Sharifi^{1,3,4}, Kari E. Neier^{1,3,4}, Anthony Valenzuela⁵, Christina G. Torres^{1,3,4}, Ian Korf^{2,3}, Pamela J. Lein⁵, Dag H. Yasui^{1,3,4#}, Janine M. LaSalle^{1,3,4#*}

¹Medical Microbiology and Immunology, School of Medicine, ²Cellular and Molecular Biology, College of Biological Sciences, ³Genome Center, ⁴MIND Institute, ⁵Department of Molecular Biosciences, School of Veterinary Medicine, University of California, Davis, CA 95616

3.2 Abstract

Rett Syndrome (RTT) is an X-linked, dominant neurodevelopmental disorder mainly caused by mutations in the *MECP2* gene that encodes Methyl CpG binding protein (MeCP2) which functions as an epigenetic regulator. RTT babies are born seemingly normal; however, in late infancy, they experience a regression in cognitive and motor functions. The variability in severity and timing of

progression in RTT suggests that *MECP2* mutations may interact with environmental factors. The MeCP2 protein binds to methylated DNA and modulates gene expression by acting as a repressor or activator. While it is clear that MeCP2 function is very important for normal neurological function, it is unclear how molecular signatures of disease progression in distinct brain cell types interact with environmental toxicants such as persistent organic pollutants (POPs). Polychlorinated biphenyls (PCBs) are POPs that are identified as probable environmental risk factors for neurodevelopmental disorders (NDD). PCB 95 has been detected in human brains of multiple NDDs including RTT. Further, the cumulative effect of NDD genetic defects and exposure to PCB 95 has been shown to alter the methylation of more than 1,000 genes in cell lines suggesting a compounding effect of genetic and environmental factors in the dysregulation of chromatin epigenetics. The Markers of Autism Risk Learning Early Signs (MARBLES) study has measured 17 PCB congeners in mothers of increased risk for autism spectrum disorders. We have exposed WT and *Mecp2^{et1/+}* dams to this 0.1mg/kg of the MARBLES PCB mix. Here we present single-nucleus 5' RNA-seq data from the cortex of the dams. Our findings establish a baseline for single cell gene expression in the brains of PCB exposed mice and suggests that PCBs and MeCP2 work on similar molecular pathways. We observe a notable overlap in dysregulated genes and molecular pathways between the effects of PCB exposure and MeCP2 mutations, indicating a potential mutual mitigation of their transcriptional impacts. Additionally, PCBs primarily influenced the non-cell-autonomous transcriptional effects of MeCP2 mutations in wild-type-expressing neurons within the mosaic mutant female cortex, suggesting that the interaction predominantly involves homeostatic gene networks.

3.3 Introduction

Rett Syndrome (RTT) is a neurodevelopmental disorder that primarily affects females and is caused by mutations in the *MECP2* gene located on the X chromosome¹. This gene encodes Methyl

CpG binding protein 2 (MeCP2), a critical epigenetic regulator that modulates gene expression by binding to methylated DNA and acting as either a repressor or activator of transcription²⁻⁴. The prevalence of RTT in females is due to the *MECP2* gene being located on the X chromosome, and males with a single X chromosome carrying a *MECP2* mutation typically do not survive to term. In contrast, females have two X chromosomes, and the presence of one functional *MECP2* allele can somewhat mitigate the severity of the disorder, though it still leads to significant neurodevelopmental impairment. RTT infants are born seemingly healthy but experience a significant regression in cognitive and motor skills during late infancy, highlighting the importance of MeCP2 for normal neurological function⁵.

The progression and severity of RTT are highly variable, suggesting that the disease's manifestations may be influenced by additional factors beyond genetic mutations alone. This variability raises the possibility that *MECP2* mutations might interact with environmental factors, further complicating the disease's clinical presentation⁶. Environmental toxicants such as persistent organic pollutants (POPs) have emerged as potential contributors to neurodevelopmental disorders (NDDs)^{7,8}. Among these, polychlorinated biphenyls (PCBs) are notable for their widespread environmental presence and their association with adverse neurodevelopmental outcomes⁹. Laufer et al. examined the differentially methylated regions (DMR) of fetal brain and placenta of mice exposed to PCBs and showed a significant overlap between MeCP2 regulated genes and PCB exposure⁶.

PCB 95, in particular, has been detected in the brains of individuals with various NDDs, including RTT^{10,11}. Research has demonstrated that exposure to PCB 95 can alter the methylation patterns of over 1,000 genes in cell lines, suggesting a potential interaction between genetic defects and environmental exposures in the disruption of chromatin epigenetics¹². In females, X-chromosome inactivation (XCI) is random, affecting either the maternal or paternal allele in each cell. Once established, this creates a cellular mosaicism throughout the body, unless there is selective loss of one

pattern¹³. This cellular mosaicism might protect females from prenatal environmental exposures and the genetic diversity from their two X chromosomes could influence gene expression and adapt their responses to exposures to PCBs better than males. Furthermore, studies such as the Markers of Autism Risk Learning Early Signs (MARBLES) study have measured PCB congeners in mothers at increased risk for autism spectrum disorders, highlighting the relevance of PCB exposure in developmental neurotoxicology¹⁴. A study by the Lein lab has shown that exposing mice to a PCB mixture based on the MARBLES study lead to an increase in dendritic branching of cortical neurons¹⁵.

To explore the intersection of genetic and environmental factors in RTT female mosaicism, we utilized a construct-relevant mouse model of RTT and performed single-nucleus 5' RNA sequencing on cortical cell types from adult female mice exposed to the MARBLES PCB mix. We demonstrate a highly significant overlap in dysregulated genes and molecular pathways and between the effects of PCB exposure and MeCP2 mutations, apparently mitigating each other's transcriptional impacts. Interestingly, PCBs preferentially impacted the non-cell-autonomous transcription effects of *Mecp2* mutation in wild-type-expressing neurons within mosaic mutant female cortex, suggesting the interaction is predominantly with homeostatic gene networks. By establishing a baseline for single-cell gene expression in PCB-exposed mice, our study aims to elucidate the complex interplay between genetic and environmental factors in RTT and related neurodevelopmental disorders, providing insights into how PCB exposures may modulate disease progression and severity.

3.4 Results

3.4.1 Experimental design to test transcriptional dysregulation in PCB-exposed WT and *Mecp2e1* heterozygous female cortical cell types

To identify cell type, and PCB exposure specific transcriptional differences in *Mecp2e1* deficient mouse cortex, single nuclei RNA sequencing (sn-RNA seq 5') analysis was performed to include the

engineered mutation at the 5' translational start site of the *Mecp2e1* isoform¹⁶. Late disease stage *Mecp2e1* (~16 weeks of age) females were chosen and compared to sex-matched wild-type (WT) littermates (**Figure 1**). Four animals from each group were exposed to the PCB MARBLES mix while the other four were treated with vehicle control for a total of 16 animals that were exposed for approximately 6 weeks (**Figure 1**).

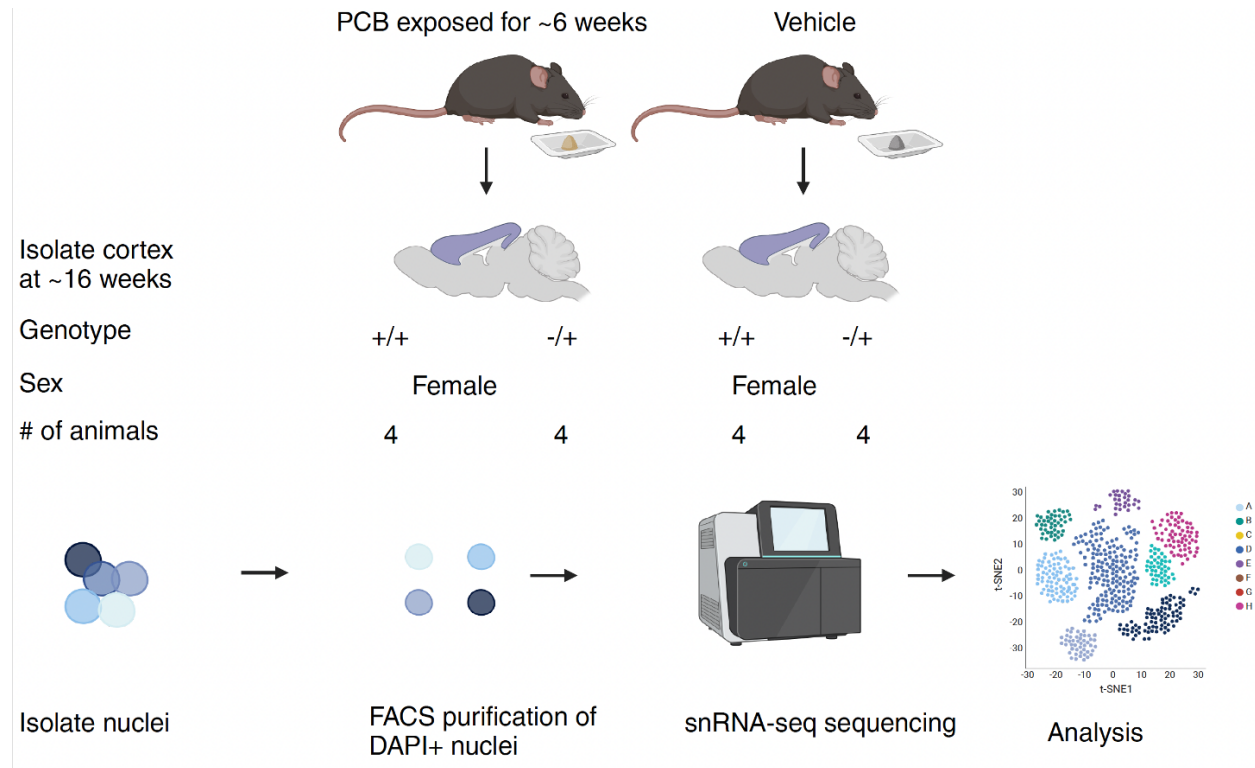


Figure 4: Study design overview showing cortical samples from PCB and vehicle control groups of two genotypes used to for snRNA-seq 5'.

3.4.2 Development of a novel bioinformatic pipeline for cell classification in a mosaic brain

To accurately distinguish between wild-type (WT) and mutant *Mecp2* expressing cells within a mosaic brain, we developed an innovative bioinformatic pipeline. This pipeline, detailed in **Figure 2**, was specifically designed to address the complexities associated with analyzing heterogeneous cell populations in the context of our gene x environment experimental design within a mosaic genetic

mouse model. **Figure 2** provides a detailed flowchart of the pipeline, illustrating each stage of the process and the interconnections between them. The figure highlights how the pipeline integrates preprocessing, feature selection, dimensionality reduction, clustering, and classification to achieve precise segregation of WT and mutant cells in the mosaic RTT brain.

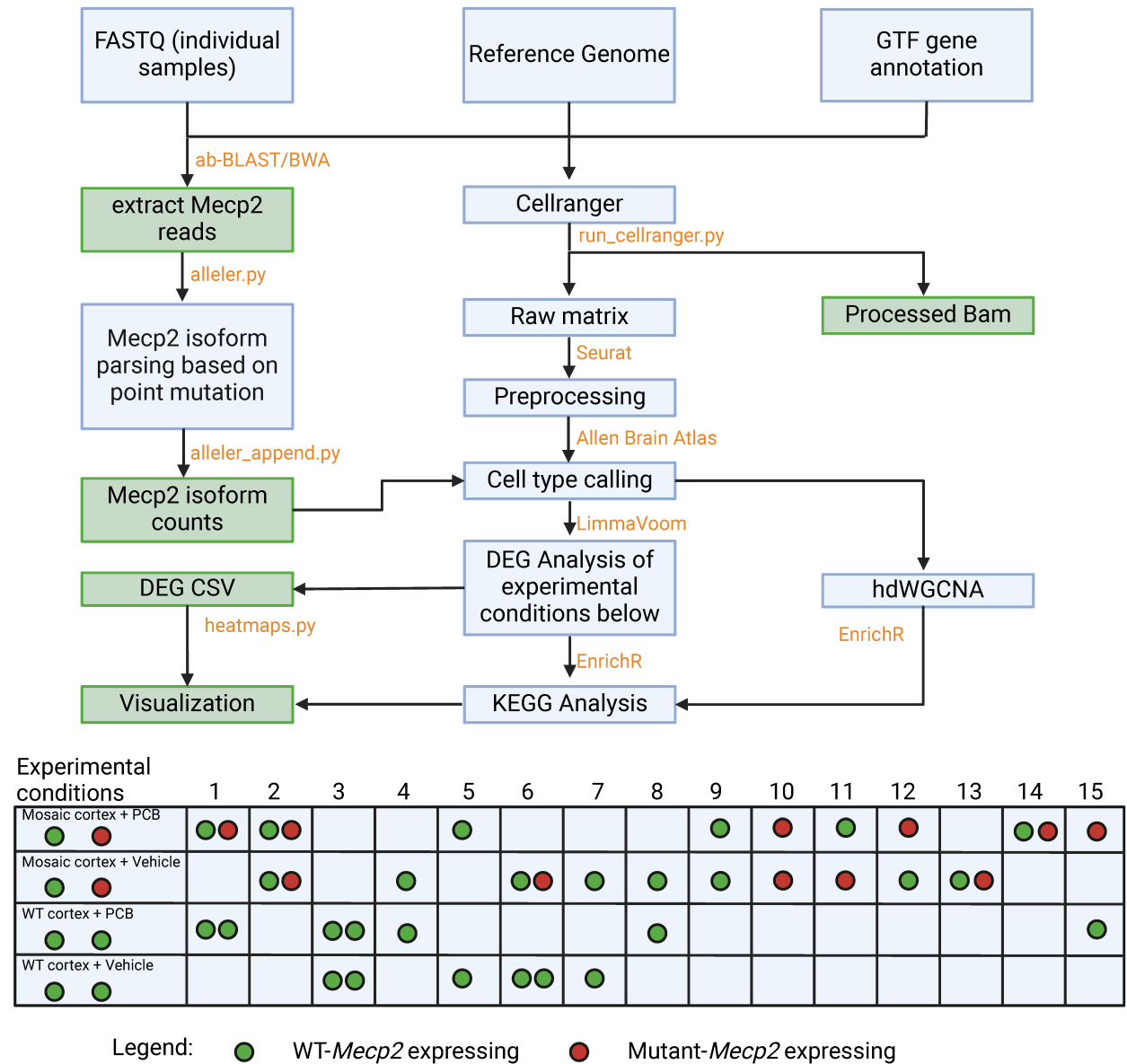


Figure 5: Bioinformatic pipeline for parsing out WT and mutant cells in the mosaic RTT cortex and differential analysis

3.4.3 Clustering and characterization of cortical cell types

To further elucidate the cellular landscape within the cortices of WT and mutant mice, we performed single-nucleus RNA sequencing (snRNA-seq 5') and subsequent clustering analysis. Our goal was to

identify distinct cell types based on a reference atlas and evaluate the impact of PCB exposure interactions with *Mecp2* genotype on these populations. We applied unsupervised clustering analysis to classify cells based on their gene expression profiles and labeled cell types based on the Allen brain atlas cortex reference¹⁷. This analysis was conducted through a combination of dimensionality reduction techniques (PCA and UMAP) and clustering algorithms. We identified 14 distinct cell types within the cortical tissue (**Figure 3a**). Cell type identification and labeling was validated based on marker genes present in each cell type (**Figure 3b**). Each marker was selected based on its known association with specific cell types, providing a robust framework for the identification and validation of the 14 cell types. These markers were instrumental in verifying the accuracy of our clustering results and ensuring that each identified cell type corresponded to its expected physiological role.

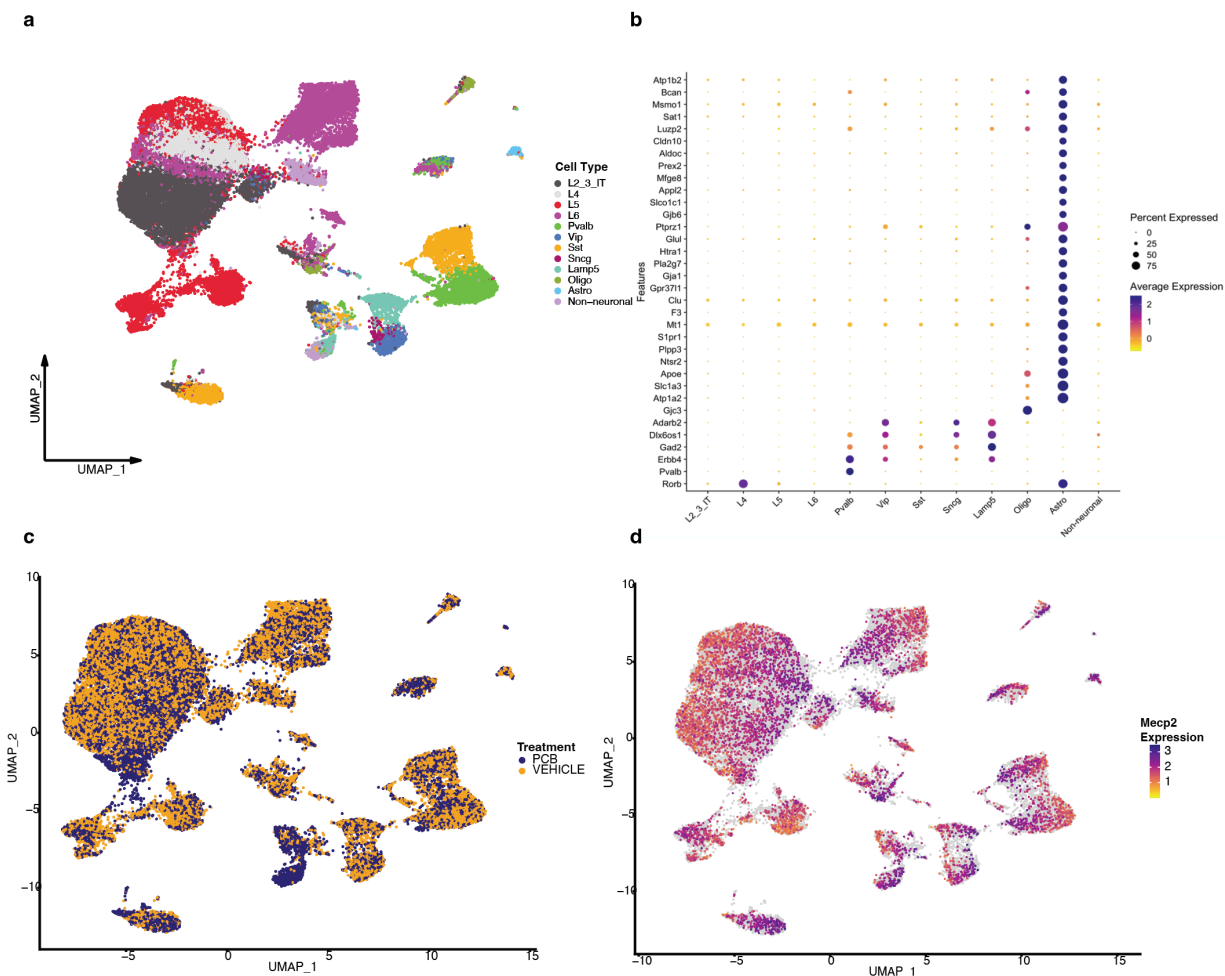


Figure 6: Cell clustering based on gene expression. **a.** UMAP of Cell Type clustering and annotation **b.** Unbiased cell type markers identified per cell type **c.** Cell clustering based on treatment **d.** *Mecp2* expression levels in each cell cluster

To assess the influence of PCB exposure on cell type clustering, we compared the clustering results of cells treated with the PCBs to those treated with a vehicle control. Cells from both treatment conditions were clustered mostly together, and the resulting cell type distributions were analyzed for differentially expressed genes (DEG) (**Figure 3c**). *Mecp2* expression was also assessed based on cell clustering showing its presence in all cell types (**Figure 3d**).

3.4.4 Cell type proportions in PCB-treated versus Vehicle-treated groups

To gain deeper insights into how PCB exposure affects the distribution of cell types in the cortex, we conducted an analysis of cell type proportions in both WT and RTT cortex. This analysis involved

comparing the proportions of different cell types between PCB-treated and vehicle-treated groups. After clustering the cells and identifying 14 distinct cell types, we quantified the proportion of each cell type within the cortical samples (**Figure 4**). The most pronounced change in cellular proportions observed are in Sst inhibitory neurons between genotypes.

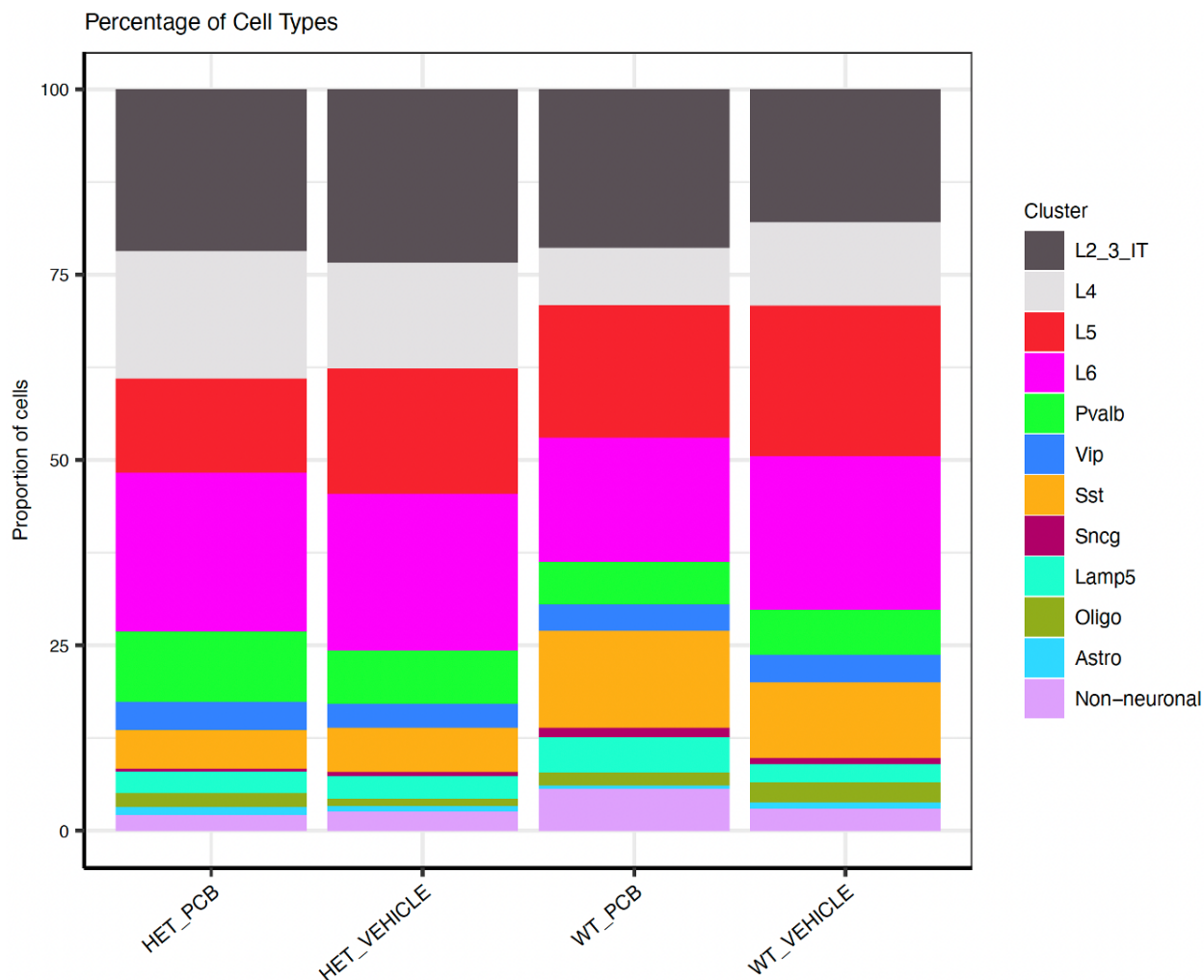


Figure 7: Bar graph showing proportion of cell types across the four experimental groups.

3.4.5 Differential gene expression (DEG) analysis of cell types in PCB-treated versus vehicle-treated groups

To further explore the impact of PCB exposure on gene expression within distinct cell types, we performed a differential expression analysis comparing vehicle-treated and PCB-treated conditions

in both WT and *Mecp2e1* mutant mice. This analysis aimed to identify changes in gene expression associated with PCB exposure and evaluate the protective or exacerbating effects of PCB treatment in the context of RTT (Figure 5).

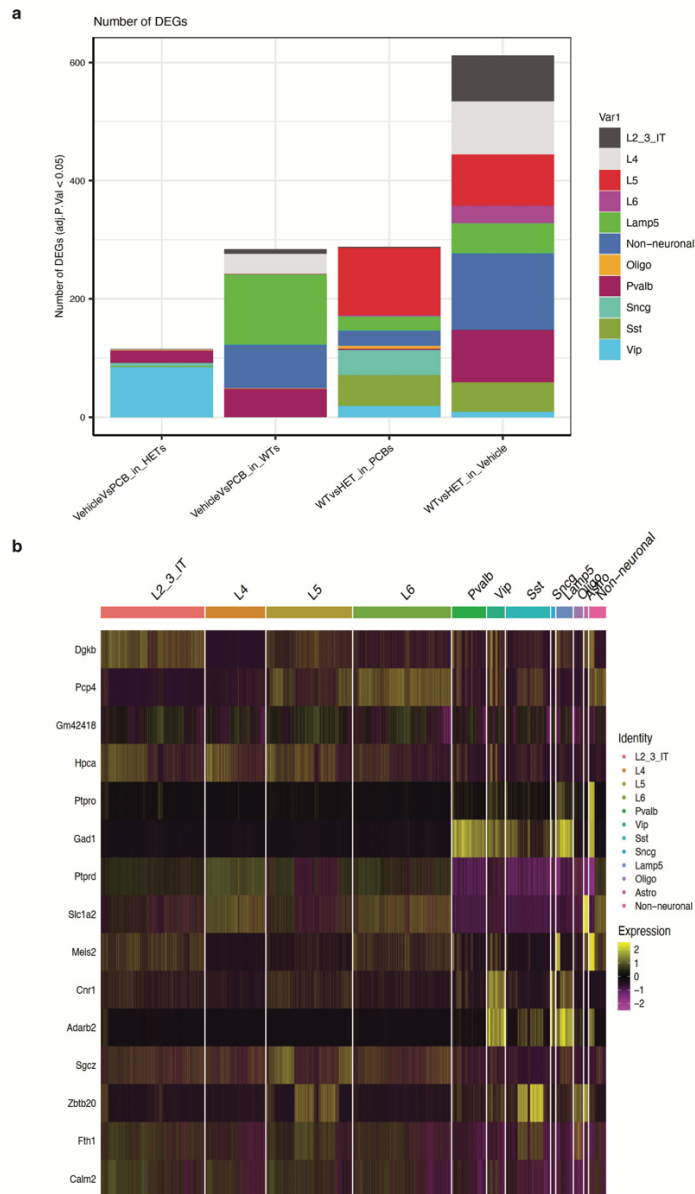


Figure 8: DEG results comparing the 4 experimental groups. **a.** Bar graph showing the significant number of DEGs for each experiment. **b.** Heatmap showing the top 15 DEGs per cell type

Additionally, we compared DEGs between WT and *Mecp2e1* mice within the vehicle-treated and PCB-treated backgrounds. This comparison aimed to discern how PCB exposure interacts with

genetic differences between WT and RTT mice. The variability in number of DEGs and expression patterns across different cell types underscores the complexity of PCB's effects, which may vary depending on the specific cellular context and the nature of the gene expression changes (**Figure 5a, b**). The highest number of DEGs were observed between WT and *Mecp2e1* cortical cells with vehicle treatment. Conversely, PCB treated RTT mice show the lowest number of DEGs compared to vehicle treated RTT cortical cells (**Figure 5a**). The lower number of DEGs in PCB-treated *Mecp2e1* mice suggests that PCB exposure may partially counteract the aberrant gene expression associated with *Mecp2* mutations. These findings highlight a potential modulatory role of PCB in the context of RTT, where PCB treatment may normalize or attenuate some of the transcriptional disruptions caused by the *Mecp2e1* mutations.

3.4.6 Dysregulated KEGG pathways across experimental conditions

To understand how PCB exposure affects cellular pathways, we analyzed dysregulated KEGG pathways across four conditions: WT vehicle, WT PCB, *Mecp2e1* mutant with vehicle treatment, and *Mecp2e1* mutant with PCB treatment. This analysis aimed to identify common and unique pathway alterations associated with the interaction between PCB exposure and the *Mecp2e1* mutation. Across all cell types, PCB exposure resulted in fewer dysregulated gene pathways in RTT cortex than vehicle control, suggesting that PCB treatment may have a protective or mitigating effect on pathway dysregulation (**Figure 6a**). A closer look at the dysregulated pathways reveals that the majority of these KEGG pathways are similar across all four conditions, with 71 overlapping pathways (**Figure 6b**). This extensive overlap indicates that both PCB exposure and RTT converge on a core set of dysregulated gene pathways. Across all experimental conditions, these overlapping dysregulated gene pathways were more impacted in neuronal than glial cells (**Figure 6c**). Further, the KEGG pathways that are consistently dysregulated across all conditions include those related to homeostatic

regulatory functions such as circadian rhythms and cAMP signaling. These neurologically relevant pathways underscore the potential interaction of PCB exposure and a neurodevelopmental disorder mutation on key aspects of neuronal function and regulation.

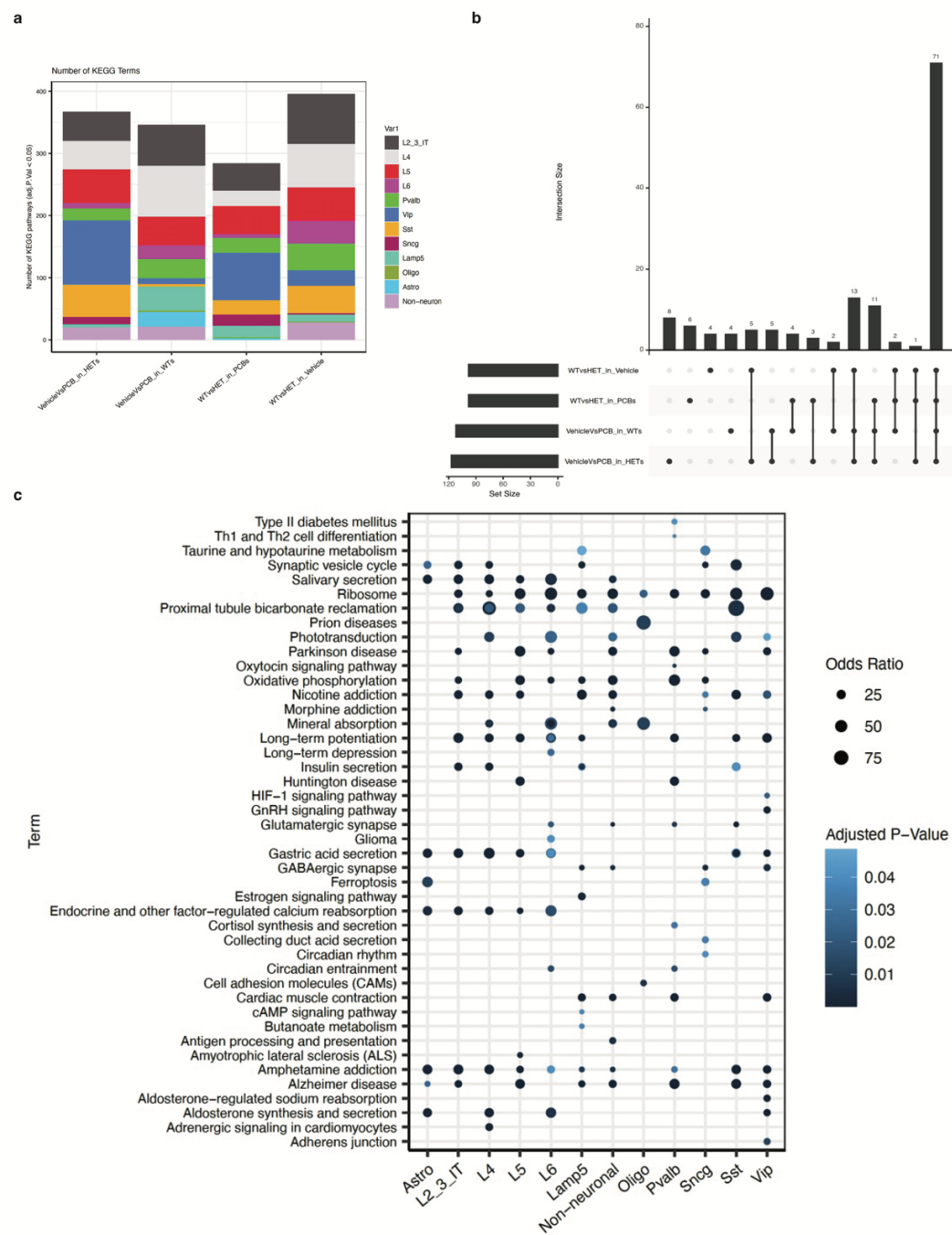


Figure 9: KEGG pathway analysis. **a.** Bar graphs showing the number of significant dysregulated KEGG pathways per cell type. **b.** Upset plot showing the number of significant

pathways overlapping across the 4 experimental conditions. **c.** Dot plot showing the top overlapping pathways per cell type.

Many of the dysregulated paths are related to neurogenerative diseases such as Parkinson's disease which is consistent with previous studies.

3.4.7 Investigating PCB effects on *Mecp2e1* mosaic cortices and cell non-autonomous interactions

To investigate how PCB exposure affects WT-*Mecp2* versus mutant-expressing cells within female mosaic cortical tissue, we first parsed *Mecp2e1* transcript genotypes in all cells (**Figure 7**). As expected for random X chromosome inactivation and what has been observed previously (Sharifi et al 2024), no mutant cells were observed in WT cortices, and female heterozygous cortical cells were a roughly equal mixture of WT- and mutant-expressing.

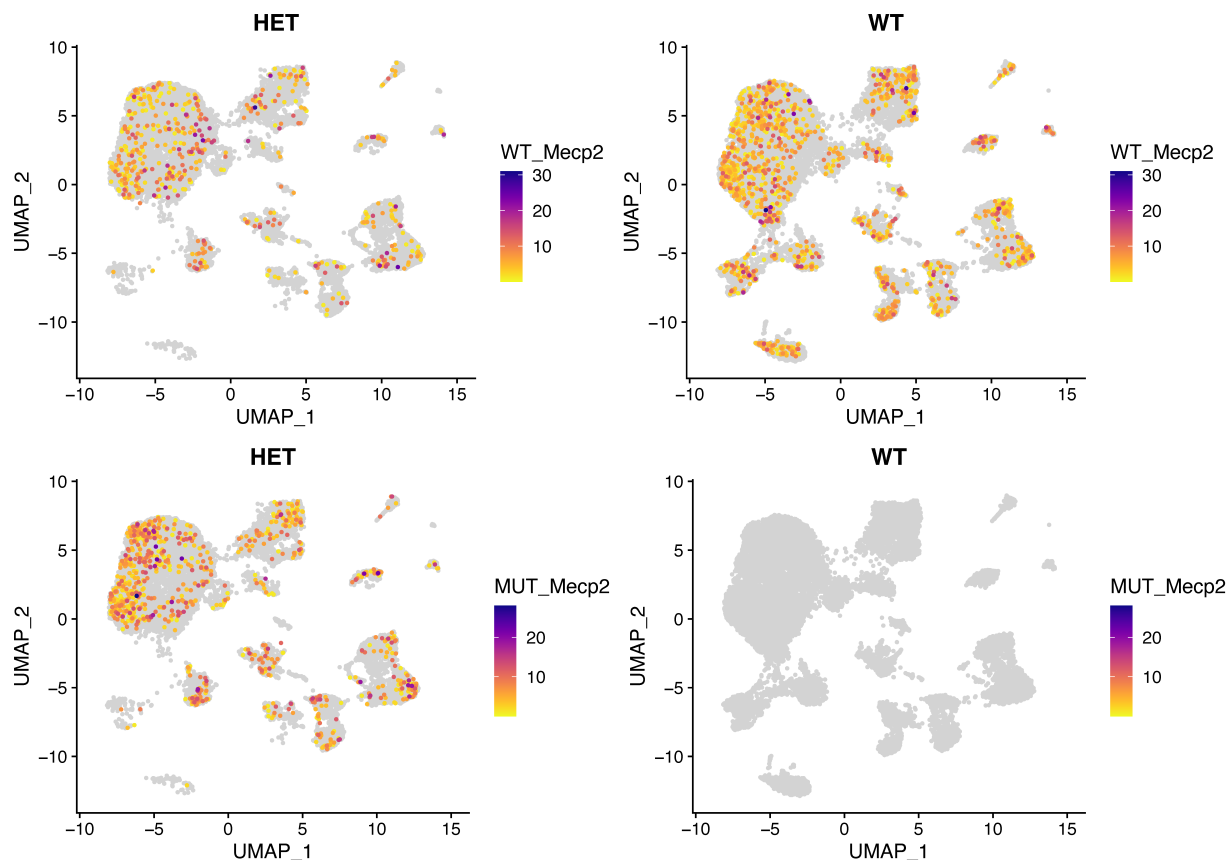


Figure 10: UMAP plots showing WT_Mecp2 and MUT_Mecp2 expression in WT cortices and *Mecp2e1* heterozygous cortices

We also specifically assessed the effects of PCB treatment on distinct cell populations within Glutamatergic and GABAergic neurons to gain insight into cell non-autonomous effects. To evaluate the effects of PCB exposure on these cell populations, we conducted 15 distinct differential expression (DEG) experiments (**Figure 8**). We quantified the number of significant DEGs in glutamatergic and GABAergic neurons for each experiment, highlighting the impact of PCB exposure on each *Mecp2e1* genotype in each of these neuronal subtypes (**Figure 8**).

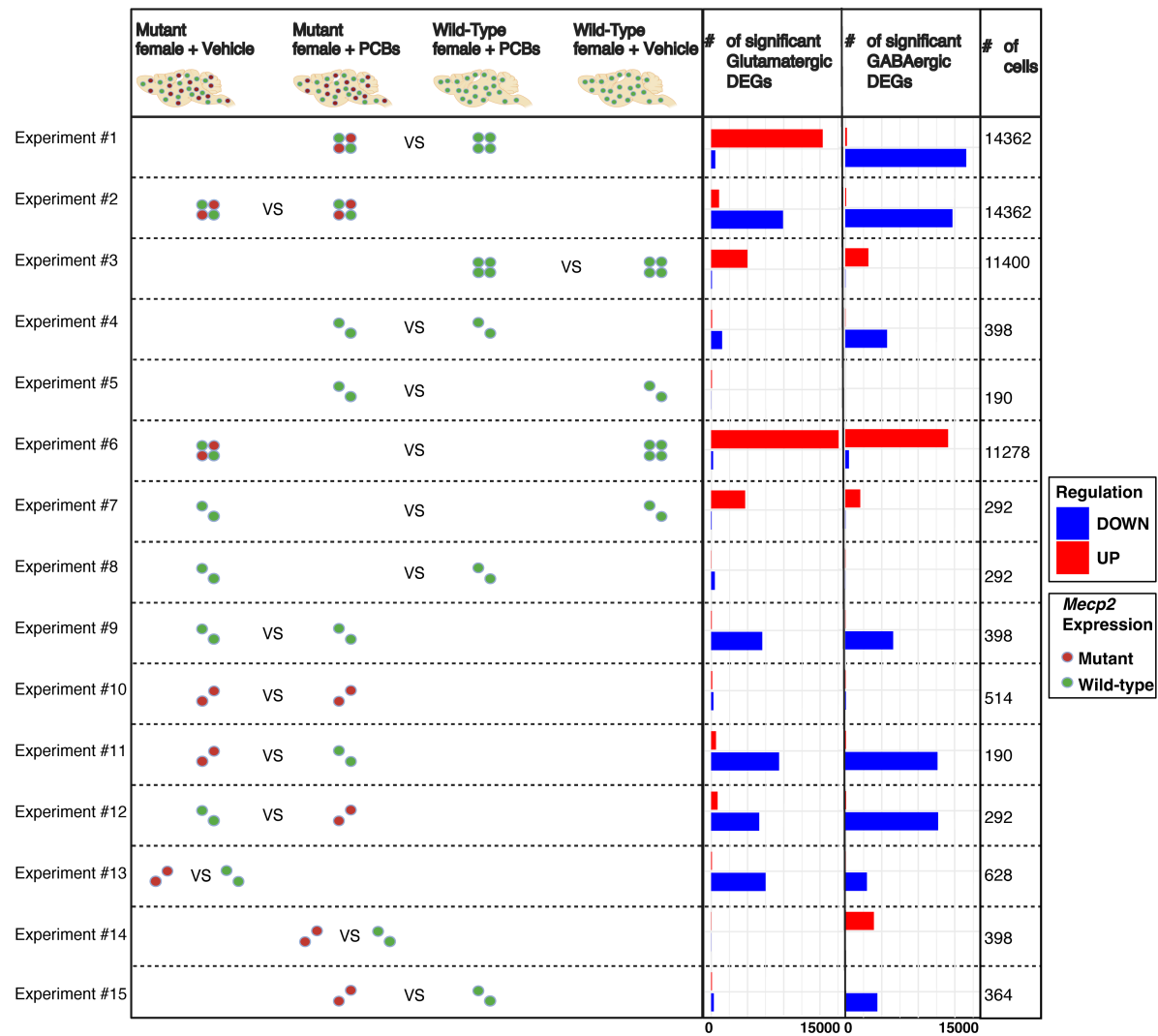


Figure 11: Distinct DEG analysis investigating cell non-autonomous effects in mosaic *Mecp2e1* cortices. In each experiment, up regulation refers to either the mutant expressing cells have higher expression, mutant mice have higher expression or PCB samples have higher expression.

Experiments 9 and 10 show that PCB treatment resulted in a higher number of dysregulated genes in WT *Mecp2*-expressing cells compared to mutant *Mecp2*-expressing cells. Remarkably, the lowest number of DEGs were seen when comparing WT-expressing neurons across both genotype and treatment groups (experiment 8). This indicates that PCB exposure predominantly affects *Mecp2* WT-expressing neurons, suggesting that PCB interaction effects within RTT cortex are indirectly acting on the homeostatic gene networks dysregulated in WT-expressing neurons. Comparisons of

Experiments 1 and 6 show that the number of DEGs with PCB treatment are different between excitatory and inhibitory neurons, indicating that PCB exposure may counteract some of the dysregulated genes associated with the *Mecp2* mutation specifically in excitatory neurons. This suggests that PCB treatment could mitigate some of the detrimental non-cell-autonomous effects caused by *Mecp2* mutation. The differences observed in DEG counts between WT and mutant cells, and the impact of PCB treatment on these cells, highlight the complex interactions within the mosaic brain environment. PCB exposure appears to interact with *Mecp2* mutation by transcriptionally dysregulating shared homeostatic pathways regulating the broader cellular milieu, demonstrating a nuanced interaction between a genetic and environmental factor.

3.4.8 Network analysis using hdWGCNA to investigate PCB and *Mecp2* genotype effects on RTT mouse phenotypes

To further explore the effects of *Mecp2* mutations and PCB exposure on RTT phenotype, we employed high dimensional Weighted Gene Co-Expression Network Analysis (hdWGCNA). This approach is relevant for assessing gene networks within snRNA-seq data and allowed us to analyze associations of experimental variables (treatment, genotype, exposure duration, and pregnancy) and measured phenotype (body weight at time of sacrifice at ~16 weeks) with gene co-expression networks. hdWGCNA analysis results show that the gene network is organized into seven distinct modules (**Figure 9a**). Each module represents a cluster of co-expressed genes with common functional characteristics. Hub genes within each module are highlighted, serving as central connectors that are crucial for the module's function and overall network stability (**Figure 9a**). The genes from these distinct modules were used to perform KEGG analysis. The top KEGG pathways for each module reflect the biological processes and pathways most prominently affected by PCB exposure and/or *Mecp2* genotype (**Figure 9b**). These pathways include many neurological disease related terms such as Parkinson's disease, Alzheimer's disease, calcium signaling and long-term

potentiation which provide insights into the functional impact of these factors on gene networks
(**Figure 9b**).

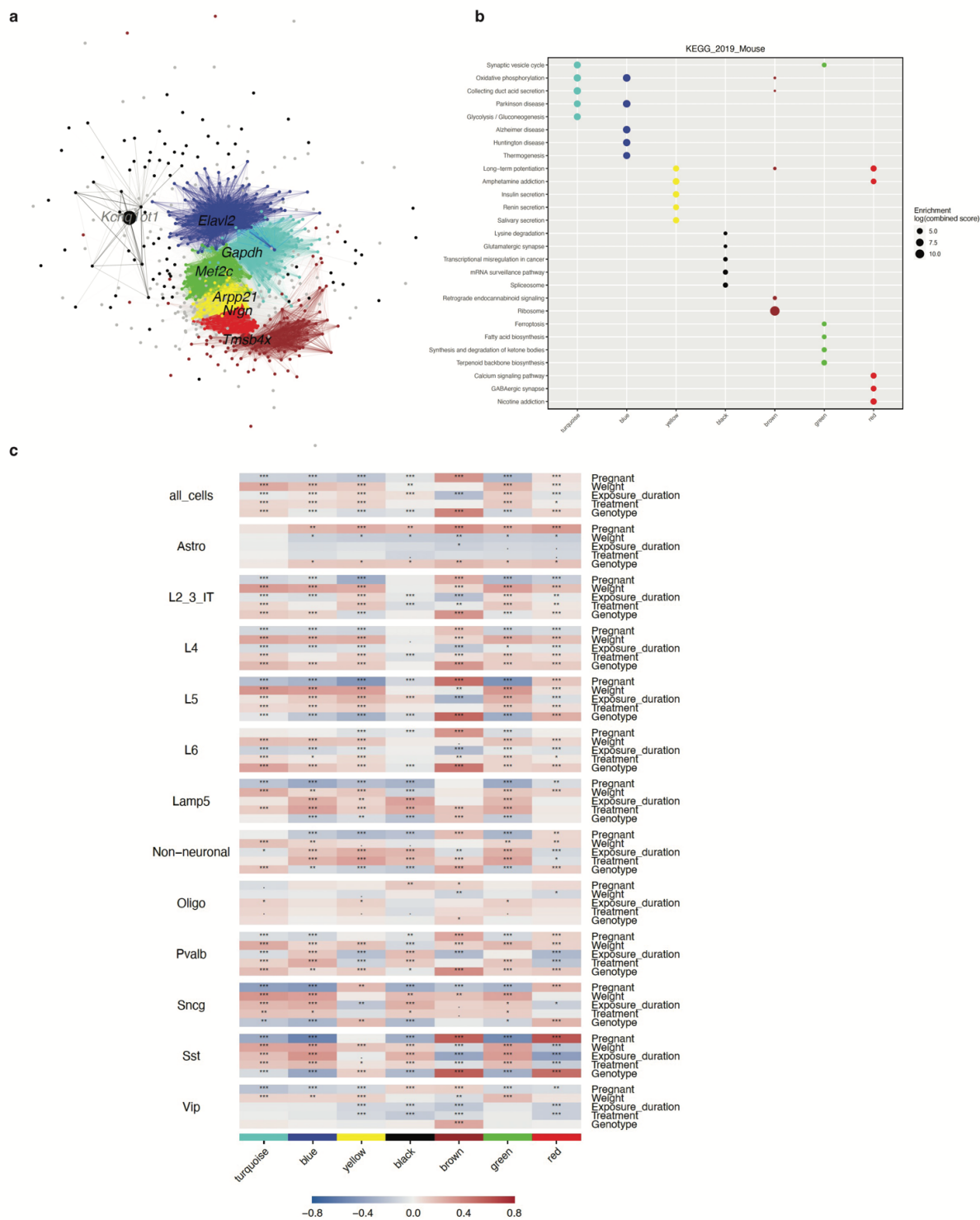


Figure 12: hdWGCNA analysis. a. Weighted gene network showing 7 modules and their corresponding hub genes. b. Dot plot showing the top 5 KEGG pathways for each module c. Heatmap showing the correlations between gene expression and phenotypes. If an

adjusted-p-value is less than 0.05, it is flagged with one star (*). If an adjusted-p-value is less than 0.01, it is flagged with 2 stars (**). If an adjusted p-value is less than 0.001, it is flagged with three stars (***).

Further, we investigated the correlation of gene expression in each cell type with various experimental variables, including pregnancy, weight at sacrifice, exposure duration, PCB treatment, and genotype (**Figure 9c**). In the context of genotype, a positive correlation indicates heterozygous mice, while a positive correlation with treatment implies an association with PCBs. Conversely, a positive correlation with pregnancy means the mice are not pregnant, whereas a negative correlation indicates that the mice are pregnant. The brown module exhibits significant correlations with several phenotypic features, particularly with genotype and PCB treatment. This module appears to be heavily influenced by both the *Mecp2* mutation and PCB exposure. Among the different cell types analyzed, the Sst cells show the strongest correlations with changes in gene expression linked to phenotypic traits (**Figure 9c**). This indicates that the Sst cell type may play a critical role in the phenotypic outcomes associated with *Mecp2* mutations and PCB exposure. Interestingly, there were cell type-specific differences in the direction of PCB-induced change in the red and brown gene networks, or *Mecp2* mutant-related change in the blue and yellow networks. These results suggest that *Mecp2* mutation and PCB exposures interact within shared excitatory/inhibitory balance and homeostatic transcriptional networks in the female adult cortex.

3.5 Discussion

This study provides a comprehensive analysis of how PCB exposure influences cortical neuronal gene expression in *Mecp2^{fl}* mosaic cortices, shedding light on the complex interplay between an environmental and a genetic factor regulating neurodevelopment. Using a combination of single-cell RNA sequencing, differential expression analysis, and network analysis, we explored the effects of PCBs on RTT pathology and identified potential mechanisms underlying these interactions. We were

able to directly differentiate WT and mutant *Mecp2*-expressing cells within the mosaic brains, facilitating a detailed comparative analysis of gene expression and pathway dysregulation between WT- and mutant-expressing neurons within and across experimental genotype and treatment groups. These experimental comparisons provided a high degree of specificity in parsing WT from mutant cells, which was crucial for understanding how PCB exposure impacts cell non-autonomous effects. Cell clustering revealed 14 distinct cell types in the cortices and demonstrated that PCB exposure affects these cell types differently. The reduction in the number of dysregulated pathways observed with PCB treatment indicates a mitigating effect of PCB exposure on RTT-related gene dysregulation. This effect was notably pronounced in the context of WT *Mecp2*-expressing cells.

Our pathway dysregulation analysis showed that PCB exposure generally led to fewer dysregulated pathways compared to vehicle treatment specifically in the RTT mouse background. This finding indicates a possible protective role of PCB exposure. The KEGG pathways commonly affected across all conditions included those relevant to adaptive pathways relevant to neurological diseases, such as circadian rhythms and cAMP signaling, highlighting the broad impact of PCB exposure on critical biological processes in RTT.

The investigation into *Mecp2^{fl}* mosaic cortices allowed us to study cell non-autonomous effects, revealing that PCB exposure transcriptionally affects WT *Mecp2*-expressing cells more significantly than mutant cells. This differential impact suggests that PCBs might modulate the effects of *Mecp2* mutations in a cell non-autonomous manner by working on similar gene pathways; however, we do not know the effects of the dysregulated genes on phenotype. The results from our differential expression experiments indicated that PCB treatment could mitigate some of the gene dysregulation associated with *Mecp2* mutations.

Weighted Gene Co-Expression Network Analysis (hdWGCNA) further elucidated the effects of PCB exposure and *Mecp2* mutations on gene networks. We identified seven distinct gene modules,

with the brown module showing significant correlations with genotype and PCB treatment. Specific pathways and cell types, particularly Somatostatin (Sst) cells, were found to be strongly influenced by PCB exposure and genetic mutations. This analysis underscored the importance of particular gene modules and pathways in mediating RTT phenotypes and the impact of environmental factors.

While the study provides valuable insights, several limitations should be considered. Our analysis focused specifically on PCB exposure, but other environmental toxins and factors may also influence RTT. Future research should explore a broader range of environmental exposures to gain a more comprehensive understanding. Longitudinal studies are needed to assess how gene expression changes evolve over time and their long-term implications for RTT. Additionally, further functional validation of the identified pathways and gene modules is necessary to confirm their roles in RTT and the impact of PCB exposure.

In conclusion, this study advances our understanding of how PCB exposure interacts with *Mecp2* mutations to influence RTT pathology. By elucidating the effects on gene expression, pathway dysregulation, and network dynamics, we could focus on therapeutic targets and mechanisms that could inform future research and clinical strategies for managing RTT.

3.6 References

1. Zoghbi, H. Y. MeCP2 Dysfunction in Humans and Mice. *J. Child Neurol.* (2005) doi:10.1177/08830738050200090701.
2. Chahrour, M. *et al.* MeCP2, a key contributor to neurological disease, activates and represses transcription. *Science* (2008) doi:10.1126/science.1153252.
3. Jordan, C., Li, H. H., Kwan, H. C. & Francke, U. Cerebellar gene expression profiles of mouse models for Rett syndrome reveal novel MeCP2 targets. *BMC Med. Genet.* **8**, 36 (2007).

4. Tudor, M., Akbarian, S., Chen, R. Z. & Jaenisch, R. Transcriptional profiling of a mouse model for Rett syndrome reveals subtle transcriptional changes in the brain. *Proc. Natl. Acad. Sci. U. S. A.* **99**, (2002).
5. Neul, J. L. *et al.* Specific mutations in methyl-CpG-binding protein 2 confer different severity in Rett syndrome. *Neurology* **70**, 1313–1321 (2008).
6. Laufer, B. I. *et al.* Placenta and fetal brain share a neurodevelopmental disorder DNA methylation profile in a mouse model of prenatal PCB exposure. *Cell Rep.* **38**, 110442 (2022).
7. Park, H.-Y. *et al.* Neurodevelopmental toxicity of prenatal polychlorinated biphenyls (PCBs) by chemical structure and activity: a birth cohort study. *Environ. Health* **9**, 51 (2010).
8. Berghuis, S. A., Bos, A. F., Sauer, P. J. J. & Roze, E. Developmental neurotoxicity of persistent organic pollutants: an update on childhood outcome. *Arch. Toxicol.* (2015) doi:10.1007/s00204-015-1463-3.
9. Jacobson, J. L. & Jacobson, S. W. Prenatal exposure to polychlorinated biphenyls and attention at school age. *J. Pediatr.* **143**, 780–788 (2003).
10. Lesiak, A. *et al.* The environmental neurotoxicant PCB 95 promotes synaptogenesis via ryanodine receptor-dependent miR132 upregulation. *J. Neurosci.* (2014) doi:10.1523/JNEUROSCI.2884-13.2014.
11. Wayman, G. A. *et al.* PCB-95 Promotes Dendritic Growth via Ryanodine Receptor–Dependent Mechanisms. *Environ. Health Perspect.* **120**, 997–1002 (2012).
12. Dunaway, K. W. *et al.* Cumulative Impact of Polychlorinated Biphenyl and Large Chromosomal Duplications on DNA Methylation, Chromatin, and Expression of Autism Candidate Genes. *Cell Rep.* (2016) doi:10.1016/j.celrep.2016.11.058.

13. Amir, R. E. *et al.* Influence of mutation type and X chromosome inactivation on Rett syndrome phenotypes. *Ann. Neurol.* (2000) doi:10.1002/1531-8249(200005)47:5<670::AID-ANA20>3.0.CO;2-F.
14. Granillo, L. *et al.* Polychlorinated biphenyls influence on autism spectrum disorder risk in the MARBLES cohort. *Environ. Res.* **171**, 177–184 (2019).
15. Keil Stietz, K. P. *et al.* Sex and Genotype Modulate the Dendritic Effects of Developmental Exposure to a Human-Relevant Polychlorinated Biphenyls Mixture in the Juvenile Mouse. *Front. Neurosci.* **15**, 766802 (2021).
16. Yasui, D. H. *et al.* Mice with an isoform-ablating Mecp2exon 1 mutation recapitulate the neurologic deficits of Rett syndrome. *Hum. Mol. Genet.* **23**, 2447–2458 (2014).
17. tasic, B. *et al.* Shared and distinct transcriptomic cell types across neocortical areas. *Nature* (2018) doi:10.1038/s41586-018-0654-5.

Chapter 4

Single cell-level transcriptomic signatures of Rett syndrome disease progression in the hypothalamus

4.1 Preface

This chapter was written by Osman Sharifi with contribution for the project from the co-authors listed below.

4.1.1 Authorship

Osman Sharifi^{1,3,4}, Kari E. Neier^{1,3,4}, Keith J. Fraga^{2,3}, Ian Korf^{2,3}, Sophia M. Hakam^{1,3,4}, Gerald Quon^{2,3}, Dag H. Yasui^{1,3,4#}, Janine M. LaSalle^{1,3,4#}

¹Medical Microbiology and Immunology, School of Medicine, ²Cellular and Molecular Biology, College of Biological Sciences, ³Genome Center, ⁴MIND Institute, University of California, Davis, CA 95616

#equal contribution

ORCID OS 0000-0002-4630-2587, KEN 0000-0002-9700-4045, KF 0000-0002-8930-6510, IFK 0000-0002-3480-2031, SH 0009-0006-1465-6571 JQ 0000-0002-1716-0153, DHY0000-0003-1178-988X, JML 0000-0002-3480-2031

4.2 Abstract

Rett syndrome (RTT) is a severe neurodevelopmental disorder primarily affecting females, caused by mutations in the *MECP2* gene. While the role of MeCP2 in neurons is well-documented, its impact on other cell types, and variations across sexes and developmental stages, remain less understood

especially in the hypothalamus. This is especially important considering the bodyweight changes in females suggesting metabolic dysfunction. We utilized single nuclei RNA sequencing (snRNA-seq 5') to examine transcriptional dysregulation in the hypothalamus of a *Mecp2^{fl}* mutant mouse model, aiming to address these gaps. We identified ten distinct cell types within the hypothalamus and confirmed high *Mecp2* expression in inhibitory and excitatory neurons. Differential gene expression analysis revealed a peak of transcriptional changes in male excitatory neurons at postnatal day 30 (P30) and in female excitatory neurons at postnatal day 60 (P60). Key dysregulated genes were involved in neurotransmission, synaptic plasticity, and metabolic pathways. Gene Ontology (GO) analysis highlighted significant disruptions in synaptic function, cellular metabolism, and neural development. MeCP2 deficiency induces distinct cellular and molecular alterations in the hypothalamus compared to the cortex, with sex-specific and developmental stage-related variations. These findings advance our understanding of RTT's metabolic and neurodevelopmental disruptions and underscore the importance of considering both cellular and sex-specific factors in RTT research. Our results offer novel insights into cell types regulating metabolic changes during RTT disease progression.

4.3 Introduction

Rett syndrome (RTT) is a debilitating neurodevelopmental disorder primarily affecting females, characterized by a spectrum of severe symptoms including loss of speech, motor abnormalities, and developmental regression occurring between 6 and 18 months of age^{1,2}. This disease predominantly results from spontaneous missense or truncation mutations in the X-linked *MECP2* gene, which encodes methyl CpG binding protein 2 (MeCP2)³. MeCP2 is an abundant neuronal chromatin factor and transcriptional regulator which is essential for normal neuronal development and brain function⁴. In RTT, *MECP2* mutation leads to transcriptional dysregulation, leading to progressive neurodevelopmental deficits observed in affected individuals⁵.

The hypothalamus, a pivotal brain region involved in regulating various metabolic processes, plays a crucial role in maintaining homeostasis, including the regulation of appetite, energy expenditure, and endocrine functions⁶. Given RTT's classification as a metabolic disease and the hypothalamus's central role in metabolic regulation, understanding the impact of MeCP2 deficiency within this region could provide significant insights into the disease's progression^{7,8}.

Studies of RTT typically utilize mouse models to investigate the underlying mechanisms of the disorder. The exon 3-4 knockout model in *Mecp2*^{+/y} males, though instrumental in elucidating MeCP2 function, does not fully capture the complexity of RTT in females, who are heterozygous for *MECP2* mutations and exhibit cellular mosaicism⁹⁻¹². In these heterozygous females, the presence of both wild-type and mutant *MECP2*-expressing cells may contribute to cell non-autonomous effects, influencing the progression and severity of the disease¹³⁻¹⁵.

Recent research suggests that the transcriptional dysregulation in RTT females extend to the wild-type expressing neurons without mutant MeCP2¹⁶. However, the extent and nature of these cell non-autonomous effects, especially in relation to metabolic regulation by the hypothalamus, remain poorly understood^{17,18}. Furthermore, while animal models such as the *Mecp2e1* deficient mouse model effectively mimic some aspects of RTT progression, the precise timing and cell-specific nature of the transcriptional changes associated with the disease progression are still poorly understood^{3,4,19}.

To address these gaps, we utilized single nuclei RNA sequencing (sn-RNA-seq 5') to analyze the hypothalamus of the *Mecp2e1* mutant mouse model in addition to wild-type (WT) controls, focusing on how sex, cell type, dysregulated genes, and cellular mosaicism influence disease progression over three time points. This approach allowed us to assess the transcriptional landscape at a high resolution and correlate these changes with progressive disease phenotypes. Our analysis reveals that MeCP2 deficiency leads to distinct cellular and molecular alterations in females compared to males, highlighting cell non-autonomous transcriptional changes in metabolic and homeostatic

pathways regulated by the hypothalamus. These findings offer new insights into how RTT impacts metabolic regulation and contribute to a deeper understanding of the disease's progression and manifestation in both sexes.

4.4 Results

4.4.1 Experimental design to test longitudinal, cellular, and sex-specific transcriptional dysregulation in hypothalamus of a symptomatically progressive mouse model of Rett syndrome

To investigate the longitudinal, cellular, and sex-specific transcriptional dysregulation in the hypothalamus associated with Rett syndrome, we designed a comprehensive experimental approach utilizing a symptomatically progressive mouse model of RTT. This model, carrying the patient relevant *Mecp2e1* point mutation, recapitulates key aspects of human RTT pathology and provides a valuable framework for studying disease progression at a granular level. **Figure 1a** illustrates the specific *Mecp2* construct utilized in this study, highlighting the ATG to TTG point mutation that introduces the *Mecp2e1* variant⁷. This mutation is crucial as it is construct-relevant in RTT research, offering insights into the disease mechanisms that are pertinent to the human condition. **Figure 1b** summarizes the study design, detailing the parameters of Age, Disease State, Brain Region (Hypothalamus), Sex, and Genotype. Single nucleus RNA sequencing (snRNA-seq 5') methodology was employed performed on duplicate hypothalamus samples in this study (**Figure 1c**). This technique enables high-resolution transcriptomic analysis by isolating individual nuclei from hypothalamic tissue samples. The snRNA-seq approach was utilized to assess cellular and molecular alterations in gene expression, allowing for a detailed examination of how MeCP2 deficiency affects different cell types and stages of disease progression. The 5' end sequencing method specifically enriches for transcript information from the

5' end of RNA, providing insights into the genotype of the *Mecp2* allele being expressed, a key component to investigating cellular mosaicism.

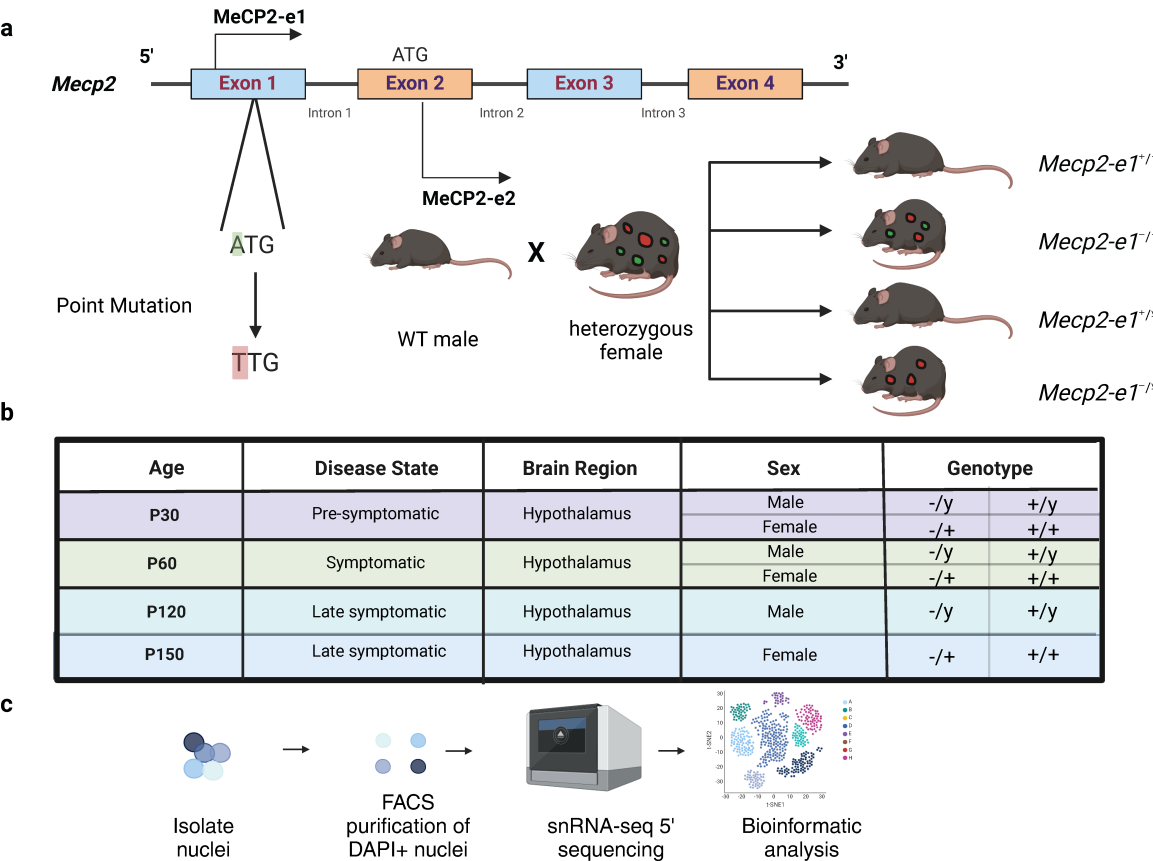


Figure 13: Study overview. **a.** A schematic of the mouse *Mecp2* construct and mouse breeding. Red cells represent mutant-expressing and green cells represent wild-type expressing cells in heterozygous females. **b.** longitudinal study design **c.** single nucleus RNA-seq 5' overview.

4.4.2 Cell type labeling and identification

To comprehensively analyze the transcriptional dysregulation in the hypothalamus of our RTT mouse model, we performed cell type labeling using established cell type markers. We leveraged the data from Steuernagel et al.²⁰, which provides a robust reference for identifying distinct neuronal and non-neuronal cell types within the hypothalamus.

We identified 10 unique cell types in the hypothalamus (**Figure 2a**). This classification was achieved through unsupervised clustering analysis of single nuclei RNA-seq data, which enabled us to distinguish between various cell populations based on their transcriptional profiles. The distinct clusters correspond to specific cell types, reflecting the diverse cellular composition of the hypothalamus and facilitating a targeted analysis of cell type-specific changes in RTT. By applying 3000 markers to our snRNA-seq dataset, we were able to map the cell type-specific gene expression patterns in the hypothalamus of RTT mice. This labeling approach was validated by looking at the top markers for each cell type (**Figure 2b**).

4.4.3 Cell clustering based on genotype, sex, and disease time point

To investigate the influence of genotype, sex, and disease progression on hypothalamic cell types, we performed clustering on these variables. **Figure 3a** demonstrates that cell clustering is not significantly affected by genotype; both wild-type (WT) and *Mecp2e1* mutant mice exhibit similar clustering patterns. This indicates that the overall distribution of cell types within the hypothalamus remains consistent regardless of the genotype, suggesting that genotype-specific transcriptional changes do not drastically alter cell type composition.

Cells were also clustered by sex, revealing that both male and female mice are present across most cell clusters (**Figure 3b**). This suggests that sex-specific differences in hypothalamic cell types or their transcriptional profiles are relatively minor, with both sexes contributing to the overall cellular diversity observed in the hypothalamus. Lastly, we clustered cells based on disease progression time points, where we observe some notable biases (**Figure 3c**). Specifically, oligodendrocytes are predominantly found in the clusters from postnatal day 30 (P30) and postnatal day 60 (P60), indicating a higher prevalence of these cells at these early stages of development. In contrast, the P150 time point is more prominent in clusters corresponding to neuronal cell populations, reflecting a shift in the cellular composition as the disease progresses.

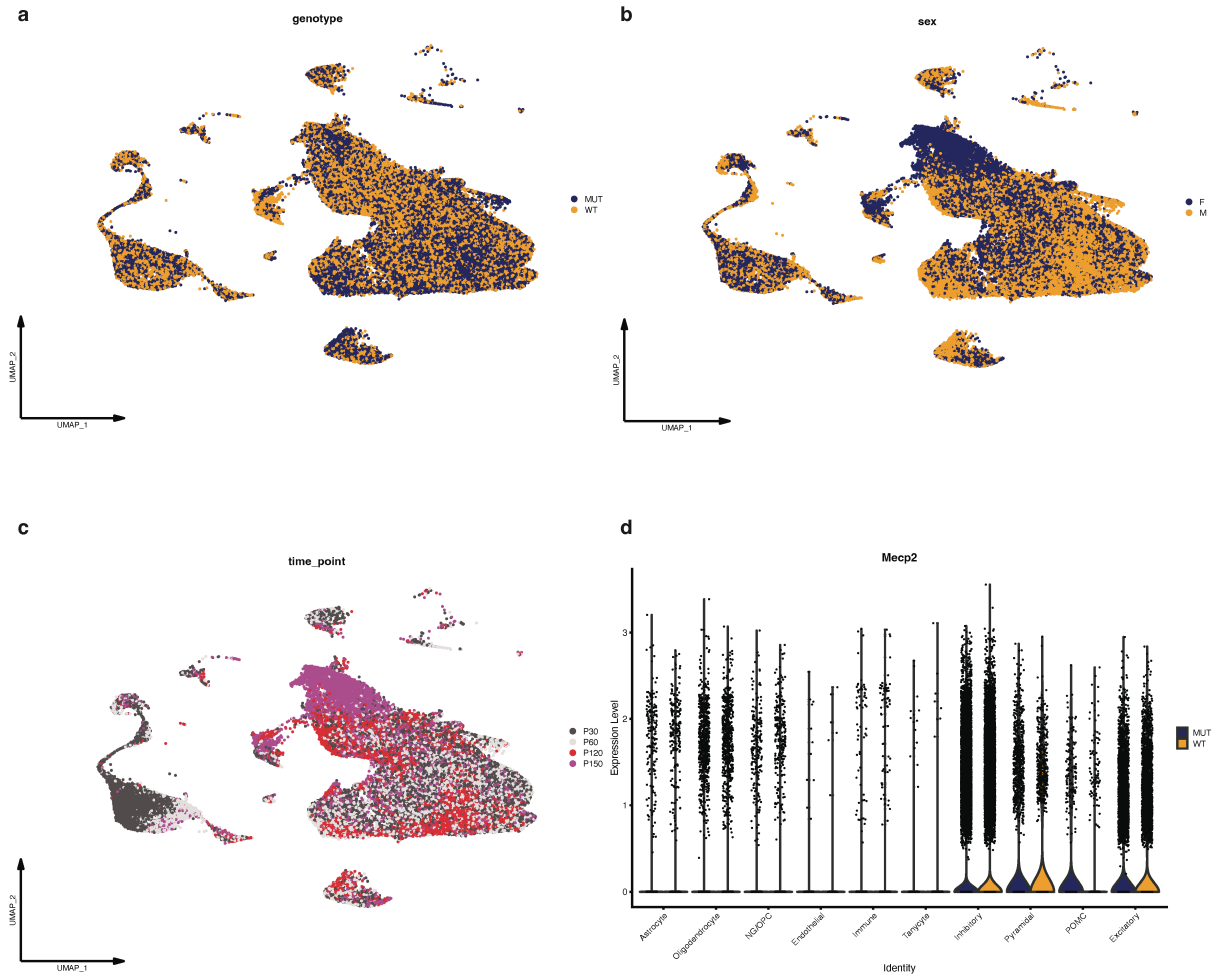


Figure 15: Cell clustering and *Mecp2* expression. **a, b, c.** UMAP plots of all cells based on genotype, sex and timepoint respectively. **d.** Violin plot of *Mecp2* expression in both WT and mutant hypothalami.

We analyzed the expression of *Mecp2* across different cell types in both wild-type and mutant mice (**Figure 3d**). The data confirms that *Mecp2* is expressed in all cell types but is most highly expressed in inhibitory and excitatory neurons. This differential expression underscores the critical role of *Mecp2* in these neuronal subtypes, which is pivotal in understanding the cellular impact of *Mecp2* deficiency and its implications for RTT pathology.

4.4.3 Differential gene expression in hypothalamic cell types across postnatal development

Analysis of differential gene expression across all identified hypothalamic cell types in mutant versus wild-type mice revealed significant temporal and cell-type-specific changes in both sexes (**Figure 4**). In male mice, the highest number of differentially expressed genes (DEGs) was observed in inhibitory neurons at P60, with a marked reduction by P120. (**Figure 4a**). In contrast, both excitatory neurons and immune cells showed the greatest number of DEGs at the presymptomatic stage of P30 in males. Other hypothalamic cell types in males, including astrocytes and oligodendrocytes, also exhibited significant DEGs, although to a lesser extent than neuronal and immune cell types.

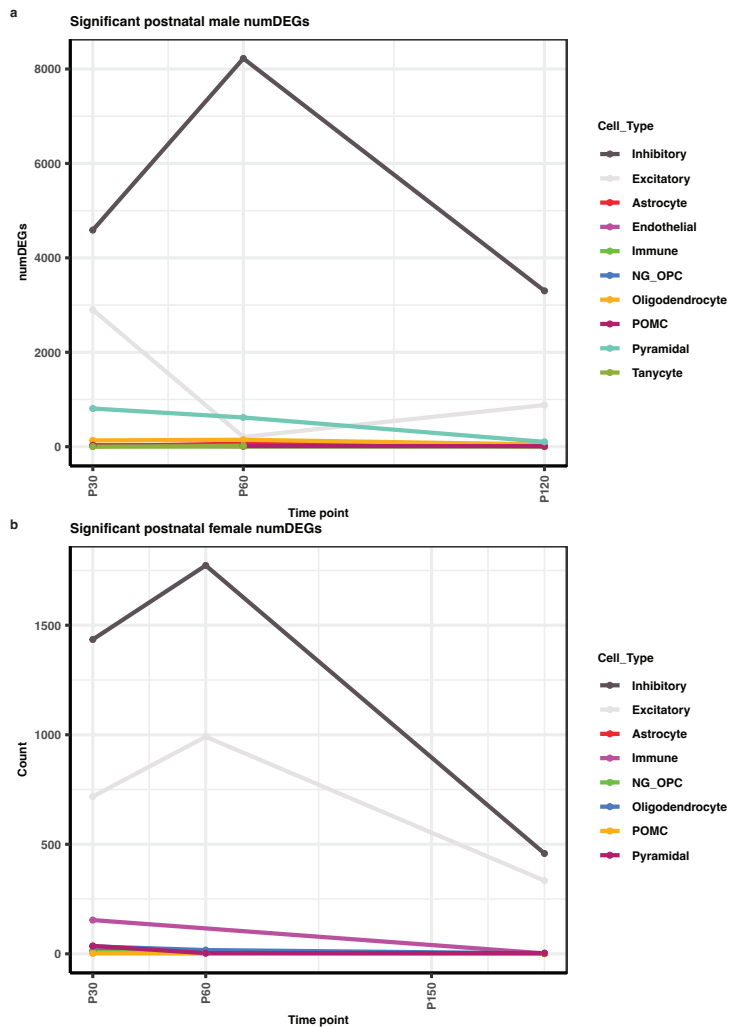


Figure 16: Differentially Expressed Genes (DEGs). **a.** Significant male DEGs over developmental time (adjusted p-value ≤ 0.05). **b.** Significant female DEGs over developmental time (adjusted p-value ≤ 0.05)

In female mice, a distinct temporal pattern was also observed, with the peak number of DEGs in both inhibitory and excitatory neurons occurring at P60 rather than P30, suggesting a delayed or prolonged period of gene expression changes compared to males (**Figure 4b**). This was followed by a decrease in the number of DEGs by P150. The DEG temporal dynamics within immune cell types were more similar between the sexes, with the most at P30 followed by a decline. Other cell types, such as POMC neurons and oligodendrocytes, showed relatively stable gene expression differences by genotype, with fewer DEGs across all time points. These findings

underscore the dynamic and cell-type-specific nature of gene expression changes in the hypothalamus during postnatal development. The pronounced differences in DEGs between males and females, specifically in neurons, reflects the sex-specific regulatory mechanisms that influence the development and function of the hypothalamus.

Further, we identified the top 10 DEGs in the hypothalamus with the largest log fold-change in excitatory and inhibitory neurons, since DEGs were predominant in these cell types (**Figure 5**). In male mice, genes such as *Tac2*, *Gpr151*, and *Ptprk* exhibited dysregulation at P30 in excitatory neurons, while *Oxt* and *Avp* were prominently dysregulated at later time points (P60 and P120 respectively) (**Figure 5a**). In females, genes including *Col27a1*, *Sparcl1*, and *Dlk1* showed differential expression at P60, with *Avp*, *Rps24* and *Nrgn* significantly altered at P150. Inhibitory neurons also demonstrated substantial changes, with *Junb*, *Auts2*, *Mid1* and *Fos* being notably down regulated at P30 in males, and *Pmcb* and *Nts* at P60 in females. The gene *AC149090.1*, interesting because of its role of phospholipid metabolism in the inner mitochondrial membrane, is dysregulated in opposite directions in male versus female inhibitory neurons from late symptomatic time points (**Figure 5b**). Top dysregulated genes for all other cell types in both males (**Supplemental Figure 1, 2, 3**) and females (**Supplemental Figure 4, 5, 6**) were also considered.

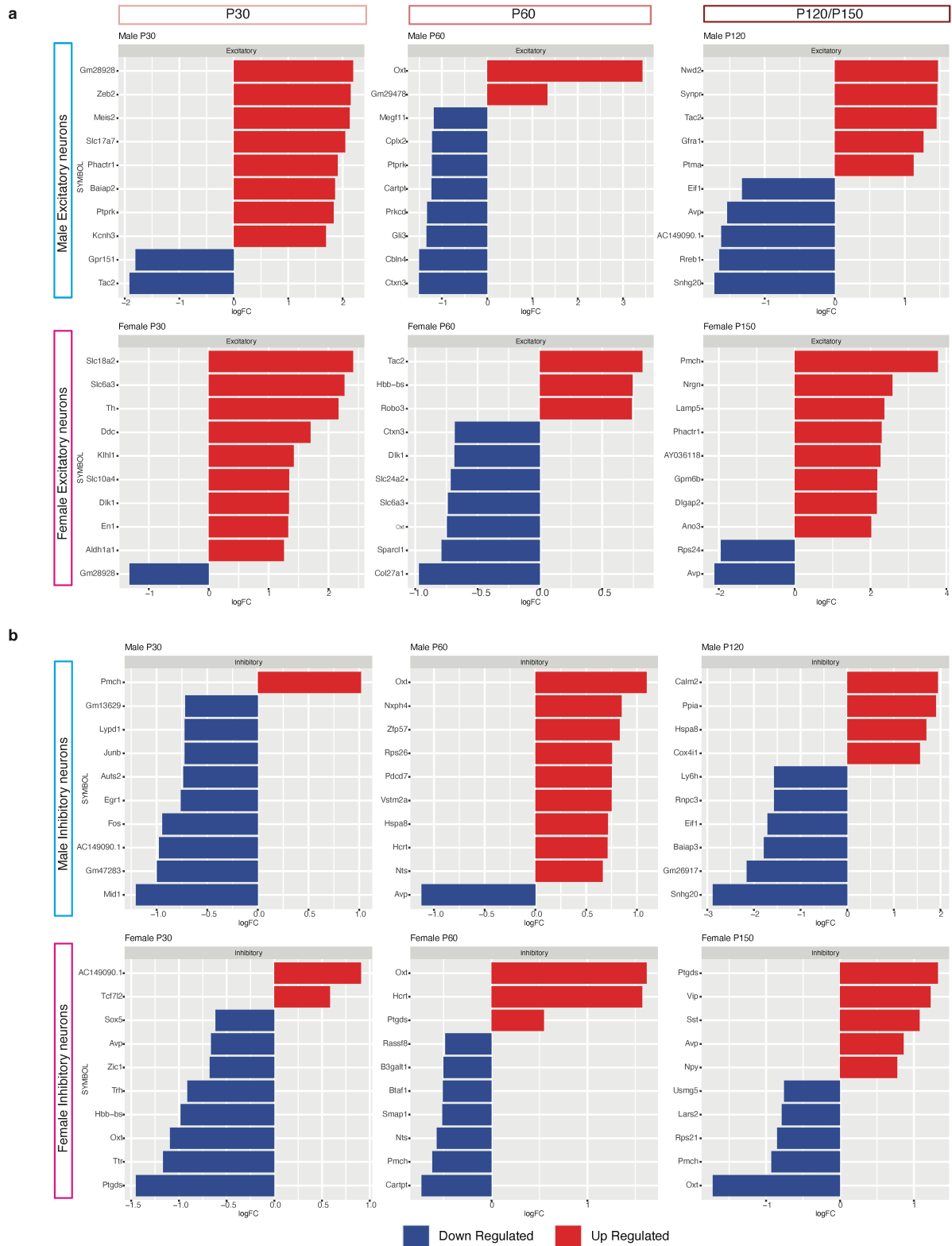


Figure 17: Top 10 DEGs. a. Significant DEGs (adjusted p-value ≤ 0.05) in Excitatory neurons. b. Significant DEGs (adjusted p-value ≤ 0.05) in Inhibitory neurons.

The observed dysregulation of genes known to be involved in neurotransmission, synaptic plasticity, and metabolic pathways and neuronal maturation suggests that these alterations may underlie the metabolic abnormalities and neurodevelopmental deficits. Furthermore, the consistent alteration of metabolic-related genes across developmental stages highlights the potential impact of hypothalamic dysregulation on the metabolic phenotype associated with RTT.

4.4.4 Functional pathway analysis of male and female excitatory and inhibitory Neurons

Gene Ontology (GO) analysis was performed on the statistically significant (adjusted p-value ≤ 0.05) dysregulated genes identified in both male and female mice to elucidate the functional enrichment of these alterations in excitatory and inhibitory neurons across the three postnatal stages, which had sufficient number of DEGs to perform these analyses.

In male excitatory neurons, the GO analysis revealed that DEGs were significantly enriched in chemical synaptic transmission and calcium ion binding molecular and biological processes across all time points (**Figure 6a**). The molecular functions of these dysregulated genes were found to be significantly associated with adenylyl ribonucleotide binding, tubulin binding, and protein kinase activity, which are vital for cellular energy balance, cytoskeletal integrity, and signal transduction. The cellular components linked to these dysregulated genes included the cytoskeleton, mitochondrial inner membrane, and nuclear body, indicating alterations in cellular structure and energy metabolism within the excitatory neurons (**Figure 6a**). Interestingly, DEGs from male inhibitory neurons were enriched in processes involved in mRNA synthesis and processing, both of which are processes known to be mediated by MeCP2 (**Figure 6b**). At all three male time points, mitochondrial inner membrane is enriched for cellular component, consistent with metabolic and energy dysregulation in RTT (**Figure 6b**). The molecular functions identified in these neurons were notably related to calcium ion binding, syntaxin binding, and voltage-gated cation channel activity, critical for

neurotransmitter release and synaptic plasticity. The dysregulated genes were also associated with cellular components like the axon, focal adhesion, and dendrite, which are essential for maintaining synaptic structure and function.

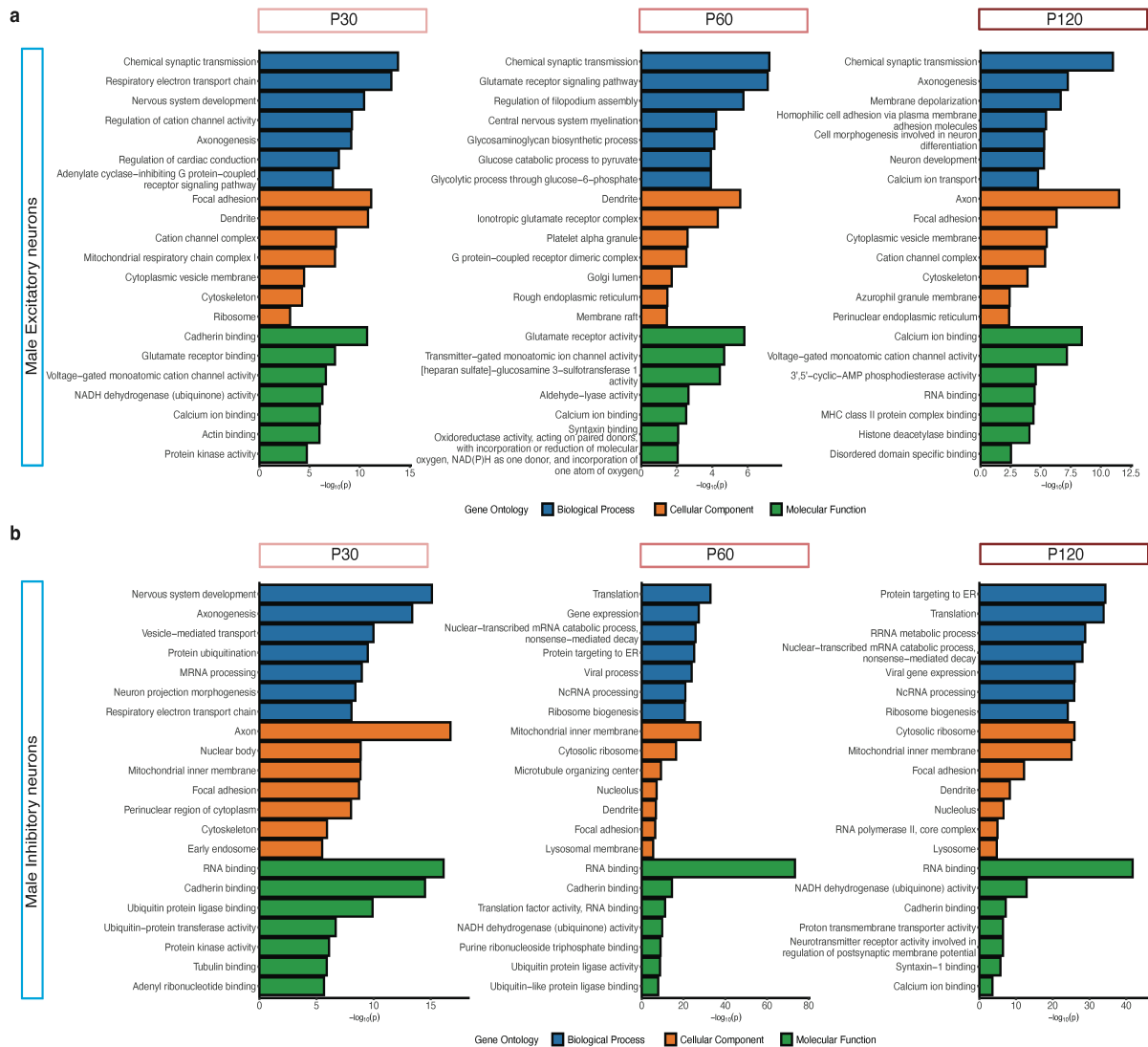


Figure 18: Top GO enrichment in male hypothalamic at P30, P60 and P120. a. Top significant GO terms in male excitatory neurons b. Top significant GO terms in male inhibitory neurons

In female neurons, the GO analysis similarly highlighted significant enrichment in nervous system development, axonogenesis, and neuron projection morphogenesis, with the strongest

enrichments in the early time points (P30 and P60) in excitatory neurons (**Figure 7a**). Interestingly, the dysregulated molecular functions show enrichment in axon and cation channel complex in all three disease stages in excitatory neurons.

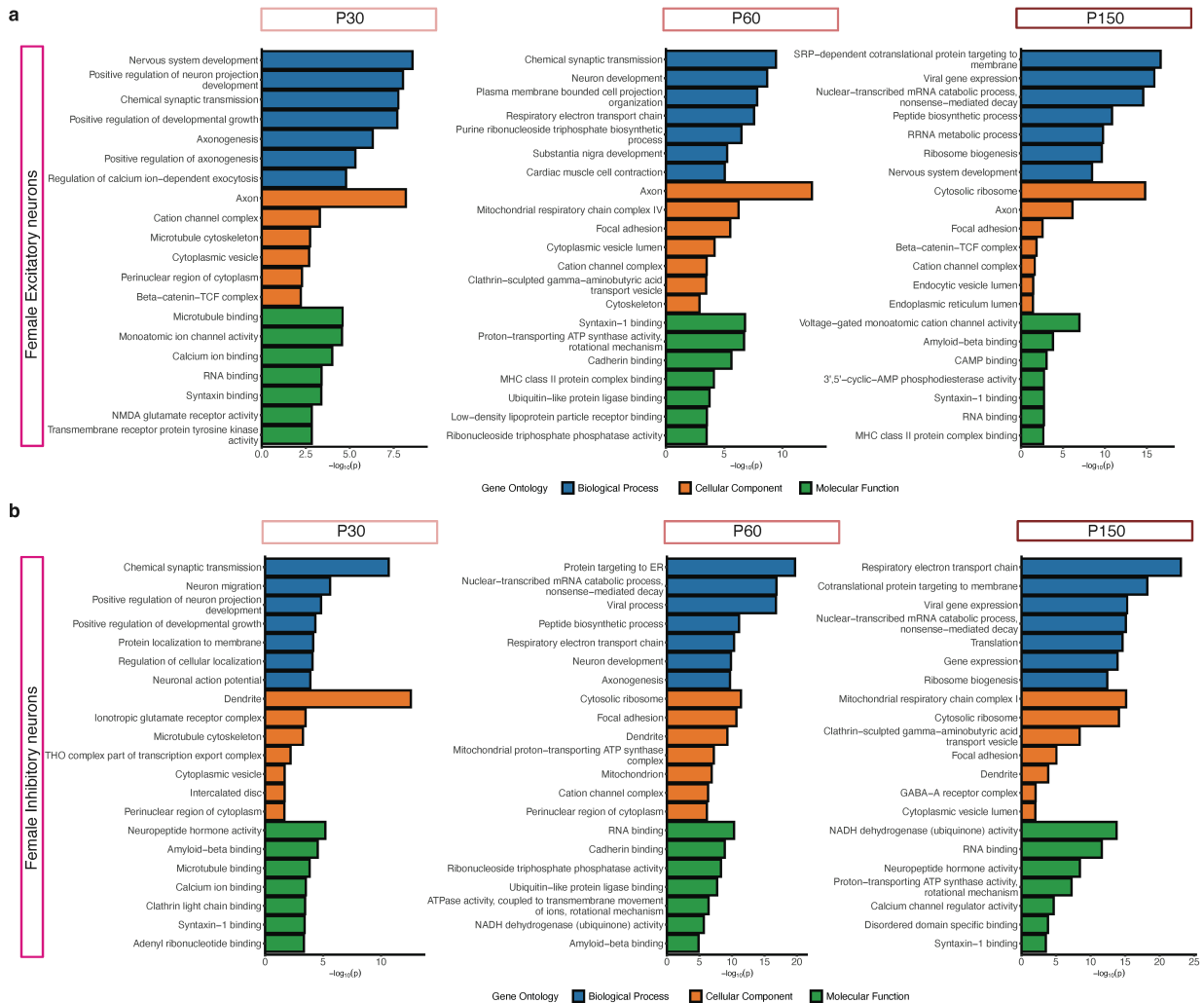


Figure 19: Top GO enrichment in male hypothalamic at P30, P60 and P150. a. Top significant GO terms in female excitatory neurons b. Top significant GO terms in female inhibitory neurons

In female inhibitory neurons, the GO analysis identified enrichment in involved in calcium signaling, mitochondrial homeostasis and RNA binding and processing (**Figure 7b**). The molecular functions related to these dysregulated genes included calcium ion binding, syntaxin binding, and ion

channel activity, further emphasizing the critical role the DEGs play in maintaining synaptic function and plasticity of inhibitory neurons (**Figure 7b**).

Overall, the GO analysis across both male and female neurons indicates that the dysregulated genes are involved in key biological processes, molecular functions, and cellular components essential for neural development, synaptic function, and cellular metabolism. These disruptions are likely contributing to the phenotypic abnormalities observed and provide insights into the molecular mechanisms underlying RTT.

4.5 Discussion

We investigated the transcriptional dysregulation in the hypothalamus of a progressively symptomatic *Mecp2^{fl}* mouse model of Rett syndrome, examining both cellular and sex-specific dimensions. Through single nuclei RNA sequencing (snRNA-seq 5'), we achieved high-resolution insights into the effects of MeCP2 deficiency across various cell types and developmental stages. Our results shed light on how MeCP2 dysfunction contributes to RTT pathology and reveal sex-specific differences in disease progression.

Our cell type labeling and clustering analysis identified ten distinct hypothalamic cell types, which were then used to investigate transcriptional changes across genotypes and developmental stages. Interestingly, cell clustering based on genotype and sex showed that while the overall distribution of cell types remained consistent, sex and disease progression influenced the prevalence of specific cell populations. Notably, oligodendrocytes were more prevalent at earlier developmental stages, while neuronal populations dominated in later stages, highlighting shifts in cellular composition as the disease progresses. *Mecp2* expression was observed across all cell types, with the highest levels in inhibitory and excitatory neurons. This underscores the critical role of MeCP2 in these neuronal

subtypes and emphasizes its significance in understanding the cellular impact of MeCP2 deficiency in RTT.

Our DEG analysis revealed notable temporal and cell-type-specific patterns. In male mice, the peak number of total DEGs were observed in inhibitory and excitatory neurons combined at P30, with a reduction in excitatory neurons by P60 and P120. In contrast, female mice exhibited a delayed peak in DEGs at P60, followed by a decrease by P150 in both inhibitory and excitatory neurons. This temporal variation suggests sex-specific regulatory mechanisms affecting gene expression and highlights the dynamic nature of gene dysregulation in RTT, particularly in females. This is consistent with the findings in the cerebral cortex showed by Sharifi et al.²¹

The top dysregulated genes identified in both sexes included those involved in neurotransmission, synaptic plasticity, and metabolic pathways. The consistent alteration of metabolic-related genes across developmental stages points to a significant impact of hypothalamic dysregulation on the metabolic phenotype associated with RTT. This finding aligns with previous research indicating that RTT involves broader metabolic disturbances beyond neuronal dysfunction^{22,23}.

Functional pathway analysis through Gene Ontology (GO) revealed significant enrichment in various biological processes and molecular functions in both male and female neurons. For male excitatory neurons, enriched pathways were associated with chemical synaptic transmission and cellular energy balance, reflecting alterations in synaptic structure and function. In male inhibitory neurons, processes related to mRNA synthesis and mitochondrial function were prominently dysregulated. In females, GO analysis highlighted disruptions in nervous system development and neuron projection morphogenesis, with specific emphasis on early developmental stages.

Overall, the dysregulated genes in both sexes were implicated in crucial processes such as neural development, synaptic function, and cellular metabolism. These disruptions likely contribute to the phenotypic abnormalities observed in RTT and provide new insights into the molecular mechanisms

underlying the disease in the hypothalamus. Our findings underscore the importance of considering both cellular and sex-specific factors when studying RTT, and they offer potential avenues for targeted therapeutic strategies aimed at addressing the complex nature of RTT.

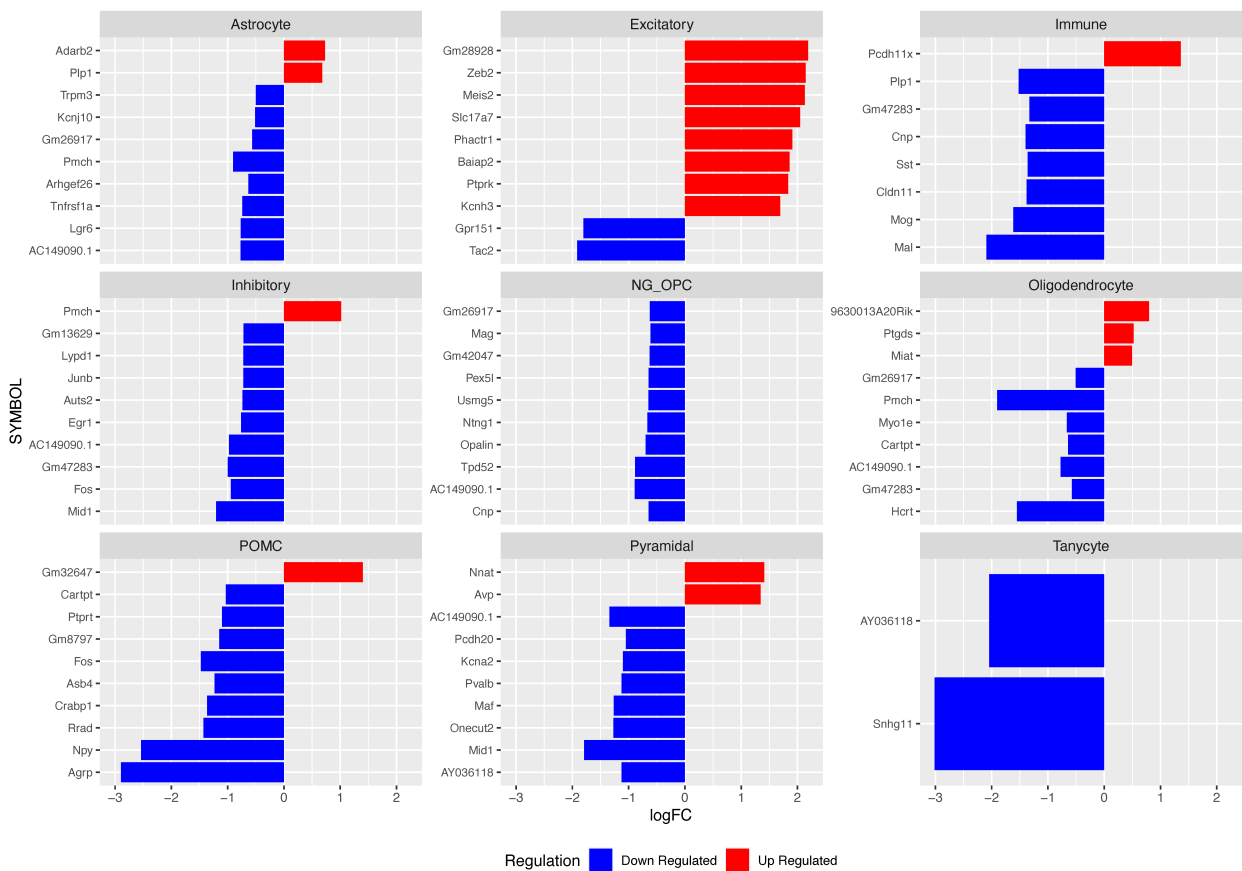
4.6 References

1. Zoghbi, H. Y. Postnatal Neurodevelopmental Disorders: Meeting at the Synapse? *Science* (2003) doi:10.1126/science.1089071.
2. Amir, R. E. *et al.* Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl- CpG-binding protein 2. *Nat. Genet.* (1999) doi:10.1038/13810.
3. Sharifi, O. & Yasui, D. H. *The Molecular Functions of MeCP2 in Rett Syndrome Pathology. Frontiers in Genetics* vol. 12 (2021).
4. Chahrour, M. *et al.* MeCP2, a key contributor to neurological disease, activates and represses transcription. *Science* (2008) doi:10.1126/science.1153252.
5. Liu, Y. *et al.* MECP2 directly interacts with RNA polymerase II to modulate transcription in human neurons. *Neuron* **112**, 1943-1958.e10 (2024).
6. Liu, T., Xu, Y., Yi, C.-X., Tong, Q. & Cai, D. The hypothalamus for whole-body physiology: from metabolism to aging. *Protein Cell* **13**, 394–421 (2022).
7. Ciernia, A. V. *et al.* MeCP2 isoform e1 mutant mice recapitulate motor and metabolic phenotypes of Rett syndrome. *Hum. Mol. Genet.* **27**, 4077–4093 (2018).
8. Kyle, S. M., Vashi, N. & Justice, M. J. Rett syndrome: A neurological disorder with metabolic components. *Open Biol.* (2018) doi:10.1098/rsob.170216.
9. Ribeiro, M. C. & MacDonald, J. L. Sex differences in Mecp2-mutant Rett syndrome model mice and the impact of cellular mosaicism in phenotype development. *Brain Research* vol. 1729 (2020).

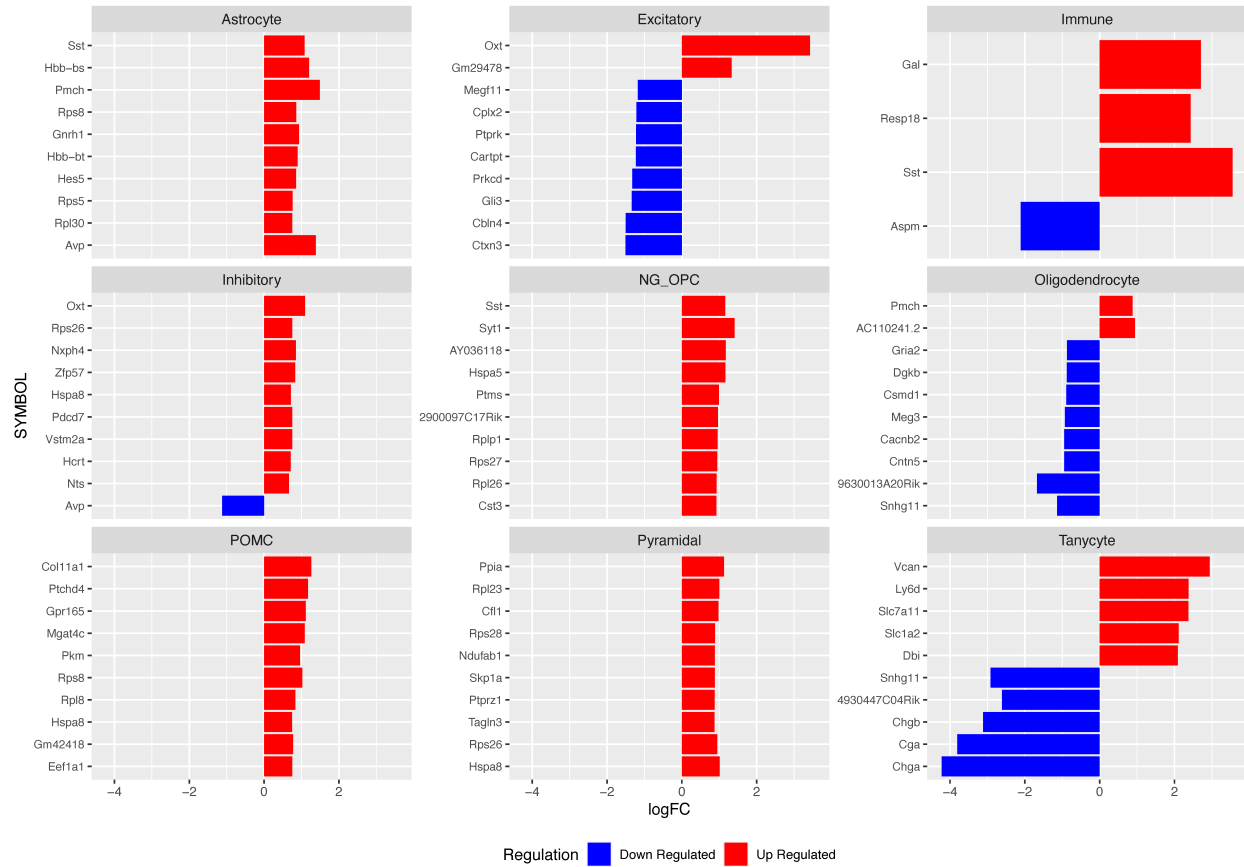
10. Braunschweig, D., Simcox, T., Samaco, R. C. & LaSalle, J. M. X-chromosome inactivation ratios affect wild-type MeCP2 expression within mosaic Rett syndrome and *Mecp2*^{-/+} mouse brain. *Hum. Mol. Genet.* (2004) doi:10.1093/hmg/ddh142.
11. Guy, J., Hendrich, B., Holmes, M., Martin, J. E. & Bird, A. A mouse *Mecp2*-null mutation causes neurological symptoms that mimic rett syndrome. *Nat. Genet.* **27**, 322–326 (2001).
12. Moretti, P. *et al.* Learning and memory and synaptic plasticity are impaired in a mouse model of Rett syndrome. *J. Neurosci.* **26**, 319–327 (2006).
13. Collins, A. L. *et al.* Mild overexpression of MeCP2 causes a progressive neurological disorder in mice. *Hum. Mol. Genet.* **13**, 2679–2689 (2004).
14. Ballas, N., Lioy, D. T., Grunseich, C. & Mandel, G. Non-cell autonomous influence of MeCP2-deficient glia on neuronal dendritic morphology. *Nat. Neurosci.* (2009) doi:10.1038/nn.2275.
15. Herring, C. A. *et al.* Human prefrontal cortex gene regulatory dynamics from gestation to adulthood at single-cell resolution In brief II Human prefrontal cortex gene regulatory dynamics from gestation to adulthood at single-cell resolution. *Cell* **185**, 4428–4447 (2022).
16. Cohen, S. *et al.* Genome-Wide Activity-Dependent MeCP2 Phosphorylation Regulates Nervous System Development and Function. *Neuron* **72**, 72–85 (2011).
17. Kishi, N. & Macklis, J. D. MeCP2 functions largely cell-autonomously, but also non-cell-autonomously, in neuronal maturation and dendritic arborization of cortical pyramidal neurons. *Exp. Neurol.* **222**, 51–58 (2010).
18. Jin, X.-R., Chen, X.-S. & Xiao, L. MeCP2 Deficiency in Neuroglia: New Progress in the Pathogenesis of Rett Syndrome. *Front. Mol. Neurosci.* **10**, (2017).
19. Buchovecky, C. M. *et al.* A suppressor screen in *Mecp2* mutant mice implicates cholesterol metabolism in Rett syndrome. *Nat. Genet.* **45**, 1013–1020 (2013).

20. Steuernagel, L. *et al.* HypoMap—a unified single-cell gene expression atlas of the murine hypothalamus. *Nat. Metab.* **4**, 1402–1419 (2022).
21. Sharifi, O. *et al.* Sex-specific single cell-level transcriptomic signatures of Rett syndrome disease progression. *BioRxiv Prepr. Serv. Biol.* 2024.05.16.594595 (2024)
doi:10.1101/2024.05.16.594595.
22. Neul, J. L. *et al.* Metabolic Signatures Differentiate Rett Syndrome From Unaffected Siblings. *Front. Integr. Neurosci.* **14**, (2020).
23. Vogel Ciernia, A. *et al.* MeCP2 isoform e1 mutant mice recapitulate motor and metabolic phenotypes of Rett syndrome. *Hum. Mol. Genet.* (2018) doi:10.1093/hmg/ddy301.

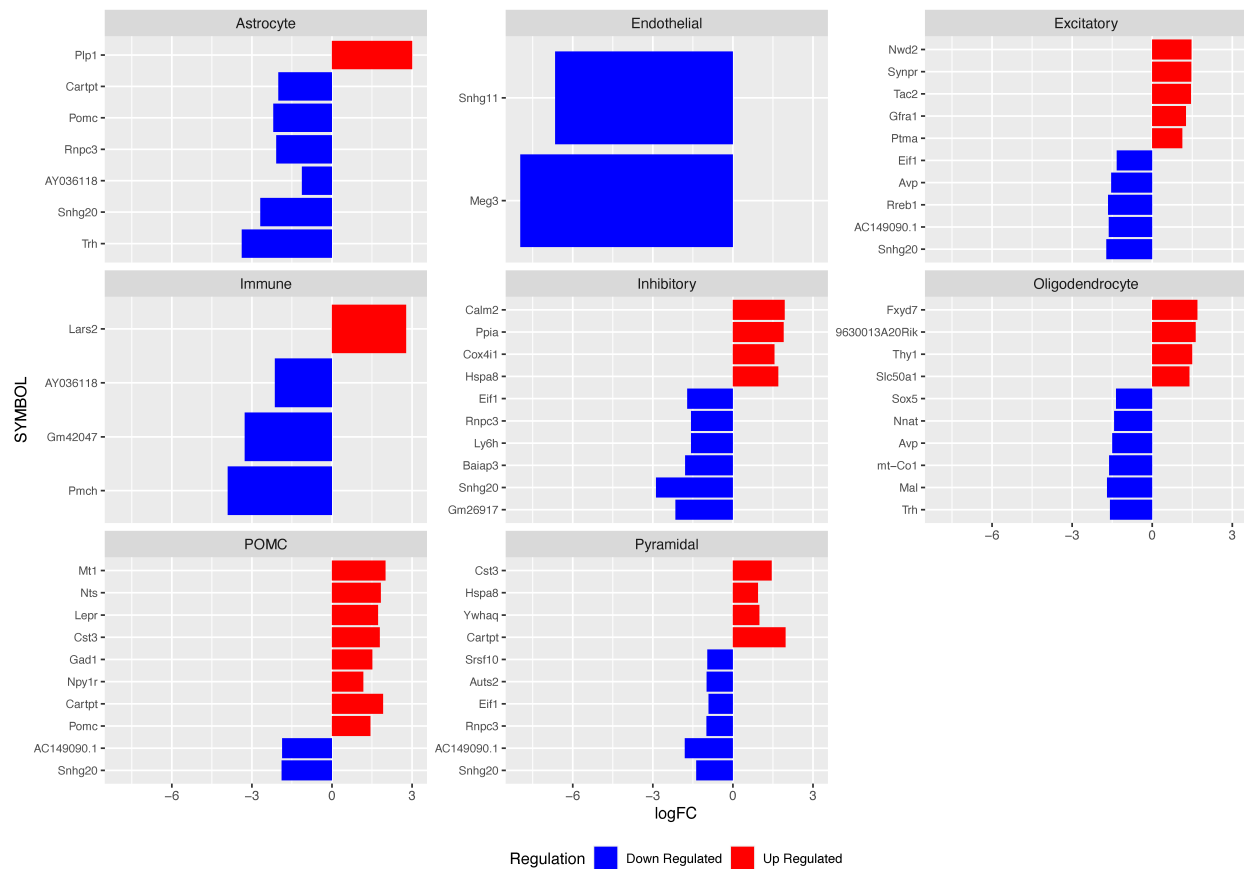
4.7 Supplemental Figures



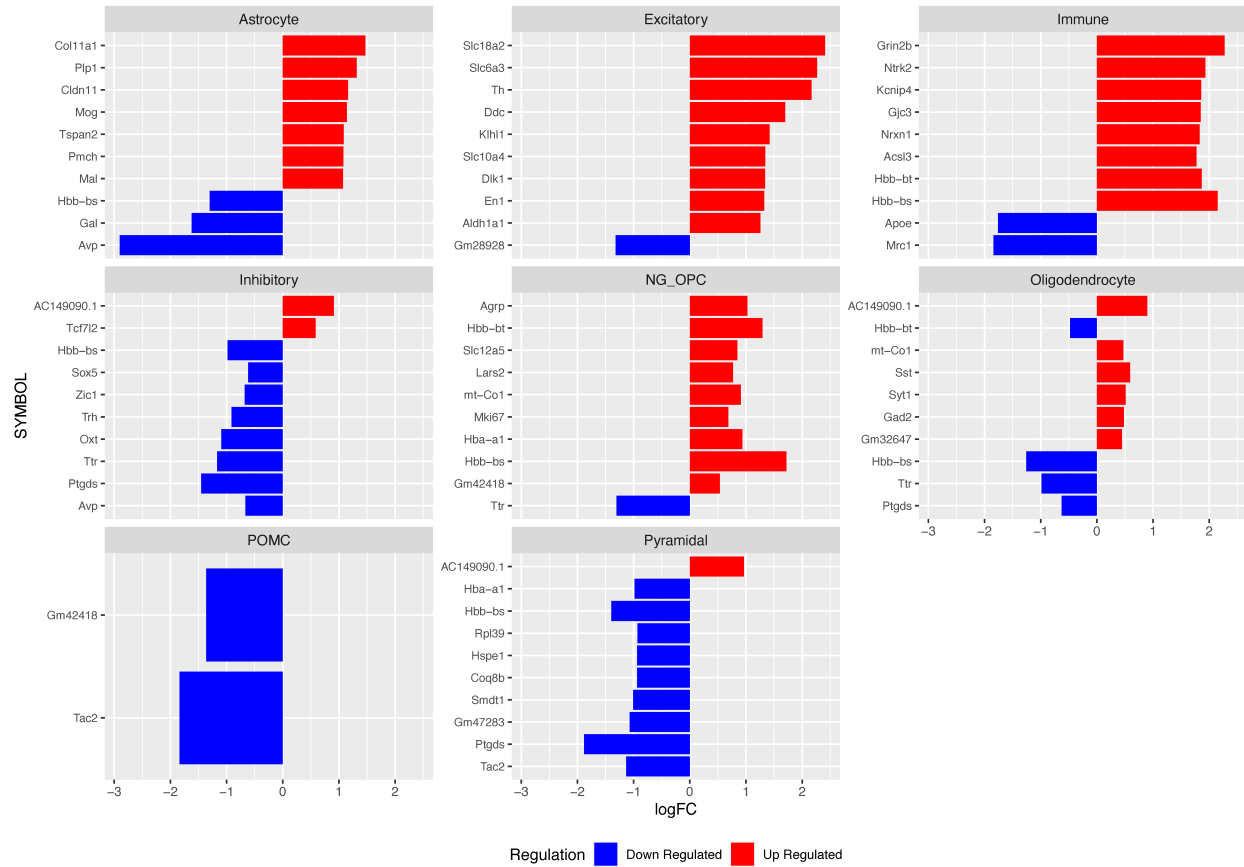
Supplemental Figure 1: Top 10 significant DEGs in males at P30 in all cell types



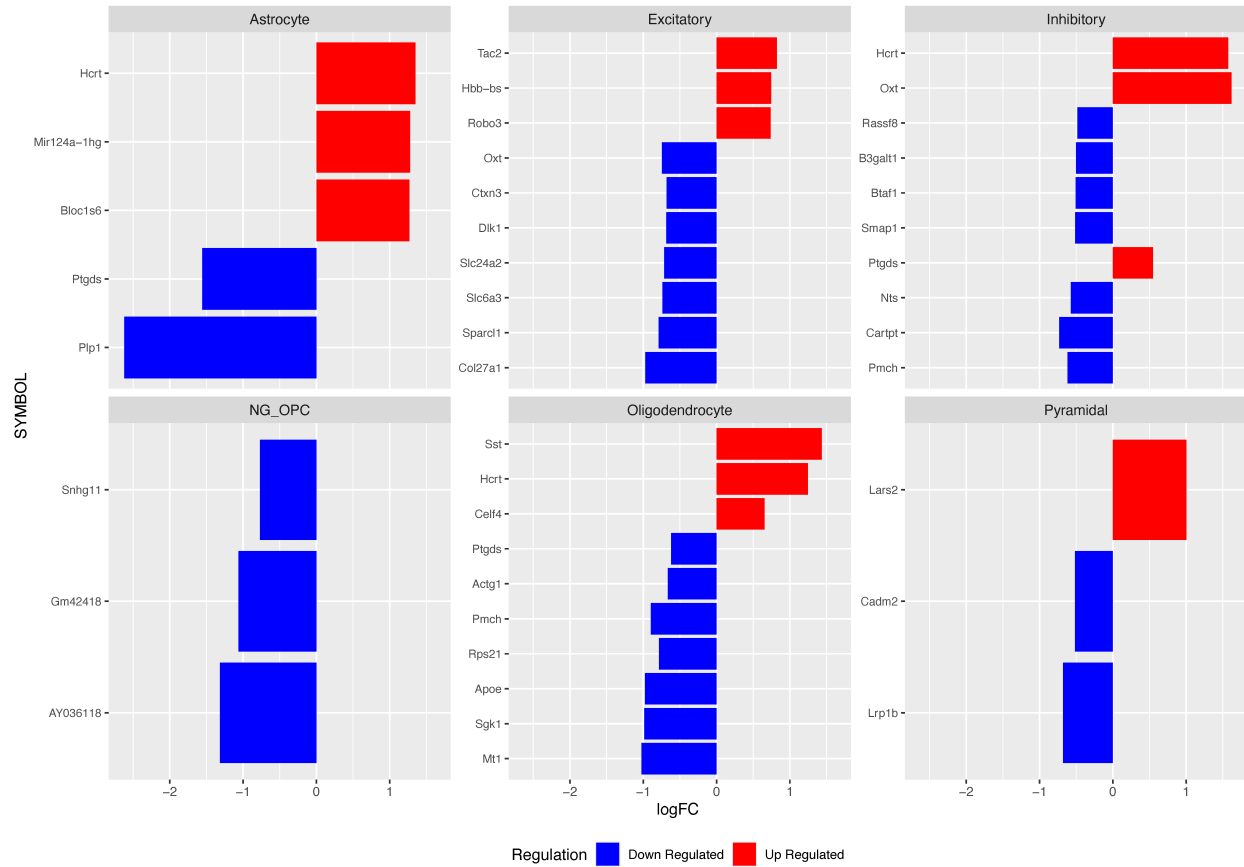
Supplemental Figure 2: Top 10 significant DEGs in males at P60 in all cell types



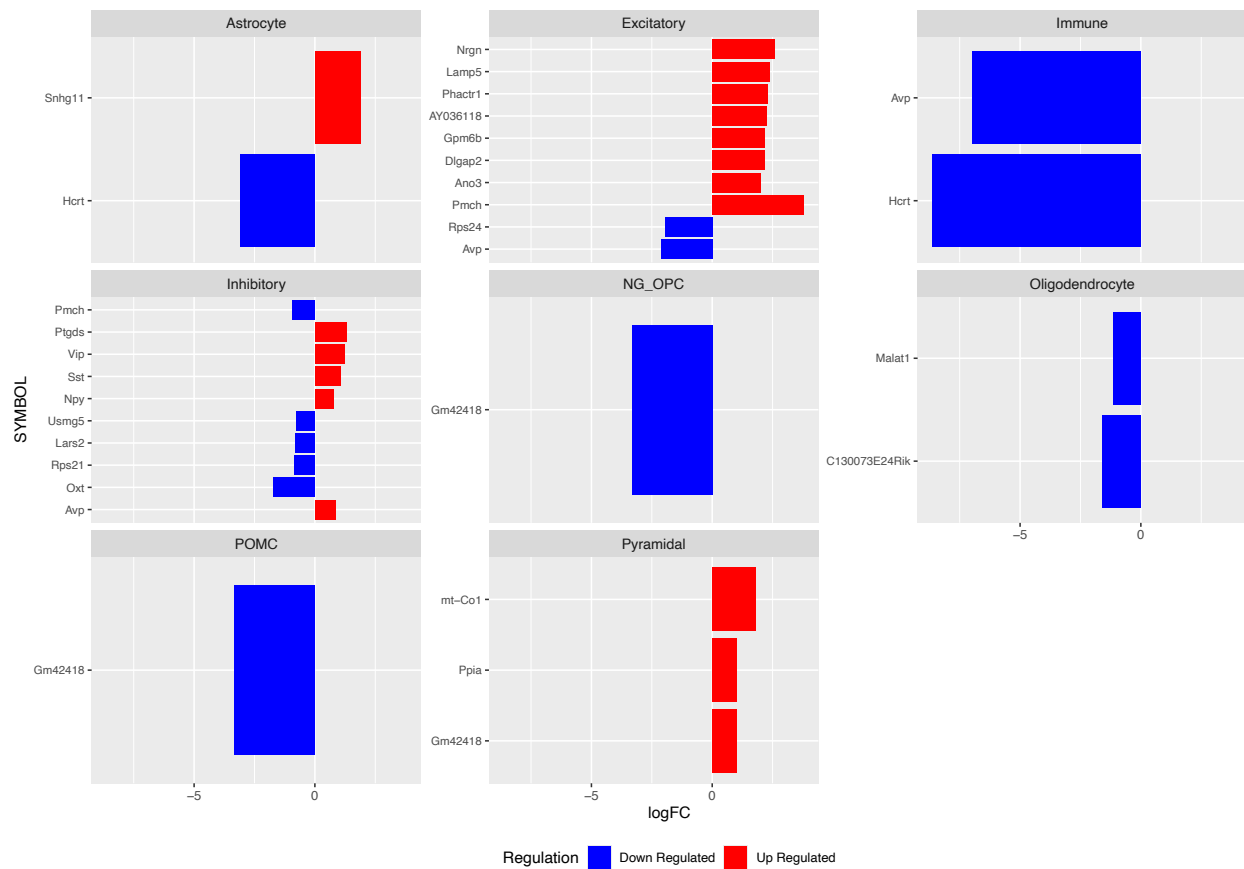
Supplemental Figure 3: Top 10 significant DEGs in males at P120 in all cell types



Supplemental Figure 4: Top 10 significant DEGs in females at P30 in all cell types



Supplemental Figure 5: Top 10 significant DEGs in females at P60 in all cell types



Supplemental Figure 6: Top 10 significant DEGs in females at P150 in all cell types