PhD Program in Neurosciences

OMICS PROFILING OF CORTICAL PROGENITOR CELLS IN EVOLUTION AND CANCER

Doctoral thesis presented by

Rafael Soler Ortuño

Thesis Director: Dr. Víctor Borrell Franco

Instituto de Neurociencias

Universidad Miguel Hernández de Elche

- 2025 -

CSIC UNIVERSITAS Miguel Hernández

INSTITUTO DE NEUROCIENCIAS

EXCELENCIA SEVERO OCHOA

2014 - 2027

UNIVERSITAS Miguel Hernández

CSIC UNIVERSITAS Miguel Hernández
INSTITUTO DE NEUROCIENCIAS

EXCELENCIA SEVERO OCHOA
2014 - 2027

QUALITY CRITERION

This Doctoral Thesis, entitled '**Omics Profiling of Cortical Progenitor Cells in Evolution and Cancer**' is submitted as conventional thesis with the following quality indicator:

Chinnappa, K., Cárdenas, A., Prieto-Colomina, A., Villalba, A., Márquez-Galera, Á., Soler, R., Nomura, Y., Llorens, E., Tomasello, U., López-Atalaya, J. P., & Borrell, V. (2022). *Secondary loss of miR-3607 reduced cortical progenitor amplification during rodent evolution*. **Science Advances**, 8(2). https://doi.org/10.1126/sciadv.abj4010

UNIVERSITAS Miguel Hernández

CSIC UNIVERSITAS Miguel Hernández
INSTITUTO DE NEUROCIENCIAS

EXCELENCIA SEVERO OCHOA
2014 - 2027

DIRECTOR'S REPORT

Sant Joan d'Alacant, 17 de junio 2025

Dr. Víctor Borrell Franco, director of the doctoral thesis entitled '**Omics Profiling of Cortical Progenitor Cells in Evolution and Cancer**'

INFORMS:

That Mr. Rafael Soler Ortuño has carried out under my supervision the work entitled '**Omics Profiling of Cortical Progenitor Cells in Evolution and Cancer**' in accordance with the terms and conditions defined in his Research Plan and in accordance with the Code of Good Practice of the *Universidad Miguel Hernández de Elche*, fulfilling the objectives satisfactorily for its public defense as a doctoral thesis.

Which I sign for the appropriate effects in San Juan de Alicante, on 17th of June of 2025.

Thesis Director

Dr. *Víctor Borrell Franco*

Report of the Coordinator of the Academic Committee of the Doctoral Program

Sant Joan d'Alacant, 17 de junio 2025

Ms. Cruz Morenilla Palao, Coordinator of the PhD Program in Neurosciences at the Institute of Neurosciences of Alicante, a joint center of the *Universidad Miguel Hernández de Elche* (UMH) and the *Consejo Superior de Investigaciones Científicas* (CSIC),

INFORMS:

That Mr. Rafael Soler Ortuño has carried out under the supervision of our Doctoral Program the work entitled '**Omics Profiling of Cortical Progenitor Cells in Evolution and Cancer**' in accordance with the terms and conditions defined in his Research Plan and in accordance with the Code of Good Practices of the *Universidad Miguel Hernández de Elche*, fulfilling the objectives satisfactorily for its public defense as a doctoral thesis.

Which I sign for the appropriate effects, in San Juan de Alicante, on 17th of June of 2025.

Dr. Cruz Morenilla Palao

Coordinator of the Doctoral Program in Neurosciences

FUNDING ENTITIES

This Doctoral Thesis, entitled "**Omics Profiling of Cortical Progenitor Cells in Evolution and Cancer**" has been carried out at the *Instituto de Neurociencias*, a joint center participated by the *Universidad Miguel Hernández de Elche* and the *Consejo Superior de Investigaciones Científicas*, with:

Grant **PRE2019-090531** funded by MICIU/AEI/10.13039/501100011033 and by "*FSE invierte en tu futuro*"

INDEX

ABBREVIATIONS

APC – Adenomatous Polyposis Coli
aRGC(s) – apical Radial Glial Cell(s)
ARHGAP11A / ARHGAP11B – Rho GTPase-Activating Protein 11A / 11B
ASCL1 – Achaete-Scute Family bHLH Transcription Factor 1
AT/RT – Atypical Teratoid/Rhabdoid Tumor(s)
ATAC-seq – Assay for Transposase-Accessible Chromatin using sequencing
ATF6 – Activating Transcription Factor 6
ATRX – Alpha-Thalassemia/Mental Retardation Syndrome X-linked
BAM – Binary Alignment Map
BCL6 – B-Cell Lymphoma 6
BCOR – BCL6 Corepressor
bHLH – Basic-Helix-Loop-Helix
BMP – Bone Morphogenetic Protein
BP(s) – Basal Progenitor(s)
bRGC(s) – basal Radial Glial Cell(s) (also referred to as oRG for “outer Radial Glial”)
C19MC – Chromosome 19 microRNA Cluster
C2H2 – Cysteine2–Histidine2 (zinc-finger motif)
CACNA1 – Calcium Voltage-Gated Channel Subunit Alpha1
CBT – Congenital Brain Tumors
CCND1/2 – Cyclin D1/2
cCRE – Candidate Cis-Regulatory Element
CD – Cytoplasmic Domain
CDC25A – Cell Division Cycle 25A
CDK6 – Cyclin-Dependent Kinase 6
CDKN1A – Cyclin Dependent Kinase Inhibitor 1A (p21)
CDKN2C (p18INK4c) – Cyclin Dependent Kinase Inhibitor 2C
ChIP-seq – Chromatin Immunoprecipitation Sequencing
CNS – Central Nervous System
CNS5 – 5th edition of the WHO Classification of CNS Tumors
CRE(s) – cis-Regulatory Element(s)
CREB3 – cAMP Response Element Binding Protein 3
CRISPR – Clustered Regularly Interspaced Short Palindromic Repeats
CTCF – CCCTC-Binding Factor
CTNNA1/2 – Catenin Alpha 1/2
CTNNB1 – Catenin Beta-1
CX/NCx – Neocortex
DAR – Differentially Accessible Region
DBD – DNA binding domain
DEG(s) – Differentially Expressed Gene(s)
DGE – Differential Gene Expression
DICER1 – Dicer 1
DLL1 – Delta-like 1
DMRTA1 – Doublesex and mab-3 Related Transcription factor a1
DN – Direct Neurogenesis
DNase-seq – DNase I Hypersensitive Sites Sequencing
DNMT3B – DNA Methyltransferase 3 Beta
DRG – Downregulated
Drosha – Ribonuclease Type III Drosha
DVR – Dorsal Ventricular Ridge
E – Embryonic Day
ED – Extracellular Domain
EMT – Epithelial-Mesenchymal Transition
ENCODE – Encyclopedia of DNA Elements

EOMES (TBR2) – T-box Brain Protein 2
ER – Endoplasmic Reticulum
ETMR – Embryonal Tumor with Multilayered Rosettes
FACS – Fluorescence-Activated Cell Sorting
FDR – False Discovery Rate
FOXP1 – Forkhead Box P1
FPKM – Fragments Per Kilobase of transcript per Million mapped reads
FZD8/9 – Frizzled Class Receptor 8/9
GBM – Glioblastoma
GFF – General Feature Format
GFP – Green Fluorescent Protein
GLI – Gli Family Zinc Finger Transcription Factors
GO – Gene Ontology
GREB1 – Growth Regulation by Estrogen in Breast Cancer 1
GRIA1 – Glutamate Ionotropic Receptor AMPA Type Subunit 1
GSEA – Gene Set Enrichment Analysis
GTF – Gene Transfer Format
H&E – Hematoxylin and Eosin
H3 – Histone H3
H3K27ac – Histone H3 Lysine 27 Acetylation
H3K4me1 / H3K4me3 – Histone H3 Lysine 4 Mono-/Trimethylation
HAR(s) – Human Accelerated Region(s)
HES1 – Hairy and Enhancer of Split-1
HGG – High-Grade Gliomas
HGG IDH – High-Grade Gliomas Isocitrate Dehydrogenase
HGG WT – High-Grade Gliomas Histone H3-wildtype
HGNET – High-Grade NeuroEpithelial Tumor
Hi-C – High-Throughput Chromosome Conformation Capture
HMGA – High Mobility Group A
HMM – Hidden Markov Model
HOPX – Homeodomain-only Protein
HPA – Human Protein Atlas
ICA domain – Intramolecular Chaperone Auto-processing domain
ICD – Intracellular Domain
IGDCC3 – Immunoglobulin Superfamily DCC Subclass Member 3
IGF1 – Insulin-Like Growth Factor 1
IGF2BP1 – Insulin-Like Growth Factor 2 Binding Protein 1
IGF1R – Insulin-Like Growth Factor 1 Receptor
IGV – Integrative Genomics Viewer
IL-6 – Interleukin 6
IN – Indirect Neurogenesis
INM – Interkinetic Nuclear Migration
INSM1 – Insulinoma-associated 1
IPC(s) – Intermediate Progenitor Cell(s)
IRS1/2 – Insulin Receptor Substrate 1/2
ISN – International Society of Neuropathology
ITD – Internal Tandem Duplication
ITGB2 – Integrin Beta 2
JAG1/2 – Jagged1/2
JPH2 – Junctophilin 2
JUN – Jun Proto-Oncogene, AP-1 Transcription Factor Subunit
KEGG – Kyoto Encyclopedia of Genes and Genomes
KNN – K-Nearest Neighbors
LD – Lumenal Domain
Let-7 – micro–RNA Let-7 family

LGG – Low-Grade Gliomas
LIN28A/B – Lin-28 Homolog A/B
LS – Lateral Sulcus
MACS2 – Model-Based Analysis of ChIP-Seq
MAPK/ERK – Mitogen-Activated Protein Kinase / Extracellular Signal-Regulated Kinase
MB – Medulloblastoma
MGMT – O-6-Methylguanine-DNA Methyltransferase
MIR100HG – MicroRNA 100 Host Gene
MIR-17-92 (MIR17HG) – microRNA-17-92 Cluster
miRNA – microRNA
MIR3607 – microRNA 3607
MORN – Membrane Occupation and Recognition Nexus
MRI – Magnetic Resonance Imaging
MSigDB – Molecular Signatures Database
MTF(s) – Membrane-Bound Transcription Factor(s)
MYCN – N-myc Proto-Oncogene
MYRF – Myelin Regulatory Factor
NBPF – Neuroblastoma Breakpoint Family
NEC(s) – Neuroepithelial stem cell(s)
NEO1 – Neogenin 1
NES – Normalized Enrichment Score
NEUROG2 – Neurogenin-2
NFIA/B – Nuclear Factor I A/B
NFR – Nucleosome-Free Region
NGS – Next-Generation Sequencing
NIH – National Institutes of Health
NKX2.2 – NK2 Homeobox 2
NLS – Nuclear Localization Signal
NOTCH2NL – Notch2 N-Terminal Like
NR2E1 – Nuclear Receptor Subfamily 2, Group E, Member 1
NR2F1 – Nuclear Receptor Subfamily 2, Group F, Member 1
OB – Olfactory Bulb
OLIG2 – Oligodendrocyte Transcription Factor 2
ORA – Over-Representation Analysis
OSVZ – Outer Subventricular Zone
OTHER EBT – Other Embryonic Brain Tumors
OXPHOS – Oxidative Phosphorylation
Padj – Adjusted P-value
PAGA – Partition-based Graph Abstraction
PAX6 – Paired-box transcription factor 6
PCA – Principal Component Analysis
PH3 – Phosphohistone H3
PI3K/AKT/mTOR – Phosphoinositide 3-Kinase / Protein Kinase B / Mechanistic Target of Rapamycin
PIK3C2B – Phosphatidylinositol-4-Phosphate 3-Kinase Catalytic Subunit Type 2 Beta
PLAGL1 – Pleiomorphic Adenoma Gene 1
PLSCR1/2 – Phospholipid Scramblase 1/2
PM – Palmitoylation Motif
PMG – Polimicrogiria
PNET – Primitive NeuroEctodermal Tumor
PPP1R17 – Protein Phosphatase 1 Regulatory Subunit 17
PRD – Proline-Rich Domain
PRTG – Protogenin
PTCH – Patched
PWM – Position Weight Matrix

RAS – Rat Sarcoma Viral Oncogene Family
RBP(s) – RNA-Binding Protein(s)
RGC(s) – Radial Glial Cell(s)
RISC – RNA-Induced Silencing Complex
RLVZ – Rhombic Lip Ventricular Zone
ROBO1/2 – Roundabout Guidance Receptor 1/2
RPKM – Reads Per Kilobase of transcript per Million mapped reads
RPM – Reads Per Million
RTPS – Rhabdoid Tumor Predisposition Syndrome
SALL1 – Spalt Like Transcription Factor 1
saPC(s) – Subapical Progenitor Cell(s)
scRNA-seq – Single-Cell RNA Sequencing
SFRP1 – Secreted Frizzled-Related Protein 1
SG – Splenial Gyrus
sgRNA – single-guide RNA
SHH – Sonic Hedgehog
SMARCA4/SMARCB1 – SWI/SNF Related, Matrix-Associated, Actin-Dependent Regulator of Chromatin Subfamily A/B, Member 4/1
SMO – Smoothened
SNP(s) – Short Neural Precursor(s)
SOX2 – SRY-Box Transcription Factor 2
SP – Septum
SREBF1/2 – Sterol Regulatory Element Binding Transcription Factor 1/2
SVZ – Subventricular Zone
SWI/SNF – SWItch/Sucrose Non-Fermentable (Chromatin-Remodeling Complex)
TAD(s) – Topologically Associated Domain(s)
TBR2 – T-box transcription factor 2 (also known as EOMES)
TD – Transactivation domain
TENM1 – Teneurin Transmembrane Protein 1
TF(s) – Transcription Factor(s)
TFBS – Transcription Factor Binding Site
TGF-β – Transforming Growth Factor Beta
TMD – Transmembrane Domain
TMM – Trimmed Mean of M-values
TMZ – Temozolomide
TP53 – Tumor Protein p53
tRG – truncated Radial Glia
TRIM71 – Tripartite Motif Containing 71
TSS – Transcription Start Site
TTP – Tristetraprolin
TTYH1 – Tweety Family Member 1
UMAP – Uniform Manifold Approximation and Projection
UMI – Unique Molecular Identifier
UPR – Unfolded Protein Response
URG – Upregulated Genes
UTR (5/3′ UTR) – Untranslated Region (5/3′ Untranslated Region)
VZ – Ventricular Zone
WHO – World Health Organization
WNT – Wingless-Related Integration Site
XBP1 – X-Box Binding Protein 1
ZNRF2 – Zinc And Ring Finger 2

ABSTRACT

The evolution of cortical neurogenesis has been fundamental to the development of the complex mammalian brain. Disruptions in these processes underlie various neurodevelopmental disorders and cancers, such as embryonal tumors with multilayered rosettes (ETMRs). This thesis integrates comparative transcriptomics, epigenomics, and single-cell RNA sequencing (scRNA-seq) to investigate the molecular and cellular mechanisms governing cortical neurogenesis and its pathological dysregulation in ETMRs.

The first part of this work explores the evolutionary mechanisms of cortical neurogenesis in amniotes by comparing transcriptomic and epigenomic profiles of apical radial glial cells (aRGCs) between mice, chicks, and snakes. We researched key regulators, such as *Sall1*, *miR-3607, Cux2, Robo1/2 or Dll1*, that govern neurogenic modes and cortical folding, shedding light on the conserved and divergent pathways that shaped cortical evolution.

In the second part, the existence of membrane-bound transcription factors (MTFs) in mammals is investigated. Hundreds of potential MTFs were identified, many localizing to the nucleus and suggesting roles in regulating gene expression during neurogenesis and cancer. Among these, *Robo1* was identified as a candidate MTF, predicted to possess transcriptional regulatory potential. These findings reveal a previously underexplored regulatory layer and its potential implications for cortical development and tumor biology.

The third part evaluates the *Rx-Cre-Dicer*$^{F/F}$ *(Rx-Dicer1)* mutant mouse as a preclinical model for ETMR and demonstrates that it is superior to the previously established GBS model. Transcriptomic profiling at embryonic stages E11.5 and E17.5 shows that *Rx-Dicer1* mutants not only recapitulate key histological and molecular features of ETMRs but also share a greater proportion of ETMR signature genes, such as *LIN28A/B*, *PRTG*, and *IGF2BPs*. Dysregulation of pathways like *PI3K/AKT/mTOR* further highlights its relevance. scRNA-seq identifies tumor-specific populations of neuroepithelial-like cells and their loss of aRGC identity. RNA velocity and transcriptional trajectory analyses reveal a disrupted differentiation process, with neuroepithelial-like cells becoming overproliferative and contributing to rosette formation.

Together, this research uncovers the evolutionary and pathological mechanisms driving cortical neurogenesis and its dysregulation in ETMRs. The findings establish the *Rx-Dicer1* mutant mouse as a more faithful model for ETMR than GBS and provide insights into potential therapeutic targets, such as *LIN28A/B, PRTG* and *PI3K/AKT/mTOR* signaling, for treating aggressive brain tumors. By bridging evolutionary biology and cancer research, this thesis advances our understanding of neurogenesis and its role in health and disease.

RESUMEN

La evolución de la neurogénesis cortical ha sido fundamental para el desarrollo del complejo cerebro de los mamíferos. Las alteraciones en estos procesos están en la base de varios trastornos del neurodesarrollo y cánceres, como los tumores embrionarios con rosetas multilaminares (ETMRs). Esta tesis integra transcriptómica comparativa, epigenómica y secuenciación de ARN de célula única (scRNA-seq) para investigar los mecanismos moleculares y celulares que regulan la neurogénesis cortical y su desregulación patológica en los ETMRs.

La primera parte de este trabajo explora los mecanismos evolutivos de la neurogénesis cortical en amniotas mediante la comparación de perfiles transcriptómicos y epigenómicos de las células gliales radiales apicales (aRGCs) entre ratones y pollos. Se investigaron reguladores clave, como *Sall1, miR-3607, Cux2, Robo1/2 y Dll1,* que controlan los modos neurogénicos y el plegamiento cortical, arrojando luz sobre las vías conservadas y divergentes que moldearon la evolución cortical.

En la segunda parte, se investiga la existencia de factores de transcripción asociados a membrana (MTFs) en mamíferos. Se identificaron cientos de MTFs potenciales, muchos de los cuales se localizan en el núcleo, lo que sugiere un papel en la regulación de la expresión génica durante la neurogénesis y el cáncer. Entre ellos, *Robo1* fue identificado como un candidato a MTF, predicho como potencial regulador transcripcional. Estos hallazgos revelan una capa reguladora previamente poco explorada y sus posibles implicaciones en el desarrollo cortical y la biología tumoral.

La tercera parte evalúa al ratón mutante *Rx-Cre-Dicer$^{F/F}$ (Rx-Dicer1)* como modelo preclínico para ETMR, demostrando que es superior al modelo GBS previamente establecido. El perfil transcriptómico en etapas embrionarias E11.5 y E17.5 muestra que los mutantes *Rx-Dicer1* no solo reproducen características histológicas y moleculares clave de los ETMRs, sino que también comparten una mayor proporción de genes característicos de ETMR, como *LIN28A/B, PRTG* e *IGF2BPs*. La desregulación de vías como *PI3K/AKT/mTOR* refuerza aún más su relevancia. La secuenciación de ARN de célula única identifica poblaciones tumorales específicas de células neuroepiteliales y su pérdida de identidad de aRGC. Los análisis de velocidad de ARN y trayectorias transcripcionales revelan un proceso de diferenciación alterado, con células neuroepiteliales que se vuelven hiperproliferativas y contribuyen a la formación de rosetas.

En conjunto, esta investigación descubre los mecanismos evolutivos y patológicos que impulsan la neurogénesis cortical y su desregulación en los ETMRs. Los hallazgos establecen al ratón mutante *Rx-Dicer1* como un modelo más fiel para ETMR en comparación con GBS y proporcionan información sobre posibles objetivos terapéuticos, como *LIN28A/B, PRTG* y la vía de señalización *PI3K/AKT/mTOR*, para tratar tumores cerebrales agresivos. Al conectar la biología evolutiva con la investigación del cáncer, esta tesis amplía nuestra comprensión de la neurogénesis y su papel en la salud y la enfermedad.

ACKNOWLEDGEMENTS

First of all, I wish to express my deepest gratitude to my thesis advisor, Víctor Borrell. Thank you for trusting me as a scientist and giving me the opportunity to carry out my PhD in your laboratory. Your guidance, your passion for doing science, and your invaluable advice have been essential to my development both as a researcher and as a person. Sharing this enthusiasm for science with you has been profoundly enriching. It not only rekindled my excitement for research, but also made me believe again in my own potential as a scientist.

To my parents, Lita and Rafa, I owe everything. Thank you for your unconditional love and for always being there at every step of the way. Your support and the values you instilled in me are the reason I stand here today. Every achievement I reach, every goal I meet, is yours as well. I love you with all my heart.

To my grandparents: thank you for being an endless source of pride and affection. Your love has been a fundamental pillar in my life. You have always believed in me, giving me strength even in the hardest moments. Some of you began this doctoral journey with me and, sadly, are no longer here to see its end, but I know you would be deeply proud. I feel you close, today and always.

My gratitude also goes to all the members of the laboratory—Lucia, Salma, Ana, Anna, Enrico, Edu, Laura, Sol, Irem, Alex, Lara, Ester, Pep, Pablo, Virginia, Adrián, and Bea. Thank you for the camaraderie, the shared laughter, and the mutual support. You have made day-to-day life in the lab not only productive but also full of joy. I know I have gained friends for life.

To Boyan Bonev, thank you for opening the doors of your lab to me and allowing me to learn new techniques and perspectives. Your hospitality, as well as Magdalena Götz's lab, made me feel welcomed and valued. Thank you for your generosity and willingness; you are wonderful researchers.

I also want to acknowledge my union colleagues and all the members of the doctoral student committee. Thank you for your commitment, your constant struggle, and for reminding me every day of the importance of unity and solidarity. You have shown me that, together, we can make a difference.

To my close friends—Luis, Álvaro, Laura, María, Josi, Javi, Héctor, David, Miranda, Iñaki, Nico, Sonia, Lorena, Pablo, Asier, Rosana… among many others—thank you for always being there. Whether with a call, a message of encouragement, or simply your presence, your support has been a necessary balm in difficult times.

Thank you to my cousin Pedro for being a constant in my life, for showing up at every important moment, and for offering your unwavering support in both joys and hardships. Your ability to listen, understand, and be there has been a fundamental pillar on my path. You are my companion in adventures, deep reflections, and inexhaustible laughter.

To Eva, thank you for sharing your life with me during these years. For your love, for your patience, and for being my refuge on tough days. To your family, thank you for welcoming me so warmly and making me feel like one of you. Reaching the end of this PhD with happiness and strength to keep moving forward is, in large part, thanks to you.

Thank you to all my relatives and friends for walking beside me on this journey. Every word of encouragement, every hug, and every small gesture of affection has reminded me that I am not alone. Everything I have achieved—and everything yet to come—is partly thanks to you.

Finally, I want to thank myself. For never giving up on the dream of being a researcher, even when the challenges seemed insurmountable. For facing the moments of uncertainty, the personal sacrifices, the long hours of work, and the weight of doubts that at times shook me. Despite everything, I kept going, driven by passion and the conviction that it was worth fighting for what I truly love. This path has not been easy, but looking back, I am proud of every obstacle overcome, every step taken, and of the person I have become. I have learned to trust my capacity to grow, to persist, and to dream, even when the road looked dark. This achievement is not only academic; it is proof of my strength, dedication, and resilience. Thank you, to myself, for believing in me even in the hardest moments.

AGRADECIMIENTOS

En primer lugar, quiero expresar mi más profundo agradecimiento a mi director de tesis, Víctor Borrell. Gracias por confiar en mí como científico y brindarme la oportunidad de realizar mi doctorado en tu laboratorio. Tu guía constante, tu motivación por hacer ciencia, y tus valiosos consejos, han sido esenciales para mi formación como investigador y como persona. Compartir contigo esta pasión por la ciencia ha sido una experiencia profundamente enriquecedora, que no solo me devolvió la ilusión por la investigación, sino también me hizo volver a creer en mi propio potencial como científico.

A mis padres, Lita y Rafa, os debo todo. Gracias por vuestro amor incondicional y por estar siempre ahí, en cada paso de este camino. Vuestro apoyo y los valores que me habéis inculcado son el motivo por el que hoy estoy aquí. Cada logro que consigo, cada meta alcanzada, es también vuestra. Os quiero con todo mi corazón.

A mis abuelos, gracias por ser esa fuente inagotable de orgullo y amor. Vuestro cariño ha sido un pilar esencial en mi vida. Siempre habéis creído en mí, y eso me ha dado fuerza incluso en los momentos más difíciles. Aunque algunos de vosotros empezasteis este camino del doctorado conmigo y, tristemente, ya no estáis para ver su final, estoy seguro de que estaríais profundamente orgullosos. Os siento cerca, hoy y siempre.

Mi gratitud también va dirigida a todos los miembros del laboratorio. Lucia, Salma, Ana, Anna, Enrico, Edu, Laura, Sol, Irem, Alex, Lara, Ester, Pep, Pablo, Virginia, Adrián y Bea, gracias por el compañerismo, las risas compartidas y el apoyo mutuo. Habéis hecho que el día a día en el laboratorio sea no solo productivo, sino también lleno de felicidad. Sé que he hecho amigos que llevaré conmigo toda la vida.

A Boyan Bonev, gracias por abrirme las puertas de tu laboratorio, por permitirme aprender nuevas técnicas y perspectivas. Tu hospitalidad y la de todo tu equipo, así como la del laboratorio de Magdalena Götz, me hicieron sentir acogido y valorado. Gracias por vuestra generosidad y disposición, sois investigadores maravillosos.

Quiero reconocer también a mis compañeros sindicales y a todos los integrantes del comité de estudiantes de doctorado. Gracias por vuestro compromiso, vuestra lucha constante y por recordarme, cada día, la importancia de la unión y la solidaridad. Habéis sido un recordatorio de que, juntos, podemos marcar la diferencia.

A mis amigos cercanos, Luis, Álvaro, Laura, María, Josi, Javi, Héctor, David, Miranda, Iñaki, Nico, Sonia, Lorena, Pablo, Asier, Rosana… entre muchos otros. Gracias por estar ahí, siempre. Ya sea con una llamada, un mensaje de aliento o simplemente con vuestra presencia, vuestro apoyo ha sido ese bálsamo necesario en los momentos difíciles.

Gracias a mi primo Pedro por ser esa constante en mi vida, por estar presente en cada momento importante, y por ofrecerme tu apoyo incondicional en las alegrías y en las dificultades. Tu capacidad para escucharme, comprenderme y estar ahí, ha sido un pilar fundamental en mi camino. Eres ese compañero de aventuras, de reflexiones profundas y de risas que nunca faltan.

A Eva, gracias por compartir tu vida conmigo durante estos años. Gracias por tu amor, por tu paciencia y por ser mi refugio en los días complicados. A tu familia, gracias por acogerme con tanto cariño y por hacerme sentir parte de vosotros. Llegar al final de este doctorado con felicidad y fuerzas para seguir adelante es, en gran parte, gracias a vosotros.

Gracias a todos mis familiares y amigos, gracias por caminar conmigo en este viaje. Cada palabra de ánimo, cada abrazo y cada pequeño gesto de cariño me han recordado que no estoy solo. Todo lo que he conseguido y lo que vendrá es, en parte, gracias a vosotros.

Finalmente, quiero darme las gracias a mí mismo. Por no haberme rendido nunca en el sueño de ser investigador, incluso cuando las dificultades parecían insuperables. Por haber enfrentado los momentos de incertidumbre, los sacrificios personales, las largas horas de trabajo y el peso de las dudas que en ocasiones me hicieron tambalear. A pesar de todo, he seguido adelante, impulsado por la pasión y la convicción de que valía la pena luchar por lo que realmente amo. Este recorrido no ha sido fácil, pero al mirar atrás, me siento orgulloso de cada obstáculo superado, de cada paso dado y de la persona en la que me he convertido. He aprendido a confiar en mi capacidad para crecer, para persistir y para soñar, incluso cuando el camino parecía oscuro. Este logro no solo es académico; es una prueba de mi fuerza, mi dedicación y mi resiliencia. Gracias, a mí mismo, por haber creído en mí, incluso en los momentos más difíciles.

INTRODUCTION

1. Cerebral cortex development

1.1 Mammals

The brain is the most complex part of the human body, being the seat of intelligence, interpreter of the senses, initiator of body movement, and controller of behavior. Lying in its bony shell and washed by protective fluid, the brain is the source of all the qualities that define our humanity.

Coating the surface of the cerebrum and the cerebellum is a vital layer of tissue about the thickness of a stack of two or three dimes, called the cerebral cortex. Most of the actual information processing in the brain takes place in the cerebral cortex, being responsible for many of the higher-level cognitive functions in humans, including language, perception, decision making, and motor planning.

In mammals, the cerebral cortex can be anatomically divided into neocortex (NCx) and allocortex, being the NCx the most developed in its organization and number of layers of the cerebral tissues (Rakic, 2009; Borrell, 2019; Moini, Gutierrez and Avgeropoulos, 2023).

1.1.1 Histological development and progenitor cell types

The onset of cerebral cortex development involves the formation of the telencephalic vesicles (Figure 1A). These areas are initially specified as the future locations of the cortex, consisting only of a single layer of stem cells on the inner surface, called neuroepithelial stem cells (NECs) (Bishop, Goudreau and O'Leary, 2000). These specialized cells are elongated and have two thin extensions reaching out from their main body in opposite directions. One extension goes up towards the surface of the neuroepithelium (apical process) and the other goes down towards the basal membrane (basal process) (Sidman and Rakic, 1973). Neighboring NECs are connected at the apical surface through adherent and tight junctions (Figure 1B) (Taverna, Götz and Huttner, 2014). NECs go through a self-amplifying type of cell division, creating two new NECs from each original cell. Throughout this cell cycle, the nucleus of the NEC moves up and down along the cell's long axis. This movement is called Interkinetic Nuclear Migration (INM) (Figure 1D) (Takahashi, Nowakowski and Caviness, 1993). Cell division itself happens at the apical surface. During the G1 phase, the nucleus moves towards the basal side, and it stays there during the S phase synthesizing DNA. Then, in the G2 phase, it moves back to the apical surface where cell division takes place again (Figure 1D) (Hinds and Ruffett, 1971; Sidman and Rakic, 1973; Caviness, Takahashi and Nowakowski, 1995; Takahashi, Nowakowski and Caviness, 1995). Importantly, not all NECs perform INM at the same time. This asynchronous timing is what gives the neuroepithelium its characteristic appearance of having many layers just one (pseudostratified). INM is a feature found across evolution in different species, although the details of the movement, cell density, and the thickness of the NEC layer may vary (Miyata *et al.*, 2014).

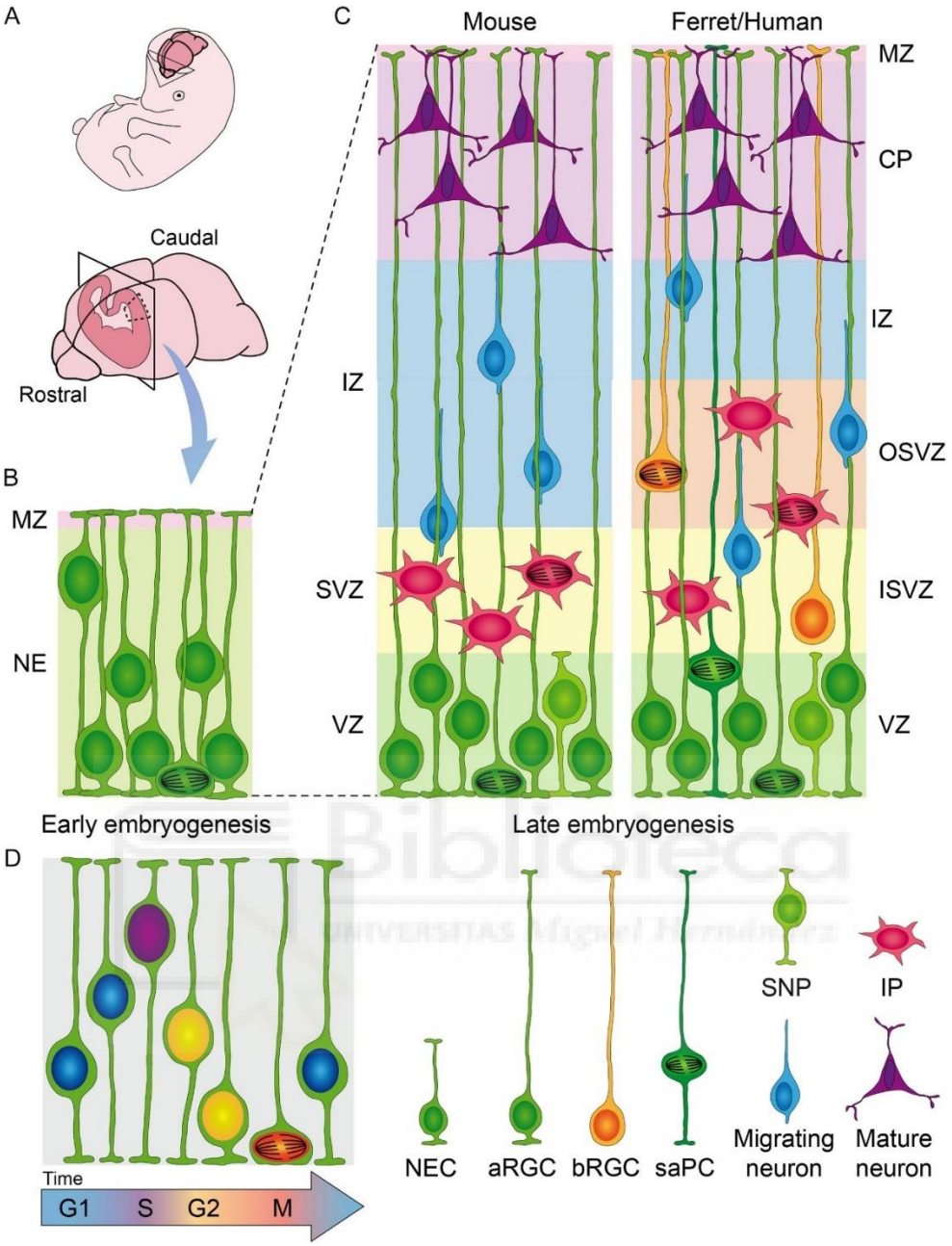


Figure 1. Development of the Mammalian Cortex and Cellular Components. A) Diagram showing a mammalian embryo and the forebrain, including a cross-sectional schematic of the telencephalon (highlighted in dark pink). **B), C)** Overview of the cortical primordium structure and its cell types during early **B)** and late **C)** developmental stages. The pseudostratified neuroepithelium (NE) initially consists of proliferative neuroepithelial cells (NECs). At later stages, distinct germinal zones (VZ, SVZ) and post-mitotic layers (IZ, CP, MZ) emerge, with more specialized organization in the folded ferret and human cortex compared to the smooth mouse cortex. While the primary cellular components are conserved across species, basal progenitors are significantly more abundant in ferrets and humans. **D)** Illustration of Interkinetic Nuclear Migration, synchronized with the phases of the cell cycle. Key cellular components include: aRGC (apical Radial Glia Cell), bRGC (basal Radial Glia Cell), saPC (subapical Progenitor Cell), SNP (Short Neural Precursor), IP (Intermediate Progenitor), VZ (Ventricular Zone), SVZ (Outer and Inner Subventricular Zone), IZ (Intermediate Zone), CP (Cortical Plate), and MZ (Marginal Zone). **Reprinted from** *Current Topics in Developmental Biology*, Volume 142, 2021, Pages 1–66, Chapter One - *The Regulation of Cortical Neurogenesis*, Ana Villalba, Magdalena Götz, and Víctor Borrell, Copyright © 2021, with permission from Elsevier (License Number: 5966970248779).

Following some cell divisions and self-amplification, NECs loosen their tight junctions but maintain apical adherent junctions, transforming into apical Radial Glial Cells (aRGCs) (Figure 1C) (Taverna, Götz and Huttner, 2014). aRGCs are the main type of cortical progenitor cell, known to divide in the apical surface. They express Paired-box transcription factor 6 (*Pax6*), and their progeny produce all the excitatory neurons in the cortex (Malatesta *et al.*, 2003). The emergence of aRGCs marks the beginning of cortical neurogenesis. Like NECs, aRGCs have processes reaching apical and basal sides of the cortical primordium (Rakic, 1972; Sidman and Rakic, 1973). They also continue the INM, but it's now limited to the apical domain, close to the ventricle surface (Takahashi, Nowakowski and Caviness, 1993; Noctor *et al.*, 2001; Haubensak *et al.*, 2004; Attardo *et al.*, 2008). This area becomes the first layer of progenitor cells in the developing cortex: the Ventricular Zone (VZ) (Figure 1C) (Angevine Jr. *et al.*, 1970). The VZ contains the cell bodies of both aRGCs and Short Neural Precursors (SNPs). SNPs are another, less common type of progenitor cell that lack either the basal process entirely or have a short one confined to the VZ (Stancik *et al.*, 2010). In the developing mouse cortex, SNPs primarily create new neurons and have a limited ability to self-renew. However, in other brain regions like the ventral telencephalon, they can divide multiple times, acting as amplifiers for cell production (Pilz *et al.*, 2013).

In the early stages of brain development, around day 11 of mouse embryonic development (E11), most aRGCs undergo self-replication, while a small portion differentiate into mature neurons (Figure 1B). As development progresses, this self-replication by aRGCs slows down, while there's an increase in asymmetric cell divisions that generate new types of secondary progenitor cells and neurons (Figure 1C) (Takahashi, Nowakowski and Caviness, 1995, 1996; Takahashi, Nowakowski and Caviness Jr., 1996; Miyata *et al.*, 2004). These new progenitors migrate from their birthplace near the apical surface to the basal border of the VZ where they divide. These characteristic progenitors are called Basal Progenitors (BPs) (Figure 1C) (Miyata *et al.*, 2004; Noctor *et al.*, 2004; Noctor, Martínez-Cerdeño and Kriegstein, 2008; Taverna, Götz and Huttner, 2014). As BPs accumulate, they form a secondary zone of cell proliferation called the Subventricular Zone (SVZ) (Angevine Jr. *et al.*, 1970; Takahashi, Nowakowski and Caviness, 1995; Noctor *et al.*, 2004; Noctor, Martínez-Cerdeño and Kriegstein, 2008). In mice, most of these BPs are Intermediate Progenitor Cells (IPCs) (Figure 1C). These are small, round cells with no polarity and short processes. They are also identified by the presence of a specific protein called T-box transcription factor 2 (*Tbr2*) (Englund *et al.*, 2005). IPCs are responsible for generating most of the excitatory neurons in all layers of the mouse cerebral cortex (Haubensak *et al.*, 2004; Noctor *et al.*, 2004; Kowalczyk *et al.*, 2009; Vasistha *et al.*, 2015). Later in development, aRGCs can also produce two other types of progenitor cells that divide in locations away from the apical surface. Subapical Progenitor Cells (saPCs) resemble aRGCs with an apical process, but they divide in between the VZ. These saPCs are rare in species with a small and smooth cortex, like mice and marmoset monkeys, but abundant in species with a larger and folded cortex, like ferrets and sheep (Figure 1C) (Pilz *et al.*, 2013). This

suggests a potential role for saPCs in cortical expansion and folding (Borrell and Götz, 2014).

The other kind of BP is known as basal Radial Glial Cell (bRGC) or outer Radial Glial (oRG) (Figure 1C) (LaMonica *et al.*, 2013). Much like aRGCs, bRGCs express *Pax6* and have a basal process that extends radially and anchors to the basal membrane (Fietz *et al.*, 2010; Hansen *et al.*, 2010; Reillo *et al.*, 2011; Wang *et al.*, 2011). However, unlike aRGCs, bRGCs lack a process reaching the apical surface and don't divide apically. Instead, their cell division happens basally within the SVZ. While bRGCs typically don't undergo INM, they do exhibit mitotic somal translocation (a rapid movement of the cell body just before division) in the developing human cortex, and in non-human primates (Hansen *et al.*, 2010; Betizeau *et al.*, 2013; LaMonica *et al.*, 2013). In the developing mouse cortex, which is small and smooth, bRGCs are rare and primarily neurogenic rather than self-renewing. They express *Pax6* but not *Eomes* (Shitamukai, Konno and Matsuzaki, 2011; Wang *et al.*, 2011). In contrast, species with large and folded brains like primates, carnivores, and ungulates have a high abundance of bRGCs with a high capacity for self-amplification, and a significant portion co-express *Pax6* and *Eomes* (Reillo *et al.*, 2011; Betizeau *et al.*, 2013; Martínez-Martínez *et al.*, 2016).

1.1.2 Modes of neurogenesis

Neurons can be generated through two primary modes: Direct Neurogenesis and Indirect Neurogenesis. In direct neurogenesis, aRGCs in the VZ directly produce neurons (Figure 2A, Figure 2B (abundant)). In contrast, indirect neurogenesis involves an intermediary step where aRGCs generate BPs, which then divide to produce neurons (Figure 2B (sparse), Figure 2C, Figure 2D). This intermediary process amplifies neuron production significantly compared to direct neurogenesis (Laguesse *et al.*, 2015; Cárdenas *et al.*, 2018). Due to the amplifying effect of BPs, indirect neurogenesis enhances neuron production manifold compared to direct neurogenesis (Hevner *et al.*, 2006; Kriegstein, Noctor and Martínez-Cerdeño, 2006). Indirect neurogenesis is the primary mode of cortical neuron generation in mammals, diverging from reptiles, and is notably augmented in primates and other species featuring a significantly expanded cortex (Figure 2C, Figure 2D) (Haubensak *et al.*, 2004; Dehay and Kennedy, 2007; Reillo *et al.*, 2011; Cárdenas *et al.*, 2018).

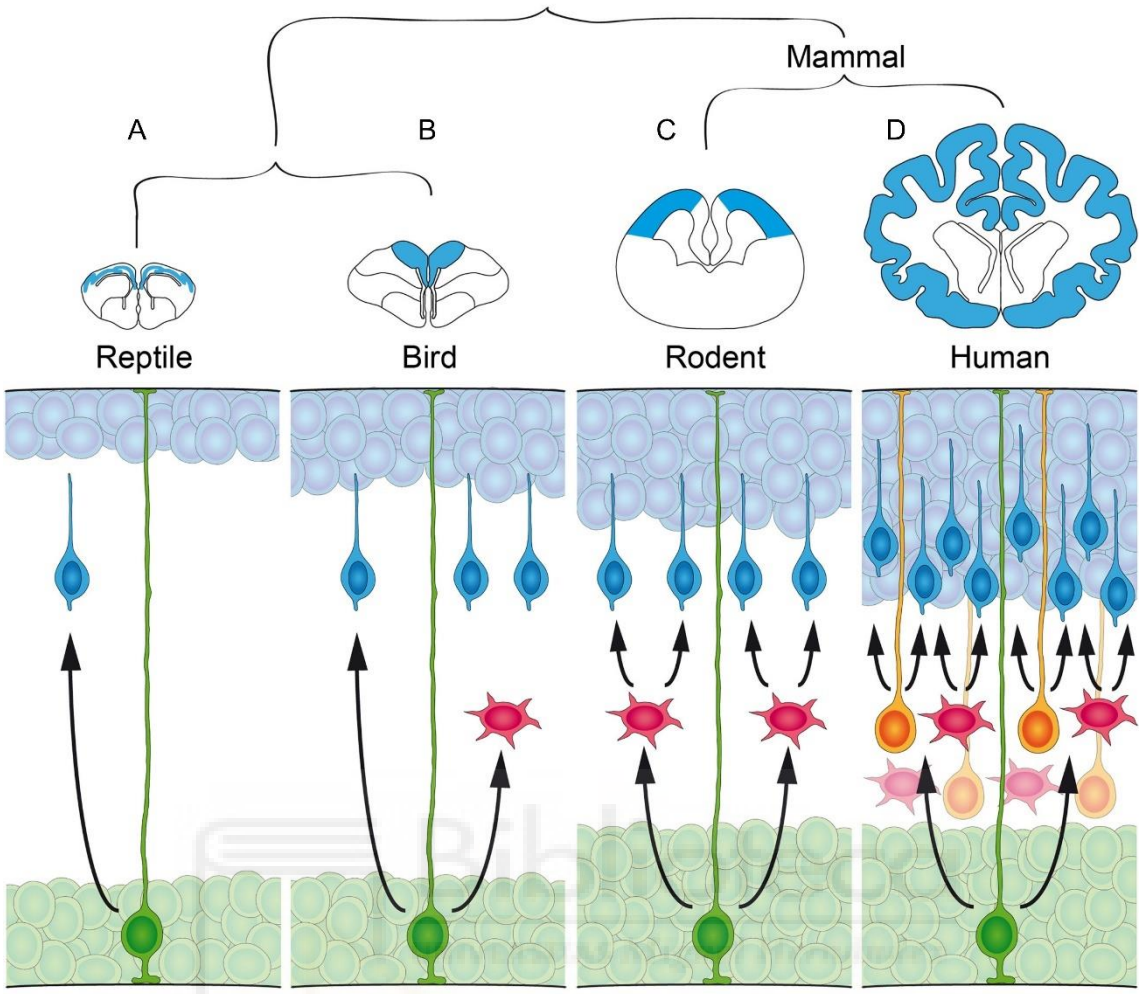


Figure 2. Evolution of Cortical Neurogenesis. Evolutionary differences in cortical neurogenesis among reptiles, birds, and mammals (rodents and humans). In reptiles **A**) and birds **B**), most cortical neurons are produced through direct neurogenesis by apical radial glial cells (aRGCs, green). In mammals **C**), **D**), cortical neurogenesis is dominated by indirect neurogenesis, which involves basal progenitors, including intermediate progenitor cells (IPCs, red) and basal radial glial cells (bRGCs, orange). This shift facilitates a significant increase in neuron production, enabling the expansion and complexity of the mammalian cortex, particularly in humans **D**), where diverse projection neuron types arise, contributing to advanced cortical functions. **Adapted from** *Current Topics in Developmental Biology*, Volume 142, 2021, Pages 1–66, Chapter One - *The Regulation of Cortical Neurogenesis*, Ana Villalba, Magdalena Götz, and Víctor Borrell, Copyright © 2021, with permission from Elsevier (License Number: 5966970248779).

Since the identification of aRGCs and BPs as the primary sources of cortical neurons (Götz, Stoykova and Gruss, 1998; Malatesta, Hartfuss and Götz, 2000; Miyata *et al.*, 2001; Noctor *et al.*, 2001), considerable efforts have been dedicated to discerning the precise roles of each neurogenic mode in cerebral cortex maturation. Kowalczyk and collaborators corroborated the presence of dividing IPCs in the emerging mouse NCx right from the outset of neurogenesis. They observed Tbr2+ dividing progenitors within the SVZ as early as E10.5, with a progressive increase until E16.5. Notably, the zenith of Tbr2+ dividing progenitors occurs between E12.5 and E13.5 in the mouse NCx. Critically, the proportion of

neurogenic divisions attributable to IPCs (indirect neurogenesis) consistently surpassed that of aRGCs (direct neurogenesis) (Kowalczyk *et al.*, 2009).

In our laboratory, we also explored the extent of direct and indirect neurogenesis in the mouse rostral NCx, proximal to the olfactory bulb, during early developmental stages (E12.5) (Cárdenas *et al.*, 2018). Our findings robustly demonstrate that direct neurogenesis accounts for less than 5% of total apical mitoses in the developing NCx during early developmental phases. Conversely, parallel analyses in the olfactory bulb at the same developmental stage revealed markedly elevated levels of direct neurogenesis (20%), highlighting a contrasting distribution of neurogenic modes between the NCx ("new cortex" in evolutionary terms) and the olfactory bulb ("old cortex" in evolutionary terms), being the NCx the region exhibiting higher levels of indirect neurogenesis and BP cell mitoses between E12 and E13 (Cárdenas *et al.*, 2018).

Nevertheless, several studies from diverse authors appear contradictory regarding the frequency of direct versus indirect neurogenesis during mouse neocortical development (Haubensak *et al.*, 2004; Miyata *et al.*, 2004; Attardo *et al.*, 2008; Noctor, Martínez-Cerdeño and Kriegstein, 2008; Kowalczyk *et al.*, 2009; Mihalas *et al.*, 2016; Cárdenas *et al.*, 2018; Mihalas and Hevner, 2018), indicating that much remains to be elucidated concerning the fate of progenitor cells and the dynamics of neurogenic modes.

Recent findings indicate that both direct and indirect neurogenesis play roles in the development of all cortical structures. However, indirect neurogenesis predominantly contributes to the hippocampus and NCx formation. In the NCx, while direct neurogenesis is responsible for producing all primary glutamatergic projection neuron types, indirect neurogenesis enhances and diversifies projection neurons within each type, creating intricate mosaics of lineage-based cortical subnetworks (Figure 2D). In conclusion, although the precise ratio of direct and indirect neurogenesis during cortical development remains undetermined, it appears that indirect neurogenesis augments the diversity of neurons (Huilgol *et al.*, 2023).

However, the genetic and molecular mechanisms that guide aRGCs in deciding between direct or indirect neurogenesis between mouse brain regions are still unknown and yet to be discovered. Understanding these mechanisms is crucial for elucidating the regulation of neurogenesis and the development of cortical structures.

1.2. Non-mammalian amniotes

1.2.1 Histological formation and progenitor cell types

Similar to mammals, the embryonic dorsal pallium of reptiles and birds consists mainly of the VZ, where aRGCs divide to give rise to daughter stem cells, neurons, and glial cells. Indeed, aRGCs are conserved across all vertebrates, indicating their early evolutionary emergence (Wullimann, Puelles and Wicht, 1999; Englund *et al.*, 2005; Götz and Huttner, 2005; Kriegstein, Noctor and

Martínez-Cerdeño, 2006; Nomura, Gotoh and Ono, 2013; Taverna, Götz and Huttner, 2014; Docampo-Seara *et al.*, 2018; García-Moreno *et al.*, 2018; Nomura *et al.*, 2018; Cárdenas and Borrell, 2020; Espinós *et al.*, 2022). In contrast, IPCs and the SVZ exhibit greater variation across species (Martínez-Cerdeño, Noctor and Kriegstein, 2006; Cheung *et al.*, 2010; Molnár, 2011; De Juan Romero and Borrell, 2015; Cárdenas *et al.*, 2018).

Lepidosaurs, such as snakes and lizards, exhibit a lack of basal mitoses in their developing dorsal pallium (Nomura, Gotoh and Ono, 2013; Martínez-Cerdeño *et al.*, 2016; Cárdenas *et al.*, 2018). In lepidosaur embryos, cells expressing *Eomes* are found within or basal to the VZ, with an absence of progenitor markers and prevalence of neuronal markers, indicating the absence of an SVZ during telencephalon embryonic development (Nomura, Gotoh and Ono, 2013; Martínez-Cerdeño *et al.*, 2016). The adult reptile dorsal cortex originates from this VZ. The reptile dorsal cortex displays a relatively simple structure comprising three layers, with two sparsely populated plexiform layers (inner and outer) flanking a densely packed neuronal layer (Nomura *et al.*, 2013). This simplistic organization is reminiscent of evolutionarily ancestral brain regions in mammals like the hippocampus and piriform cortex, suggesting its existence in the common ancestor of exhaust reptiles and mammals (Dugas-Ford and Ragsdale, 2015; Luzzati, 2015; Cárdenas and Borrell, 2020).

In Galliformes (chickens and related species), a distinct SVZ is present in the dorsal ventricular ridge (DVR), subpallium, and hyperpallium during late neurogenesis stages, confirming its existence with IPCs, as evidenced by the evolutionary conservation of *Eomes* as a marker of the glutamatergic cell lineage (Englund *et al.*, 2005; Cheung *et al.*, 2007; Charvet, Owerkowicz and Striedter, 2009; De Juan Romero and Borrell, 2015; Martínez-Cerdeño *et al.*, 2016; Nomura *et al.*, 2016; Cárdenas and Borrell, 2020). However, only a few basal mitoses occur in the lateral dorsal pallium, decreasing in a lateral-to-medial gradient (Cheung *et al.*, 2007; Charvet, Owerkowicz and Striedter, 2009; Martínez-Cerdeño *et al.*, 2016).

In the adult hyperpallium of birds, unlike the laminar organization seen in mammals and reptiles, projection neurons exhibit a distinctive nuclear-like or semi-layered arrangement (Jarvis *et al.*, 2005; Montiel *et al.*, 2016). Despite the absence of layers in the avian pallium, excitatory neurons establish connections in a manner strikingly like the mammalian NCx (Briscoe and Ragsdale, 2018). Hence, while evolution appears to have conserved the basic layout of circuits, cell types, and embryological territories of the amniote pallium, it has not preserved a specific cytoarchitecture (Briscoe and Ragsdale, 2018; Martínez-Cerdeño *et al.*, 2018; Tosches *et al.*, 2018).

1.2.2 Modes of neurogenesis

IPCs (or any BP type) are notably absent in the reptilian pallium (Charvet, Owerkowicz and Striedter, 2009; Nomura, Gotoh and Ono, 2013; Martínez-Cerdeño *et al.*, 2016), requiring that all neurons be generated by direct

neurogenesis from aRGCs (Cárdenas and Borrell, 2020). In fact, the limited productivity of this neurogenic mode is further reduced by the extended length of cell cycles of reptilian cortical progenitors, resulting in a relatively small number of total cell cycles during their brief neurogenic period (Nomura, Gotoh and Ono, 2013). This combination leads to a minimal production of cortical neurons, resulting in the formation of the dorsal cortex, which is relatively small and thin, organized into three layers (Figure 2A).

In birds, specifically in chicks, there is a rudimentary SVZ containing a relatively small number of basal mitoses (Charvet, Owerkowicz and Striedter, 2009; Cárdenas and Borrell, 2020), some of which are *Tbr2*-positive, typical of IPCs (Cheung *et al.*, 2007; Charvet, Owerkowicz and Striedter, 2009; Martínez-Cerdeño *et al.*, 2016; Cárdenas *et al.*, 2018). This indicates that in the embryonic avian pallium, there is a partial restriction of the direct neurogenic program along with a partial activation of the indirect neurogenesis program. Consequently, this leads to a modest occurrence of BP, ultimately contributing to the expansion of the avian pallium compared to reptiles (Figure 2B).

Nevertheless, the genetic and molecular mechanisms that guide aRGCs in deciding between direct or indirect neurogenesis across species—mammalian, avian, and reptilian—remain unknown and are a subject of ongoing research. Understanding these mechanisms is crucial for elucidating the regulation of neurogenesis and the development of cortical structures in different vertebrate lineages.

2. Disease of cortical development: pediatric brain tumors

Pediatric brain tumors are the most common type of solid childhood cancer and only second to leukemia as a cause of pediatric malignancies. They are classified into supra and infratentorial tumors. They are also classified according to the age of diagnosis into congenital brain tumors (CBT) (diagnosed antenatally in the first 60 days of life), tumors of the infancy (younger than 1 year of age), and older children. With the advances in imaging techniques, molecular biology, and genetics, pediatric brain tumors are increasingly being diagnosed early in the disease course, sub-grouped, and treated with more targeted strategies (Ostrom *et al.*, 2015, pp. 2007–2011; Subramanian and Ahmad, 2024).

Historically, tumors of the central nervous system (CNS) were categorized based on their presumed cell of origin and developmental differentiation states. This foundational principle still influences several tumor names included in the current 5th edition of the World Health Organization (WHO) Classification (Louis *et al.*, 2021). For instance, astrocytomas and oligodendrogliomas—once believed to originate from neoplastic astrocytes and oligodendrocytes, respectively—are prime examples of this traditional nomenclature. However, recent research has revealed a far more complex scenario. Experimental evidence demonstrates that astrocytomas can arise from various cell types, including oligodendrocyte

precursor cells, neural precursor cells, and astrocytes themselves (Liu *et al.*, 2011; Lindberg *et al.*, 2014; Tirosh *et al.*, 2016; Venteicher *et al.*, 2017; Viaene, 2023).

In the past, the histologic classification of CNS tumors primarily relied on the examination of hematoxylin and eosin (H&E)-stained sections under a light microscope. While this method remains a cornerstone of diagnostic practice, additional ancillary studies have become instrumental in tumor classification and grading, which estimate a tumor's biological behavior. CNS tumors are assigned grades from 1 to 4, with grade 1 representing low-grade tumors and grade 4 indicative of malignancy. Although grading criteria differ across tumor types, certain features, such as increased mitotic activity and focal necrosis, are generally associated with higher grades. While histological evaluation continues to be the primary method for grading many CNS tumors, some, like diffuse midline glioma with H3K27 alteration, are graded based on molecular and transcriptomic profiles. Regardless of histological findings, this tumor is assigned a CNS WHO grade 4 due to the poor prognosis linked to its molecular characteristics (Mackay *et al.*, 2017).

The integration of molecular and transcriptomic data with histological findings to establish an "integrated diagnosis" was first introduced by the International Society of Neuropathology (ISN)-Haarlem guidelines in 2014 (Louis *et al.*, 2014) and officially incorporated into the revised 4th edition of the WHO CNS tumor classification (Louis *et al.*, 2016). This approach has been significantly expanded in the 5th edition of the WHO classification (Louis *et al.*, 2021). Within this framework and focusing on the primary pediatric brain tumors that will be discussed in this thesis, they can be categorized into the following groups:

2.1. Gliomas

2.1.1 Low-grade gliomas (LGG)

Low-grade gliomas are the most frequent brain tumors of childhood, accounting for a third of all cases if mixed glioneuronal and neuronal tumors are included. This group of tumors is heterogeneous; unlike low-grade gliomas in adults, those in children rarely transform into higher-grade tumors. Common IDH1 or IDH2 mutations seen in adult gliomas are much less common in pediatric cases. The initial treatment for most low-grade gliomas in children is surgery to establish a tissue diagnosis and achieve maximal safe resection. In a large international study, the 5-year progression-free survival for children with low-grade gliomas was 69%, and overall survival was 95%. Risk factors for progression include young age, incomplete resection, specific histological features, and certain tumor locations (Ryall, Tabori and Hawkins, 2020).

2.1.2 High-grade gliomas (HGG)

Pediatric-type high-grade gliomas, accounting for 10% of childhood brain tumors, have a poor prognosis, with 70-90% of affected children dying within 2 years despite surgery and adjuvant therapy. The WHO CNS5 reclassified "glioblastoma multiforme" with an emphasis on molecular markers, removing the term "glioblastoma" from pediatric neoplasms (Louis *et al.*, 2021). Advances in understanding include identifying driver mutations in the chromatin-remodeling gene family, with four subtypes of gliomas recognized:

Diffuse Midline Glioma (H3K27-altered): A lethal tumor in young children, previously termed “H3K27M-mutant,” affecting the brain's midline structures and associated with worse survival. These tumors include the previously named diffuse intrinsic pontine gliomas, along with aggressive gliomas involving the thalamus and other midline structures.

Diffuse Hemispheric Glioma (H3G34-mutant): Occurs in older children and young adults, linked with *ATRX* and *TP53* mutations and MGMT promoter methylation.

High-Grade Glioma (IDH mutant): Less common in children compared to adults, these tumors have mutations in the IDH gene and may have a different prognosis and response to therapy than IDH wild-type gliomas.

Diffuse Pediatric-Type High-Grade Glioma (H3 and IDH wild-type): An aggressive tumor typically found in the cerebral hemispheres with a poor prognosis.

Infant-Type Hemispheric Glioma: Found in newborns and infants, often harboring receptor tyrosine kinase gene fusions, which are targetable and suggest better outcomes.

2.2 CNS embryonal tumors

2.2.1 Medulloblastomas

Medulloblastomas are the most common malignant brain tumors in children, typically arising in the cerebellum and presenting symptoms of increased intracranial pressure or cerebellar dysfunction. They account for over 60% of childhood embryonal tumors, with 70% occurring in children under 10 years old, predominantly in boys. One-third of cases occur in children under 3 years old. Poor prognostic factors include large tumor size, disseminated disease at presentation, young age (<3 years), and significant residual tumor post-surgery.

Historically, medulloblastomas were classified morphologically into four subtypes: classic, large-cell anaplastic, desmoplastic-nodular, and medulloblastoma with extensive nodularity, with the latter two having better prognoses. The current CNS5 classification system (Louis *et al.*, 2021) divides medulloblastomas into molecularly defined and histologically defined categories, with the molecularly

defined group further split into four subtypes, each with distinct clinical behaviors and genetic profiles:

***WNT*-Activated Medulloblastomas** account for 10% of cases and have an excellent prognosis, with a 10-year event-free survival rate exceeding 95%. These tumors often have β-catenin accumulation due to *CTNNB1* mutations.

***SHH*-Activated Medulloblastomas** make up 30% of cases, affect young children and adults, and are linked to *SHH–PTCH–SMO–GLI* pathway mutations. Prognosis varies based on *TP53* mutation status, with *TP53*-mutated tumors having a poor outcome. Targeted therapies like *SMO* inhibitors are being trialed for these tumors.

Non-*WNT*, Non-*SHH* Medulloblastomas include group 3 and group 4 subtypes. Group 3 tumors, which constitute 25% of cases, have the poorest prognosis and often feature isochromosome 17q. Group 4 tumors, the most common subtype at 35%, have an intermediate prognosis, with genetic alterations such as *MYCN* and *CDK6* amplifications.

2.2.2 Other CNS embryonal tumors

2.2.2.1 Atypical teratoid rhabdoid tumors (ATRT)

Atypical teratoid rhabdoid tumor (ATRT) is a CNS cancer in young children, characterized by multi-lineage differentiation and a primitive phenotype. In the 1990s, studies revealed the loss of chromosome 22's long arm as a common event in rhabdoid tumors, including ATRT, and identified the loss of SMARCB1, a key component of the SWI/SNF chromatin-remodeling complex, as the primary genetic alteration in most ATRTs (Versteege *et al.*, 1998; Sévenet *et al.*, 1999; Biegel *et al.*, 2002). Despite this genetic simplicity, recent epigenetic research has uncovered shared and subtype-specific epigenetic disruptions that contribute to tumorigenesis, prompting efforts to develop new therapies targeting these changes (Torchia *et al.*, 2015, 2016; Johann *et al.*, 2016; Ho *et al.*, 2020; Federico *et al.*, 2022).

ATRT accounts for 1-2% of pediatric CNS tumors but is more common in early childhood, with three-quarters of cases occurring in children under three years old. ATRT represents about 20% of embryonal CNS tumors in this age group and 40-50% of CNS malignancies in the first year of life. Although rare in teens and adults, the median age at diagnosis is 16-30 months, with a male predominance (male-to-female ratio of 1.1 to 2). ATRTs can develop anywhere in the CNS and should be considered in aggressive intracranial tumors in young children. These tumors typically appear as large, heterogeneous masses with necrosis, hemorrhage, and peritumoral edema. They often show restricted diffusion on MRI due to dense cellularity, and a distinctive thick, irregularly enhancing wall surrounding a central cystic region may be present in up to 28% of cases (Nesvick *et al.*, 2020).

Specific MRI patterns are associated with ATRT subtypes: ATRT-MYC tumors tend to have more peritumoral edema, and ATRT-SHH tumors may show less enhancement. Multiple intracranial or extracranial lesions, especially in very young children, suggest rhabdoid tumor predisposition syndrome (RTPS), often linked to *SMARCB1* or *SMARCA4* mutations. Up to 35% of ATRT patients may have *SMARCB1* germline mutations, with higher rates in children under one year old. In fact, modern diagnosis relies on detecting the absence of *SMARCB1* expression, confirming ATRT in the appropriate clinical and histopathologic context. Some ATRTs retain *SMARCB1* expression but have *SMARCA4* mutations instead. Additionally, *DICER1* alteration syndrome can also give rise to ATRTs (Biegel, Busse and Weissman, 2014; Nesvick *et al.*, 2020; Reddy *et al.*, 2020). *DICER1* is involved in microRNA processing, and its mutation can contribute to tumorigenesis in various tissues, including the CNS (Kamihara *et al.*, 2020).

2.2.2.2 High-grade neuroepithelial tumor (HGNET)

CNS high-grade neuroepithelial tumor with *BCOR* alteration (CNS HGNET-BCOR) is a newly identified entity, characterized by internal tandem duplication (ITD) in exon 15 of the *BCOR* gene, which acts as a corepressor of *BCL6*, a gene crucial for immune responses. These tumors resemble high-grade gliomas histologically and exhibit BCOR immunopositivity. However, *BCOR* fusions are also found in lower grade gliomas, questioning the sensitivity and specificity of *BCOR* immunohistochemistry for identifying BCOR-ITD (Rao *et al.*, 2021).

CNS HGNET-BCOR was classified after the 2016 WHO revision, which removed the term primitive neuroectodermal tumor (PNET) from diagnostic terminology. The molecular signature includes ITD in the *BCOR* gene's 3' end, leading to *BCOR* overexpression. BCOR is an epigenetic regulator on the short arm of chromosome X11.4, playing roles in embryological development and tumorigenesis through interactions with histone deacetylases and polycomb repressive complex 1. Typically affecting children with a mean age of 5.5 years and a male/female ratio of 1.4, the cerebellar hemisphere is the most frequently involved region (De Lima *et al.*, 2020).

Symptoms vary by tumor location. Infratentorial tumors often cause headache and vomiting due to raised intracranial pressure and hydrocephalus, while supratentorial tumors may present with seizures, headache, and motor deficits. DNA methylation profiling is necessary for diagnosis, as there are no specific radiological or pathological features distinguishing HGNET BCOR from other malignant neuroepithelial tumors. Metastases are typically absent at diagnosis but common in recurrences (De Lima *et al.*, 2020; Sugino *et al.*, 2023). Clinical research on CNS HGNET-BCOR is in its early stages, and more extensive studies and clinical trials are needed to understand their behavior and develop optimal management strategies.

2.2.2.3 Embryonal tumor with multilayered rosettes (ETMR)

Embryonal tumors with Multilayered Rosettes (ETMR) is a rare and highly lethal type of WHO-grade IV brain tumor that predominantly affects infants under the age of 3 years (Figure 4A). This tumor type was recognized relatively recently, and advances in molecular diagnostics have improved our understanding and identification of ETMRs (Lambo *et al.*, 2019, 2020).

ETMRs are marked by the amplification of the chromosome 19 miRNA cluster (C19MC), a genetic hallmark found in approximately 90% of cases (Figure 4C, 4E). Another critical feature is the high expression of the RNA-binding protein *LIN28A* (Figure 4K), which is used as a diagnostic marker due to its rarity in other brain tumors. Despite the diverse histological appearance of ETMRs, these molecular markers provide a reliable means of diagnosis.

Around 10% of ETMR cases lack the C19MC amplification but often have biallelic mutations in the *DICER1* gene (Figure 4C, 4E), causing a deregulation in the maturation of miRNAs (Figure 4H). These tumors, despite lacking C19MC, show molecular similarities to those with the amplification. These miRNA alterations result in the accumulation of R-loops (Figure 4F), activation of DNA repair pathways (Figure 4G), and unique methylation patterns (Figure 4I) (Lambo *et al.*, 2019). Interestingly, tumors associated with *DICER1* syndrome often acquire copy-number alterations (CNAs) (Figure 4D) (Seki *et al.*, 2014; Koelsche *et al.*, 2018). However, the potential link between R-loop formation and these CNAs remains unclear and warrants further investigation.

This molecular understanding has been crucial in classifying ETMRs, previously grouped under the broader category of primitive neuroectodermal tumors (PNETs). Features commonly observed in ETMRs include large areas of neuropil and characteristic rosette structures formed by layers of NECs and multiple distinct cell populations (Figure 3, Figure 4J) (Lambo *et al.*, 2019, 2020).

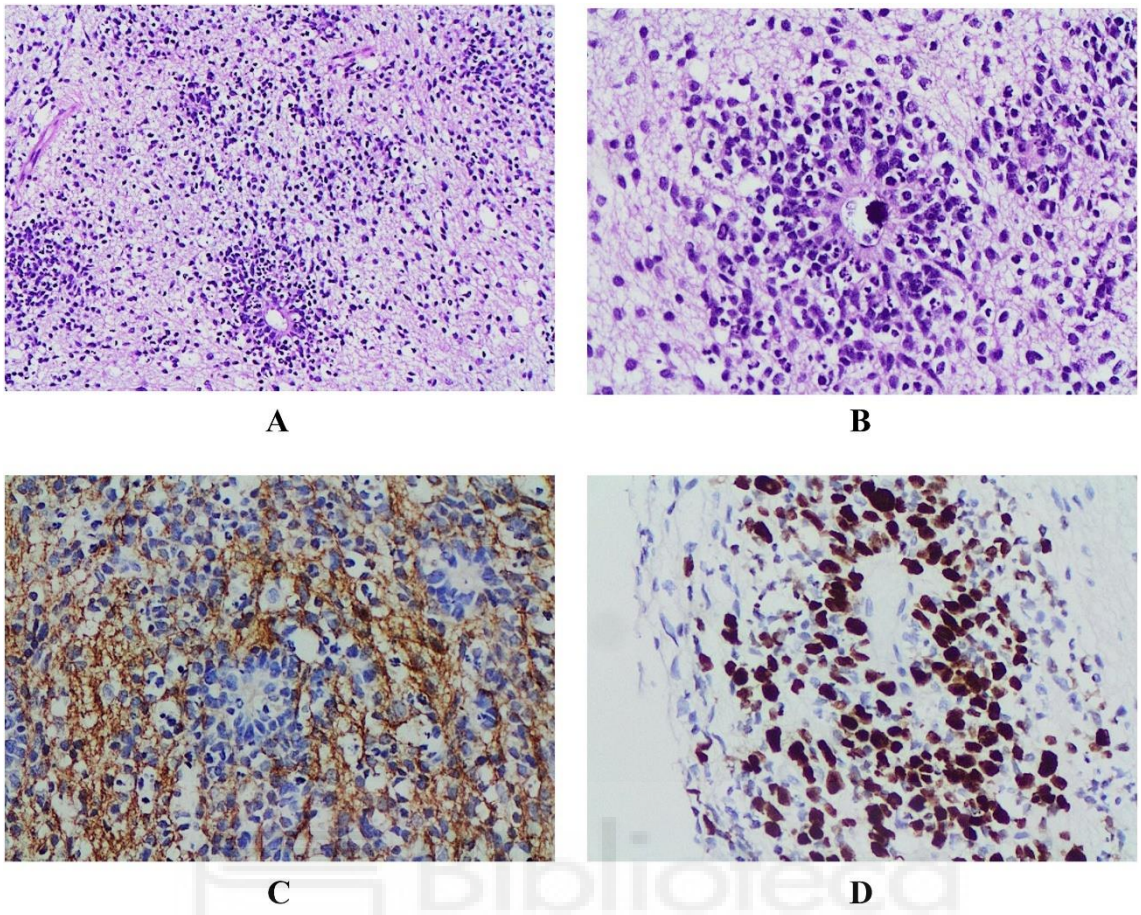


Figure 3. Histological and Immunohistochemical Features of ETMR. A) Low magnification reveals a biphasic architecture consisting of primitive and more differentiated tumor components (Hematoxylin and Eosin staining; 10x). **B**) True multilayered rosettes are observed within the primitive regions, accompanied by numerous apoptotic cells (Hematoxylin and Eosin staining; 20x). **C**) Immunohistochemical staining for synaptophysin demonstrates positivity within the differentiated neuropil areas (20x). **D**) The Ki-67 proliferation index is markedly high, with approximately 70–80% of tumor cells in the primitive embryonal regions showing positive labeling (20x). **Reprinted from** *Embryonal tumor with multilayered rosettes; rare pediatric CNS tumor*. A case report and review of literature, Volume 9, Issue 3, 2022, Pages 174-178, *Case Reports in International Journal of Pediatric and Adolescent Medicine,* Abdelrazak Meliti, Wedad Gasim, Haneen Al-Maghrabi, and Ghadeer Mokhtar, under CC BY-NC-ND 4.0. Available at: https://doi.org/10.1016/j.ijpam.2021.11.002.

ETMRs are primarily found in the brain, with about 70% in supratentorial regions and 30% in infratentorial regions (Figure 4L). Rarely, these tumors can also occur in the spine (Gessi *et al.*, 2009; Korshunov *et al.*, 2014; Horwitz *et al.*, 2016). On MRI, ETMRs typically present as large, heterogeneous masses with cystic components, hemorrhage, and rapid growth, which often leads to severe clinical symptoms. ETMRs predominantly affect infants under three years old, and reliable epidemiological data are still being developed (Figure 4A). Historically, these tumors were often misdiagnosed, complicating accurate incidence assessments. The incidence appears to be evenly distributed between males and females (Lambo *et al.*, 2019, 2020).

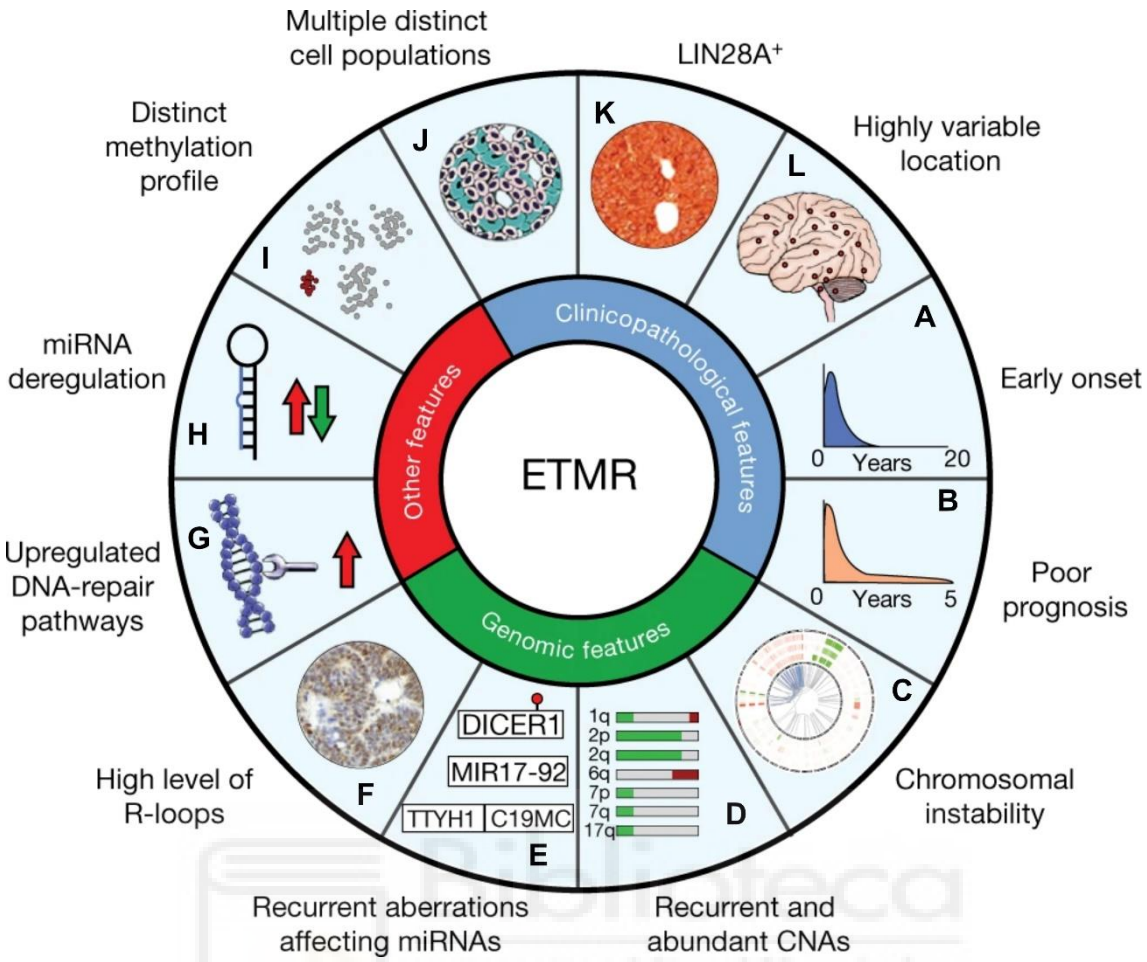


Figure 4. Key Features of Embryonal Tumor with Multilayered Rosettes (ETMR). This figure summarizes the genomic, clinicopathological, and additional features of ETMR, a highly lethal WHO-grade IV brain tumor predominantly affecting infants under three years of age. **A**) Clinicopathological features highlight its early onset, **B**) poor prognosis with 5-year survival rates below 30%, **L**) highly variable tumor locations, **J**) and a heterogeneous composition of multiple distinct cell populations forming rosettes along with **K**) the overexpression of *LIN28A* as a diagnostic marker. **C**), **D**) Genomic features include widespread chromosomal instability with recurrent CNAs, **E**) amplification of the chromosome 19 miRNA cluster (C19MC) in ~90% of cases and *DICER1* mutations in ~10% of cases, **F**) and high levels of R-loops contributing to genomic instability. Other features include **G**) upregulated DNA-repair pathways, **H**) deregulation of miRNAs, and **I**) distinct methylation profiles unique to ETMRs. Together, these features underscore the aggressive nature of ETMR and the critical need for further research and therapeutic advancements. **Adapted from** *The Molecular Landscape of ETMR at Diagnosis and Relapse*, Lambo, S., Gröbner, S.N., Rausch, T. et al., *Nature* 576, 274–280 (2019). Reproduced with permission from Springer Nature (License Number: 5967521286928).

The aggressive nature of ETMRs necessitates intensive treatment. Surgical resection is commonly attempted, with complete resection potentially offering a survival advantage. However, the size and location of these tumors in very young patients make surgery highly risky. Chemotherapy regimens for ETMRs often borrow from treatments for other high-risk CNS tumors, including combinations of drugs. Radiation therapy is generally avoided due to the high risk of toxicity in young children. Some older patients have received radiation therapy, with mixed results. A few long-term survivors have been reported, often following a

combination of surgery, chemotherapy, and in some cases, radiation therapy (Bandopadhayay and Chi, 2022; Gualano *et al.*, 2023).

Despite aggressive and multimodal treatment approaches, the prognosis for ETMR patients remains poor, with 5-year survival rates between 0% and 30% (Figure 4B). Tumors frequently progress or recur locally, and treatment-resistant cases are common. Factors such as the absence of metastases, complete resection, and high-dose chemotherapy or irradiation have been associated with better outcomes, but many patients still experience early relapse and death. There are sporadic reports of long-term survivors, suggesting potential variability in tumor behavior and response to treatment (Bandopadhayay and Chi, 2022; Gualano *et al.*, 2023).

2.3 ETMR proliferative tumoral cells

Recent research has identified stalled developmental programs as the root cause of pediatric brain tumors, including Embryonal Tumor with Multilayered Rosettes (ETMR) (Jessa *et al.*, 2019). In a landmark study, Jessa et al. mapped the bulk and single-cell transcriptomes of ETMR and other pediatric brain tumors, comparing this data to reference datasets from healthy human fetal brain. Using these data, they postulated that the cell of origin for ETMR likely arises from an oncogenic transformation in prenatal RGCs (Jessa *et al.*, 2019).

In fact, ETMR tumor cells are strikingly similar to embryonic RGCs. However, we find these cells in children around three years old. As mentioned before, RGCs play a crucial role in the development of the cortex, acting as progenitors for neurons and glia and facilitating the migration of newly formed neurons. In ETMR, tumor cells mimic these progenitor characteristics, exhibiting similar phenotypic traits and behaviors. This resemblance underscores the theory that ETMR arises from disruptions in normal developmental processes, where cells intended to contribute to brain development in the embryo become immortal and tumorigenic in development or early childhood. Understanding this relationship between ETMR cells and RGCs may provide further insights into potential therapeutic targets and strategies for overcoming the developmental arrest that characterizes these tumors (Jessa *et al.*, 2019; Gualano *et al.*, 2023).

2.4 ETMR mouse models

Mouse models play a critical role in cancer research, offering invaluable insights into tumor biology, genetics, and potential therapeutic approaches. These models are created through various methods such as genetic manipulation, chemical induction, or transplantation of human tumor cells into mice. They mimic human cancer characteristics, allowing researchers to study tumor development, progression, and response to treatments in a controlled environment. The use of mouse models has been instrumental in understanding the mechanisms of oncogenesis and testing new cancer therapies, ultimately contributing to the development of more effective and targeted treatments for human cancers.

Specifically, ETMR mouse models have been developed to closely mimic the human condition, enabling the study of tumor biology and the evaluation of potential therapies. A notable ETMR mouse model was detailed by Julia E. Neumann and colleagues (Neumann *et al.*, 2017). This research reveals key insights into the molecular mechanisms driving ETMRs and highlights promising therapeutic treatments.

However, despite these advancements, the model is not without its limitations. The mouse model does not express *LIN28A*, the main gene marker for human ETMRs, which is crucial for accurately mimicking the human condition. Additionally, the histology of the tumors in this mouse model does not fully resemble the histology of human ETMRs. Because of these limitations, a better mouse model that includes *LIN28A* expression and more closely mirrors the histological features of human ETMRs is urgently needed. Such an improved model would provide a more accurate platform for studying ETMR biology and for the development of effective therapies.

Interestingly, a previous study from our lab that employed *Rx-Cre-Dicer*$^{F/F}$ *(Rx-Dicer1)* mutant mice to investigate how the absence of mature miRNAs influences cortical development, revealed a phenotype that closely resembles to ETMR (Fernández *et al.*, 2020). This model was generated through the conditional deletion of *Dicer1*, a key component of the miRNA processing pathway, in neural progenitors of the embryonic telencephalon. The *Rx-Cre* line was chosen for its early, region-specific expression pattern, as *Rx* is first detected in the anterior neural fold (the prospective forebrain) at E7.5 (Furukawa, Kozak and Cepko, 1997).

By disrupting miRNA maturation at an early stage of forebrain development, these *Rx-Dicer1* mice recapitulated several key features of ETMR biology, including the formation of multilayered rosettes and aberrant gene expression profiles. Consequently, their phenotype and transcriptomic profile closely resemble those of ETMR patients, suggesting that *Rx-Dicer1* mutants may serve as an improved model for studying this aggressive pediatric brain tumor. Nevertheless, further validation and direct comparisons with human ETMR samples are necessary to confirm their reliability as an ETMR model.

3. Genetic regulation of modes of cortical neurogenesis.

3.1 New genes

Genes are one of the main drivers of evolutionary change, particularly when new functions are restricted to specific tissues or organs (Stern, 2000; Carroll, 2005, 2008; Chen, Krinsky and Long, 2013). In the evolution of the cortex, numerous studies have documented changes in the genetic toolkit associated with cortical development. For instance, over 30 gene families have expanded uniquely in the human lineage, many of which are involved in neurodevelopment (Fortna *et al.*, 2004; Sudmant *et al.*, 2010; Zhang *et al.*, 2011). Additionally, several genes that

emerged in primates and humans are highly expressed in the developing cortex (Camp *et al.*, 2015; Florio *et al.*, 2015, 2018; Suzuki *et al.*, 2018).

New genes often arise from the duplication of existing ones followed by sequence modifications (Bailey *et al.*, 2002; Fortna *et al.*, 2004; Chen, Krinsky and Long, 2013). *ARHGAP11B*, a human-specific gene, is expressed in embryonic NCx progenitor cells and originated from a partial duplication of *ARHGAP11A*, followed by a critical sequence change (Sudmant *et al.*, 2010; Antonacci *et al.*, 2014). This modification resulted in a new protein that localizes to the mitochondrial membrane and modulates mitochondrial metabolism, enhancing BP proliferation and NCx expansion (Figure 5D) (Florio *et al.*, 2016; Namba *et al.*, 2021).

Another significant example is *NOTCH2NL*, a gene involved in neural progenitor proliferation that originated from a partial duplication of *NOTCH2*. This gene has several paralogues, three of which are located near *NBPF* genes, suggesting co-evolution (Heide and Huttner, 2021). *NOTCH2NL* proteins enhance *Notch* signaling, which is crucial for neural stem cell maintenance and proliferation (Figure 5F) (Fiddes *et al.*, 2018; Florio *et al.*, 2018; Suzuki *et al.*, 2018).

Comparative genomics has identified 87 proteins with single-amino-acid changes that distinguish modern *Homo sapiens* from *Neanderthals*, with some expressed in the cortical germinal layers and involved in mitosis (Prüfer *et al.*, 2014). For example, *NOVA1*, differing from its Neanderthal ortholog by a single amino acid, impacts synapse formation and neural network function (Trujillo *et al.*, 2021). These subtle genetic changes suggest ongoing adjustments in cortical development throughout human evolution.

3.2 Conserved genes

The emergence of new genes significantly contributes to phenotypic diversity in evolution, but it alone cannot explain the vast diversity seen in nature. Many key genetic regulators of cortical development are conserved throughout the vertebrate lineage (Fernandez *et al.*, 1998; Puelles *et al.*, 2000; Yamashita *et al.*, 2018). However, their expression patterns vary widely across species and even different brain regions within a single species.

3.2.1 *Pax6*

Pax6, a paired-box transcription factor (TF), plays a crucial role in cortical development by regulating gene expression in progenitor cells, with effects that vary significantly depending on the context (Mi *et al.*, 2013; Manuel *et al.*, 2015). *Pax6* affects progenitor cell proliferation and re-entry into the cell cycle in a dose- and context-dependent manner. High *Pax6* expression at early corticogenesis stages and loss of *Pax6* at later stages both lead to premature cell-cycle exit (Haubst *et al.*, 2004; Manuel *et al.*, 2015; Wong *et al.*, 2015). Additionally, *Pax6* acts as a neural fate determinant in the developing cortex, influencing the number

of neurons produced by radial glial cells (RGCs); its absence results in a 50% reduction in neuron numbers, while overexpression increases neuron production (Schmahl *et al.*, 1993; Heins *et al.*, 2002). Overall, *Pax6*-mediated regulation of cell proliferation and neuronal differentiation is critical for maintaining the precise developmental timing necessary for normal brain formation (Figure 5B) (Ochi *et al.*, 2022).

Interestingly, *Pax6* exhibits species-specific functions during pallial development. In the mouse embryonic NCx, sustained *Pax6* expression in progenitors induces primate-like dynamics and increases BP numbers, suggesting a link between prolonged *Pax6* expression, SVZ enlargement, and NCx expansion (Wong *et al.*, 2015). Conversely, in the avian brain, *Pax6* has distinct regulatory roles concerning cellular dynamics and target genes compared to mammals. Specifically, while *Pax6*'s role in progenitor cell maintenance is unique to mammals, its function in promoting neuronal differentiation is conserved across both mammals and birds (Yamashita *et al.*, 2018). *Pax6* also regulates neuronal differentiation, particularly affecting upper layer neurons (Tuoc *et al.*, 2009), through the positive regulation of proneural genes such as *Ngn2* and *Eomes* (Heins *et al.*, 2002; Scardigli *et al.*, 2003; Sansom *et al.*, 2009).

Recently, it has been discovered that within mammals, coupling of *Pax6* and H3K9ac caused BP pool expansion, leading to enhanced neurogenesis, which evoked expansion and proto folding of the mouse NCx (Sokpor *et al.*, 2024).

3.2.2 *Eomes* (*Tbr2*)

Eomes (*Tbr2*), a member of the T-box family of transcription factors, plays a crucial role in the amplification of IPCs in mammals. Since IPCs contribute significantly to neuronal output, *Eomes* is directly linked to cortical expansion (Sessa *et al.*, 2008). In the developing ferret cortex, the loss of *Eomes* function results in reduced cortical folding (Toda *et al.*, 2016), while its deletion in mice leads to a marked reduction in IPC-dependent neurogenesis (Mihalas *et al.*, 2016), underscoring its role in the amplification of BPs.

Cell lineage and single-cell transcriptomic studies in mice indicate that most cortical glutamatergic neurons pass through intermediate states of *Eomes* expression (Mihalas *et al.*, 2016; Mihalas and Hevner, 2018). This suggests that *Eomes* regulates the transition from RGCs to IPCs by gradually deactivating the RGC genetic program in favor of neurogenesis. In fact, in *Pax6* null NCx, *Eomes* levels are profoundly reduced (~2.5-fold), and abventricular mitoses are moderately decreased, being a key gene for BP formation (Figure 5D) (Hevner, 2023).

However, mitotic *Eomes*+ cells have been observed in the developing pallium of various vertebrates, including sharks, lizards, turtles, chicks, and doves, some of them without BPs, as we mentioned before (Clinton *et al.*, 2014; Martínez-Cerdeño *et al.*, 2016; Cárdenas *et al.*, 2018; Docampo-Seara *et al.*, 2018). This observation suggests that the amplification of neuronal output through Eomes+

progenitors may be an ancient mechanism in vertebrate evolution, but its function varies between species.

3.2.3 bHLH transcription factors

Basic-helix-loop-helix (bHLH) TFs are key downstream effectors of Notch signaling, crucial for regulating neural fate. *Hes1*, a transcriptional repressor expressed in neural progenitor cells, exhibits oscillatory mRNA expression. Disruption of these oscillations affects progenitor cell proliferation maintenance (Imayoshi *et al.*, 2013; Imayoshi, Ishidate and Kageyama, 2015; Ochi *et al.*, 2020). *Hes1* oscillations may result from intronic-dependent expression delays, with alterations in their length leading to unstable oscillations (Ochi *et al.*, 2020). Negative feedback through *miR-9* may enhance and stabilize *Hes1* oscillations (Bonev, Stanley and Papalopulu, 2012; Goodfellow *et al.*, 2014). Mathematical models suggest that the intrinsic biological noise of the *Hes1*/*miR-9* oscillatory system favors progenitor fate, whereas *Hes5* oscillations become periodic during differentiation (Phillips *et al.*, 2016; Manning *et al.*, 2019).

Other mechanisms regulating *Hes1* expression and cortical progenitor cell fate include Robo signaling. In the mouse cerebral cortex, *Robo1* and *Robo2* receptors enhance *Hes1* transcription, restraining IPC generation in favor of aRGC self-renewal (Borrell *et al.*, 2012). *Hes1* oscillations repress the expression of other bHLH proneural genes like *Ascl1, Neurog2, and Dll1*, driving their oscillatory expression (Kageyama *et al.*, 2008; Imayoshi *et al.*, 2013; Imayoshi, Ishidate and Kageyama, 2015; Shimojo and Kageyama, 2016; Kaise and Kageyama, 2021).

Ascl1 oscillations induce progenitor cell proliferation, whereas sustained expression promotes neuronal differentiation (Imayoshi *et al.*, 2013; Ochi *et al.*, 2020), and *Neurog2* is necessary for neural stem cells (NSCs) to produce neuron-fated IPCs via cell division. *Neurog2* oscillations induce the accumulation of *Tbr2*, which suppresses *Hes1* expression, generating an IPC-like gene expression state in NSCs (Figure 5D). In the absence of *Tbr2*, *Hes1* expression is up regulated, decreasing the formation of IPCs. In conclusion, the *Neurog2*-*Tbr2* axis forms a continuous transcriptional trajectory to an IPC-like neurogenic state in NSCs, which then generate IPCs via cell division (Shimojo, Masaki and Kageyama, 2024).

3.2.4 *Robo*/*Dll1* signaling pathway

Robo/*Dll1* signaling is crucial in balancing direct and indirect neurogenesis in the developing cerebral cortex, as we saw in previous research from our lab by modulating the Notch pathway (Cárdenas *et al.*, 2018). In this research across species, we demonstrated that during amniote evolution, the attenuation of *Robo1* and *Robo2* and the increase in *Dll1* expression in RGCs shifted the balance from direct to indirect neurogenesis. This genetic evolution was crucial for cortical expansion in amniotes and the emergence of key features of the

mammalian NCx, such as the formation of IPCs, the SVZ, and layer 2/3 neurons (Cheung *et al.*, 2010).

High *Robo* signaling drives direct neurogenesis, limiting neuron number and the size of telencephalic structures, such as the olfactory bulb (OB) and the non-mammalian cortex. Conversely, low *Robo* signaling allows indirect neurogenesis, producing BPs and more neurons, as observed in the mammalian NCx. However, the complete absence of *Robo*, as in knockout mice, is severely detrimental to BPs, impairing further NCx expansion. In the OB of these mutants, this deleterious effect combines with suppressed direct neurogenesis, resulting in reduced neuron production and OB size. RGCs in direct neurogenesis mode have a longer cell cycle than in indirect mode, so the loss of *Robo* also alters the frequency of their apical mitoses. The magnitude of this effect varies between cortical regions, likely due to axial gradients in cortical development (Borrell *et al.*, 2012).

High levels of *Robo* block *Dll1* expression, promoting the expression of *Jag1* and *Jag2*, which drive direct neurogenesis, as seen in the OB (Figure 5C). In contrast, the NCx favors indirect neurogenesis due to low *Robo* signaling. Current and previous results from studies on OB and NCx in *Robo* mutants consistently suggest that the presence or absence of *Jag1* and *Jag2*, along with varying levels of *Robo* expression, modifies how *Robo* influences *Hes1* expression in the telencephalon.

The influence of *Robo* signaling on cortical expansion extends beyond driving direct neurogenesis. High *Robo* signaling prevents BP formation in the snake dorsal cortex, where they are virtually absent. Remarkably, experimental loss of *Robo* and gain of *Dll1* in snake embryos induces de novo formation of BPs, aligning basal to the ventricular zone and forming a proto-SVZ, a process typically restricted to the mammalian NCx (Figure 5D). These findings suggest that attenuation of *Robo* receptor signaling intensity during evolution was crucial for the emergence of BPs and the SVZ, as well as the blockade of direct neurogenesis, which collectively allowed the significant expansion and complexity of the mammalian cerebral cortex.

3.2.5 *Hopx*

Homeodomain-only protein (*Hopx*) is an unusual homeodomain protein that lacks DNA binding sites (Zweifel *et al.*, 2018) and serves as a marker for bRGCs, including their basal fibers (Pollen *et al.*, 2015). *Hopx* is expressed in the developing cerebral cortex of both ferrets and mice, where it promotes the expansion of bRGCs without affecting the dynamics of aRGCs (Figure 5D) (Vaid *et al.*, 2018). Despite the presence of bRGCs in the lissencephalic mouse medial NCx and the near-lissencephalic primate SVZ (Fietz *et al.*, 2010; García-Moreno *et al.*, 2012; Kelava *et al.*, 2012; Vaid *et al.*, 2018), their role in gyrencephaly remains unclear. bRGCs may be necessary but not sufficient for cortical folding, and the exact role of *Hopx* in this process is still undetermined (Vaid *et al.*, 2018; Matsumoto *et al.*, 2020).

Recent scRNA-seq from individual cortical germinal zones in ferrets revealed that *Hopx* is expressed across RGC subclusters, with a noticeable enrichment in the VZ compared to the outer subventricular zone (OSVZ) (Del-Valle-Anton *et al.*, 2024). This expression pattern mirrors that seen in the human embryonic cortex before gestational week 16 (GW16), aligning with previous studies that question *Hopx* as a universal marker for bRGCs (García-Moreno *et al.*, 2012; Kelava *et al.*, 2012). The relative abundance of *Hopx* is emerging as a key factor in understanding its role, as recent research has identified both Hopx-positive and Hopx-negative bRGCs, suggesting a more complex and nuanced role for *Hopx* in cortical development (Vaid *et al.*, 2018).

3.2.6 *Sall1*

Sall1 is a putative TF containing a C2H2 zinc-finger, highly expressed in the developing CNS and peripheral organs. Previous research has demonstrated that Sall gene family members are involved in various critical cellular processes such as cell cycle regulation, proliferation, neuronal differentiation, migration, and cell adhesion across different species (Basson and Horvitz, 1996; de Celis, Barrio and Kafatos, 1999; Cantera *et al.*, 2002; Barembaum and Bronner-Fraser, 2004; Sakaki-Yumoto *et al.*, 2006; Harrison, Nishinakamura and Monaghan, 2008; Harrison, Parrish and Monaghan, 2008).

Sall1 is expressed in aRGCs during neurogenesis. Using knockout studies, researchers found that *Sall1* is necessary for proper development in the dorsal aRGCs. At embryonic day (E) 18.5, the cerebral cortex of *Sall1*−/− mice was smaller in both surface area and depth compared to controls, due to changes in the division type of aRGCs. Early in development, *Sall1*−/− aRGCs exited the cell cycle more often, while later, IPCs re-entered the cycle more frequently. This led to more neurons adopting an early cortical fate and fewer adopting a later cortical fate compared to controls, suggesting that *Sall1* regulates cortical neurogenesis timing in aRGCs (Figure 5B) (Harrison *et al.*, 2012).

Sall1's role in IPC function is intriguing. Although highly expressed in RGCs and downregulated in IPCs, *Sall1* significantly impacts IPC properties. Several molecules regulate IPC proliferation and differentiation, with *Tbr2* promoting an IPC phenotype and other proteins like *Frs2α, Foxg1* (Figure 5E), and *Insm1* (Figure 5D) supporting IPC expansion. *Sall1* might interact with the *Wnt* pathway, influencing RGC to IPC transition (Harrison *et al.*, 2012). Interestingly, in colorectal cancer cells, *SALL1* upregulates p-p65 and p-JUN expression, and the c-Fos/activator protein (AP)-1 inhibitor (T-5224) can reverse the induction of CRC progression by *SALL1* overexpression, further highlighting a strong relationship between *SALL1* and the *Wnt* pathway (Yuan *et al.*, 2023).

3.2.7 *Dmrta1*

Dmrt (double-sex and mab-3 related transcription factor) genes, first identified in Drosophila and C. elegans, are critical for sexual development and CNS

development. The Dmrta subfamily, including *Dmrt3, Dmrta1, and Dmrta2,* contains a conserved DMA domain necessary for their activity. These genes can form heterodimers on DNA, suggesting combinatorial gene regulation. In vertebrates, while Dmrt genes are essential for sexual organ development, recent research highlights their importance in CNS development. Specifically, *Dmrta1* is expressed in aRGCs in the brain and specific neurons in the spinal cord during development (Kikkawa and Osumi, 2021).

Dmrta1 is expressed in the VZ of the developing mouse telencephalon as early as E9.5, peaking around E10.5-E12.5. In the rat, its expression begins on E10.5. *Dmrta1* then is crucial for the proper patterning of the telencephalon, being a direct target of the TF *Pax6*. *Dmrta1* also plays a significant role in neurogenesis by influencing the expression of *Neurog2*. *Neurog2*, as mentioned before, is a key TF involved in the commitment of aRGCs into IPCs. In the dorsal telencephalon, *Dmrta1* promotes indirect neurogenesis by repressing *Ascl1*, and facilitating the expression of *Neurog2* (Figure 5D). This regulation ensures the correct production of excitatory neurons in the dorsal telencephalon (Kikkawa *et al.*, 2013).

3.3 Genetic cis-regulatory elements

Cis-regulatory elements (CREs) are key genomic sequences that regulate gene expression by serving as binding sites for TFs and other regulatory proteins. The core types of CREs include promoters, enhancers, silencers, and insulators, each with distinct roles in orchestrating transcription. Promoters are typically located near the transcription start site (TSS) of genes and function as platforms for the assembly of the transcriptional machinery (Shibata, Gulden and Sestan, 2015). These can be further classified as proximal promoters, which are closely associated with the TSS and actively recruit RNA polymerase II and general TFs, or distal promoters, which are situated farther from the TSS and exert weaker regulatory influence. Active promoters are often marked by histone modifications such as trimethylation of histone H3 at lysine 4 (H3K4me3), which indicates a transcriptionally permissive chromatin state (Heintzman *et al.*, 2007).

Enhancers are another essential type of CRE, often located at significant distances from their target genes. They enhance gene expression by recruiting TFs and cofactors with histone acetyltransferase activity, which facilitates chromatin decondensation and promotes access of the transcriptional machinery to promoters (Shibata, Gulden and Sestan, 2015). Enhancers are characterized by specific histone modifications such as H3K27ac, H3K4me1, H3.3, and H2A.Z (Barski *et al.*, 2007; Wang and Hayes, 2008). Among enhancers, super enhancers represent a specialized subset that consists of large clusters of densely occupied enhancers, spanning up to 50 kilobases. These super enhancers are strongly associated with the regulation of genes critical for cell identity and differentiation, making them particularly sensitive to perturbations in TFs or mediator complexes.

In contrast to enhancers, silencers are regulatory elements that repress gene expression. Silencers can act in a temporally or tissue-specific manner, ensuring that genes are expressed only when and where they are needed. Silencers achieve this by recruiting proteins such as REST and SUZ12 (Lee *et al.*, 2006; Johnson *et al.*, 2007), which promote chromatin compaction and the establishment of repressive histone modifications. Insulators, on the other hand, play a distinct role in maintaining the structural and functional organization of the genome. By preventing interactions between unrelated enhancers and promoters or blocking the spread of heterochromatin, insulators help define independent regulatory domains within the genome. The CCCTC-binding factor (CTCF) is a key marker of insulator activity and is crucial for genome integrity (Wendt *et al.*, 2008).

Beyond these core CREs, other regulatory elements also contribute to gene expression regulation. These include 3′ untranslated regions (3′ UTRs), which influence mRNA stability and translation, and splicing regulatory elements that modulate alternative splicing. A recent SLAM-seq study of the developing mouse cortex showed that transcript-specific control of mRNA half-life, shaped by 3′ UTR motifs and the CCR4-NOT decay complex, is essential for proper neurogenesis (Serdar *et al.*, 2025). Additionally, epigenetic factors such as DNA methylation and histone modifications play dynamic roles in regulating chromatin state and transcriptional activity.

Human accelerated regions (HARs), for example, are conserved CREs in nonhuman primates but highly mutated in humans, with many functioning as regulatory enhancers (Capra *et al.*, 2013; Girskis *et al.*, 2021; Uebbing *et al.*, 2021). HARs are often near genes crucial for brain development, suggesting roles in proliferation and differentiation during cortical development (Won et al., 2019). HARs have been linked to cortical expansion, as shown by differing enhancer activity in human and chimpanzee HARE5, influencing brain size in mouse models (Boyd *et al.*, 2015). HAR-regulated genes like *PPP1R17*, expressed in primate cortex progenitor cells, highlight regulatory mechanisms' importance in cortical evolution (Girskis *et al.*, 2021). Recently, researchers used sgRNA-Cas9 screens in human neural stem cells (hNSCs) to disrupt 10,674 genes and 26,385 conserved regions in 2,227 enhancers active in the developing human cortex, examining their effects on proliferation. Although enhancer disruptions generally have weaker effects than gene disruptions, some enhancer disruptions significantly alter hNSC self-renewal. Disruptions in HARs affected clearly proliferation (Geller *et al.*, 2024).

Enhancers are also key in the NCx evolution. For instance, neCtnnb1 is an evolutionarily conserved CRE located 55 kilobases upstream and spatially close to the promoter of *Ctnnb1* and plays a crucial role in gene regulation. Activation or interference of the neCtnnb1 locus using CRISPR/Cas9 technology resulted in increased or decreased transcription of *Ctnnb1*, respectively. The knockout of neCtnnb1 led to reduced expression of *Ctnnb1*, which in turn impaired the production and transit-amplification of BPs (Wang *et al.*, 2022).

Enhancers are pivotal in driving evolutionary innovation by enabling rapid and flexible adaptations without altering the core genetic code. This regulatory evolution allows species to diversify and adapt to new environments efficiently, underscoring that the evolution of gene regulation is a primary driver of evolutionary change.

3.4 Chromatin organization and TADs

Chromatin structure and organization also regulate gene expression. DNA wraps around histones, forming nucleosomes, and histone modifications affect transcription (Stillman, 2018). For example, human BPs show higher H3 acetylation than their mouse counterparts, promoting their proliferation (Kerimoglu *et al.*, 2021). DNA methylation, particularly on cytosines in CG pairs, can reduce gene expression and is relevant in cortical evolution (Lister *et al.*, 2013; Moore, Le and Fan, 2013). The human prefrontal cortex exhibits distinct epigenetic and transcriptional features during development that contribute to its unique functional and structural properties. Compared to chimpanzees, the human prefrontal cortex has more hypomethylated sites, many of which contain binding motifs for transcription factors (TFs) such as *FOXP1* (Jeong *et al.*, 2021). *FOXP1* plays a critical role in early cortical development, particularly at a developmental time point when its expression peaks in cortical progenitors, specifically the bRGCs. Loss of *FOXP1* leads to a significant reduction in the number of bRGCs, along with impaired proliferation and differentiation of the remaining cells, ultimately resulting in a decreased production of excitatory cortical neurons (Park, Kulkarni and Konopka, 2023). Chromatin conformation shifts from open during neurogenesis to compact after postnatally. This is mediated by HMGA proteins, affecting gene expression (Vogel and Marcotte, 2012)(Kishi et al., 2012).

Topologically associated domains (TADs) are conserved across bilaterian mammals and bring gene regulatory elements together, facilitating gene expression regulation (Acemel *et al.*, 2016). However, emerging high-resolution Hi-C and multi-omics data indicate that specific TAD boundaries and broader 3D chromatin architecture can be species-specific, differing between humans, chimpanzees and macaques (Keough *et al.*, 2023; Brand *et al.*, 2024). TADs' role in cortical development has been studied in humans, macaques, and mice, revealing human-specific TADs and new enhancer interactions (Luo *et al.*, 2021).

Additionally, the proneural TF *Neurog2* plays a crucial role in 3D chromatin structure, by acting as an epigenome remodeler. It induces local chromatin accessibility and DNA demethylation at its binding sites, enhancing chromatin looping between enhancers and promoters. This facilitates stronger transcriptional activation of target genes. The simultaneous binding of *Neurog2* to enhancers and promoters contributes to dynamic and cell-type-specific chromatin interactions, ensuring robust changes in gene expression and cell fate during cortical development and neurogenesis (Noack *et al.*, 2022).

3.5 Post-transcriptional mechanisms

Differences in mRNA levels alone cannot account for the full variability observed in protein abundance, highlighting the importance of posttranscriptional mechanisms (Vogel and Marcotte, 2012). Posttranscriptional regulation is a complex process involving numerous players that interact both in the nucleus and cytoplasm.

Chronologically, the first mechanism is alternative splicing, which is linked to the transcriptional process itself. The spliceosome complex assembles while RNA polymerase is still active, regulating the splicing process and the selection of specific exons. This selection is also influenced by chromatin structure (Naftelberg *et al.*, 2015). Alternative splicing can affect posttranslational modifications and protein-protein interactions, such as through the insertion of microexons (Ule and Blencowe, 2019).

Like DNA, mRNA is susceptible to covalent modifications of its nucleotides, known as epitranscriptomics (Fu *et al.*, 2014; Yoon *et al.*, 2018). Although not extensively explored, the most common modification, methylation on adenosine N6, affects nuclear export, microRNA-mediated decay, and sometimes splicing (Noack and Calegari, 2018). This modification is significant for the NPC cycle and the stability of specific neuronal genes (Yoon *et al.*, 2017). Its dynamic patterning is related to cell identity specification, increasing during neuronal differentiation in the developing brain (Meyer, 2017). Additionally, mRNAs can bind multiple RNA-binding proteins (RBPs) that regulate splicing, polyadenylation, nuclear export, cytoplasmic localization, and stabilization (Hsu *et al.*, 2019; Salamon and Rasin, 2021; Sena *et al.*, 2021).

mRNA stability is another regulatory mechanism. miRNAs play a well-documented role in this process. miRNAs are short noncoding RNAs that guide the RISC complex to bind specific mRNAs through complementary sequence recognition, usually at the 3′-UTR, leading to translation inhibition or mRNA degradation (Prieto-Colomina *et al.*, 2021). The canonical miRNA biosynthesis pathway involves the sequential processing of an immature miRNA by *Drosha* in the nucleus and *Dicer1* in the cytoplasm (Miyoshi, Miyoshi and Siomi, 2010; Prieto-Colomina *et al.*, 2021).

miRNAs are also crucial in brain evolution, either through the differential expression of conserved miRNAs or the emergence of new ones, such as a eutherian-specific cluster involved in neuronal identity acquisition (Diaz *et al.*, 2020; Martins *et al.*, 2021). The proliferation of NPCs and the emergence of features crucial for cortical expansion and gyrification, like the OSVZ expansion, are influenced by the evolution of posttranscriptional regulation by miRNAs. For example, *miR-137* and *miR-122* are highly expressed in the OSVZ of developing ferret, macaque, and human cerebral cortex, but not in the embryonic mouse cortex, which lacks an OSVZ. *miR-137* and *miR-122* play essential roles in regulating crucial aspects of cortical expansion. *miR-137* enhances the self-renewal of BPs while directing them toward a superficial layer neuronal identity (Figure 5E). In contrast, *miR-122* slows down the neuronal differentiation

process, maintaining the progenitor pool for extended periods (Figure 5F) (Tomasello *et al.*, 2022).

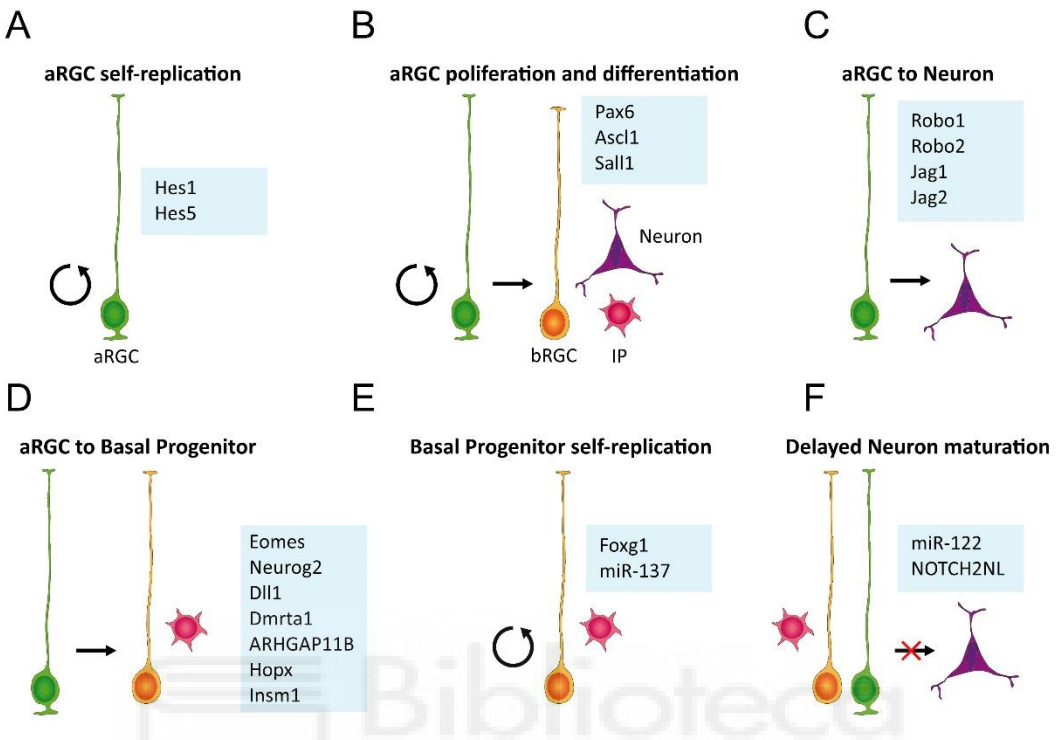


Figure 5. Genetic regulation of the main modes of cortical neurogenesis. A) aRGC self-replication. Apical radial glial cells (aRGCs) maintain their stem-cell identity and proliferate under the influence of *Hes1* and Hes5, which bolster Notch signaling and sustain progenitor pools. **B**) aRGC proliferation and differentiation. Factors such as *Pax6*, *Ascl1*, and *Sall1* balance aRGC proliferation with the onset of neurogenic programs, guiding progenitors to either re-enter the cell cycle or initiate neurogenesis via direct or indirect modes. **C**) aRGC to neuron (direct neurogenesis). High *Robo1/2* activity suppresses *Dll1* and promotes *Jag1/Jag2*, driving the direct differentiation of aRGCs into neurons. **D**) aRGC to basal progenitor (indirect neurogenesis). Multiple genes (*Eomes, Neurog2, Sall1, Dll1, Dmrta1, ARHGAP11B, Hopx, and Insm1*) orchestrate the transition of aRGCs to basal progenitors, boosting neuronal output and enabling cortical expansion. **E**) Basal progenitor self-replication. Basal progenitors, marked by *Foxg1* and *miR-137*, undergo additional rounds of division, further enhancing neuron production. **F**) Delayed neuron maturation. *miR-122* and *NOTCH2NL* can prolong the progenitor state or postpone neuronal maturation, extending progenitor pool maintenance.

4. Genetic mechanisms driving ETMR tumors.

Embryonal Tumors with Multilayered Rosettes (ETMR) are a rare and aggressive form of pediatric brain tumors characterized by unique genetic and molecular aberrations. These tumors exhibit high cellular heterogeneity and rapid proliferation, making them particularly challenging to treat. While the pathological hallmark of ETMRs is the presence of multilayered rosettes, recent research has uncovered key genetic and molecular mechanisms underlying their development

and progression. This chapter explores the genetic drivers of ETMR, focusing on their diverse post-transcriptional and epigenetic alterations. Mutations in genes like *DICER1*, amplification of critical microRNA clusters such as C19MC and *miR-17-92*, and downstream effects on pathways like *LIN28A*/*let-7*, are strong candidates to produce a highly aggressive tumor phenotype. The following sections delve into these molecular alterations, highlighting their role in disrupting normal developmental pathways and fostering oncogenesis in ETMR.

4.1 miRNA dysregulation in ETMR: *DICER1*, C19MC, and *miR-17-92* clusters

Posttranscriptional mechanisms play a crucial role in ETMR tumors, including mutations in key genes such as *DICER1*, overexpression of the C19MC miRNA cluster, and other microRNA-related aberrations like the somatic amplification of the *miR-17-92* cluster (also known as *MIR17HG*) (Lambo *et al.*, 2019).

DICER1 is an essential enzyme of the microRNA (miRNA) biogenesis pathway, responsible for processing precursor miRNAs into mature miRNAs. Mutations in *DICER1* can disrupt this process, leading to widespread changes in miRNA expression and function. In ETMR, *DICER1* mutations result in the impaired production of specific miRNAs, which can affect various cellular pathways involved in tumorigenesis. This disruption contributes to uncontrolled cell proliferation, survival, and differentiation (Lambo *et al.*, 2019)

Another important miRNA mutation is in the C19MC miRNA cluster, located on chromosome 19q13.42, is the largest known human miRNA cluster, consisting of over 50 miRNA genes. This cluster is normally expressed during early embryonic development but is typically silenced in differentiated tissues.

ETMRs are characterized by high-level amplification and overexpression of the C19MC miRNA cluster. This overexpression drives the oncogenic processes by targeting multiple tumor suppressor genes and altering cell cycle regulation, apoptosis, and differentiation pathways. In fact, the fusion of *TTYH1* with the C19MC cluster leads to the activation and overexpression of these miRNAs, which play a crucial role in the aggressive nature of ETMR (Lambo *et al.*, 2019)

The *miR-17-92* cluster, also known as *MIR17HG*, is another critical miRNA cluster implicated in various cancers, including ETMR. This cluster is involved in regulating cell proliferation, apoptosis, and angiogenesis. Somatic amplification of the *miR-17-92* cluster leads to its overexpression, which promotes oncogenic activities by targeting multiple tumor suppressor genes and enhancing cell cycle progression. In ETMR, the amplification of *miR-17-92* contributes to the tumor's malignant phenotype by further disrupting the regulatory networks control cell growth and apoptosis (Lambo *et al.*, 2019).

A critical question remains: how do mutually exclusive genetic aberrations, such as C19MC amplification, *DICER1* mutations, or *MIR17HG* overexpression lead to a common downstream mechanism, given the molecular similarity of these tumors? All these conditions have been shown to increase the efficiency of reprogramming stem cells and delay differentiation. This occurs because the

three conditions cause dysregulation of the *LIN28A*/*let-7* pathway, leading to the upregulation of several critical oncogenes and factors implicated in ETMR pathogenesis. Currently, there is evidence suggesting that C19MC may affect *LIN28A* indirectly. For instance, C19MC can downregulate Tristetraprolin (TTP), a protein that degrades LIN28A. Other factors may also be involved, as many different miRNAs have been shown to target *LIN28A/B* directly or indirectly, including *let-7* miRNAs, which generate a negative feedback loop. This suggests that deregulated miRNA processing could underlie changes in the regulation of *LIN28A* (Lambo *et al.*, 2019).

A previous study in our lab showed that early *DICER1* mutation in the telencephalon disrupts the processing of *let-7* microRNAs. This mutation impairs their production and maturation, resulting in decreased levels of mature *let-7*. Normally, mature *let-7* represses several important genes for development, including Protogenin (*Prtg*), *Lin28a/b*, and others involved in the *p53* pathway. However, with reduced *let-7* levels, due to the *Dicer1* mutation, these genes are not effectively repressed and are upregulated. The upregulation of *Prtg, Lin28a/b*, *Irs2*, and *p53* pathway genes leads to significant developmental consequences, such as the formation of neural proliferative rosettes (Fernández *et al.*, 2020).

4.2 Coding genes targeted by *Let-7*

The *let-7* family of microRNAs plays a critical role in regulating key developmental and cellular processes by targeting a wide range of coding genes involved in cell proliferation, differentiation, and apoptosis. Dysregulation of *let-7* leads to the activation of oncogenes and the suppression of tumor suppressors, contributing to the development and progression of aggressive tumors like ETMR. This section explores the coding genes directly influenced by *let-7* dysregulation, including *LIN28A, LIN28B, PRTG, IGDCC3, DNMT3B, MYCN, IRS1* and *IRS2*, highlighting their roles in neural development, tumorigenesis, and the unique characteristics of this pediatric brain tumor.

4.2.1 *LIN28A* and *LIN28B*

LIN28A is an RNA-binding protein that plays a crucial role in the regulation of microRNA processing, specifically inhibiting the maturation of the *let-7* family of microRNAs. *Let-7* miRNAs are known to suppress various oncogenes involved in cell proliferation, differentiation, and apoptosis (Viswanathan, Daley and Gregory, 2008).

By inhibiting *let-7* miRNAs, *LIN28A* indirectly leads to the upregulation of oncogenes such as *MYC, RAS*, and *HMGA2*, which drive cellular proliferation and survival. Its overexpression enhances cell proliferation, promotes the reprogramming of metabolic pathways, and maintains the undifferentiated state of cancer stem cells. These effects contribute to the aggressive nature of tumors (Lin *et al.*, 2022).

LIN28B functions similarly to *LIN28A* by binding to the precursors of *let-7* miRNA and preventing their maturation. It is also involved in post-transcriptional regulation of mRNA, affecting various cellular processes. *LIN28B* is highly expressed during embryonic development and is involved in maintaining the proliferative and pluripotent state of embryonic stem cells. In cancer, *LIN28B* overexpression leads to decreased *let-7* miRNA levels, resulting in the activation of target oncogenes. This promotes increased cell proliferation, invasion, and resistance to apoptosis. *LIN28B* has been implicated in liver cancer, neuroblastoma, and other malignancies, where it enhances the self-renewal and tumor-initiating capacity of cancer stem cells (Molenaar *et al.*, 2012; Nguyen *et al.*, 2014).

In fact, *LIN28A* is a canonical marker for ETMR diagnosis. Using LIN28A-specific antibodies, immunohistochemical analysis on 37 ETMR samples resulted in a strong LIN28A immunoexpression. Only a small proportion (6/50) of AT/RT samples exhibited focal LIN28A reactivity, while all other pediatric brain tumors tested were LIN28A-negative (Korshunov *et al.*, 2012; Rao *et al.*, 2017).

4.2.2 *PRTG* and *IGDCC3*

Protogenin (*PRTG*) is a membrane protein that is part of the immunoglobulin superfamily. It plays a significant role in neural development by mediating cell-cell adhesion, signaling, and interactions essential for processes such as neural tube closure and neurogenesis. *PRTG* is crucial for the regulation of neural progenitor cell behaviors, including their proliferation, migration, and differentiation (Wong *et al.*, 2010).

During embryonic development, *PRTG* is highly expressed in neural progenitor cells. It helps guide these cells to their correct locations and supports their proliferation and differentiation into various neural cell types. *PRTG*'s role in cell adhesion and signaling is vital for the proper formation of the nervous system (Wong *et al.*, 2010).

Although direct evidence linking *PRTG* to ETMR remains limited, emerging findings underscore its critical involvement in both neural development and tumorigenesis. Recent work has identified *PRTG* as a defining marker of a population of high-*MYC*, low-*NESTIN* stem cells in the four-week-old human embryonic hindbrain, localized specifically to the ventricular zone of the rhombic lip (RLVZ). Oncogenic transformation of these early PRTG+ rhombic lip stem cells can initiate group 3 medulloblastoma (Gr3-MB)-like tumors, highlighting the importance of *PRTG* in tumorigenesis (Visvanathan *et al.*, 2024).

Immunoglobulin Superfamily DCC Subclass Member 3 (*IGDCC3*) is a gene that plays a role in neural development, particularly in cell adhesion and axon guidance. *IGDCC3* is involved in the interactions between neural cells, guiding axons to their targets, which is crucial for the proper formation of neural circuits.

Dysregulation of *IGDCC3* can disrupt normal neural cell interactions and signaling pathways, potentially contributing to tumorigenesis. Changes in

IGDCC3 expression may lead to abnormal cell proliferation and differentiation, like *PRTG*. While direct evidence linking *IGDCC3* to ETMR is not well-documented, its role in neural development suggests it could play a part in the development of neural tumors if its function is altered (Maness and Schachner, 2007).

4.2.3 *DNMT3B*

DNA Methyltransferase 3 Beta (*DNMT3B*) is a DNA methyltransferase enzyme responsible for de novo DNA methylation, which involves adding methyl groups to cytosine residues in CpG dinucleotides. This epigenetic modification is critical for regulating gene expression, maintaining genomic stability, and ensuring proper cellular differentiation without altering the DNA sequence (Robertson, 2001).

DNMT3B plays a crucial role in neural development by regulating the expression of genes involved in neural progenitor cell proliferation, differentiation, and maturation. Proper DNA methylation patterns are essential for the correct timing and spatial expression of neural genes, ensuring the development of a functional nervous system. Disruption in *DNMT3B* activity can lead to abnormal neural development and has been implicated in various neurodevelopmental disorders (Okano *et al.*, 1999).

DNMT3B is frequently overexpressed or mutated in various cancers, leading to abnormal DNA methylation patterns that contribute to tumorigenesis. *DNMT3B* can cause hypermethylation of the promoter regions of tumor suppressor genes, leading to their silencing. This loss of tumor suppressor gene function can remove critical controls on cell proliferation, apoptosis, and DNA repair, promoting tumor growth and survival. Conversely, *DNMT3B* can contribute to the hypomethylation and activation of oncogenes, further driving uncontrolled cell proliferation and enhancing tumor aggressiveness. Additionally, abnormal *DNMT3B* activity can lead to widespread genomic instability, increasing the likelihood of mutations and chromosomal rearrangements that drive cancer progression (Rauch *et al.*, 2009; Loaeza-Loaeza *et al.*, 2022).

In the context of ETMR, *DNMT3B*'s role is particularly important due to its ability to epigenetically reprogram neural progenitor cells. Overexpression or mutations in *DNMT3B* lead to the aberrant methylation of key regulatory genes, supporting the aggressive growth and poor differentiation characteristic of ETMR cells. In fact, ETMR was identify an extreme overexpression of a isoform of *DNMT3B*, originating from an alternative promoter active only during the first weeks of neural tube development. This suggests an oncogenic re-engagement of an early developmental program in ETMR through epigenetic alteration mediated by this embryonic, brain-specific *DNMT3B* isoform (Kleinman *et al.*, 2014).

In glioblastoma (GBM), overexpression of *DNMT3B* via the *PI3K/Akt* signaling pathway contributes to temozolomide (TMZ) resistance by promoting tumor cell survival and proliferation. Additionally, *DNMT3B* downregulation enhances the

chemosensitivity of GBM cells to TMZ. Treatment with LY294002, a *PI3K* inhibitor, suppresses this signaling pathway, reduces PI3K phosphorylation, and significantly lowers *DNMT3B* expression in U251-TMZ cells, further increasing chemosensitivity to TMZ. These findings highlight the *PI3K/Akt-DNMT3B* axis as a potential therapeutic target in GBM, and potentially to ETMR (Kan *et al.*, 2024).

4.2.4 *MYCN*

MYCN, a member of the *MYC* family of TFs, plays a critical role in the regulation of gene expression, cell growth, proliferation, differentiation, and apoptosis. It is particularly important during embryonic development and in the maintenance of stem cell populations (Westermark *et al.*, 2011). *MYCN* regulates the expression of genes involved in cell cycle progression and proliferation of neural progenitor cells. Proper expression of *MYCN* ensures the rapid expansion of neural cell populations during brain development (Otte *et al.*, 2021). Also, *MYCN* amplification or overexpression is a hallmark of several aggressive cancers, particularly neuroblastoma, a pediatric cancer of the sympathetic nervous system.

In the context of ETMR, *MYCN* amplification has been observed and is associated with the aggressive behavior of these tumors. *MYCN* works in concert with other oncogenic drivers, such as the C19MC microRNA cluster and *LIN28A*, to promote rapid tumor growth and resistance to differentiation and apoptosis (Kleinman *et al.*, 2014). *MYCN* also drives the expression of genes that promote cell cycle progression, such as cyclins and cyclin-dependent kinases (CDKs), leading to increased cell proliferation (Otte *et al.*, 2021), it suppresses differentiation pathways, maintaining cells in a proliferative, undifferentiated state (Westermark *et al.*, 2011), and influences cellular metabolism, supporting the high metabolic demands of rapidly proliferating cancer cells by upregulating genes involved in glycolysis and other metabolic pathways (Otte *et al.*, 2021).

4.2.5 *IRS1* and *IRS2*

IRS1 (Insulin Receptor Substrate 1) and *IRS2* (Insulin Receptor Substrate 2) are key adaptor proteins that play critical roles in transmitting signals from the insulin and insulin-like growth factor (*IGF*) receptors to intracellular signaling pathways. These pathways regulate various cellular processes, including metabolism, growth, proliferation, and survival .

IRS1 is primarily involved in mediating metabolic functions of insulin. It plays a significant role in glucose uptake and metabolism by activating the *PI3K/Akt* pathway (LeRoith, Scheinman and Bitton-Worms, 2011), while *IRS2* has overlapping but distinct functions compared to *IRS1*, including more prominent roles in regulating cell growth and survival, especially in the context of *IGF* signaling (Lomperta *et al.*, 2020). Both *IRS1* and *IRS2* are expressed in the brain and contribute to neural development by promoting neuronal growth, differentiation, and survival through *IGF* signaling pathways.

IRS1 and *IRS2* are implicated in various cancers due to their roles in promoting cell proliferation and survival through the activation of downstream signaling pathways, such as *PI3K/Ak*t and *MAPK/ERK*. The overexpression or dysregulation of these proteins can contribute to tumorigenesis. *IRS1* is often overexpressed in several cancers, including breast, prostate, and colorectal cancers. Its overexpression enhances cell proliferation, survival, and resistance to apoptosis (LeRoith, Scheinman and Bitton-Worms, 2011). Similar to *IRS1*, *IRS2* is also implicated in cancer, with its overexpression linked to increased tumor growth, metastasis, and poor prognosis. It plays a crucial role in cancer cell migration and invasion (Jo *et al.*, 2024).

In summary, Figure 6 offers a concise overview of the genetic mechanisms driving ETMR, demonstrating how *Let-7* downregulation upregulates the previously discussed genes, fosters proliferative rosette formation, and ultimately leading to oncogenesis.

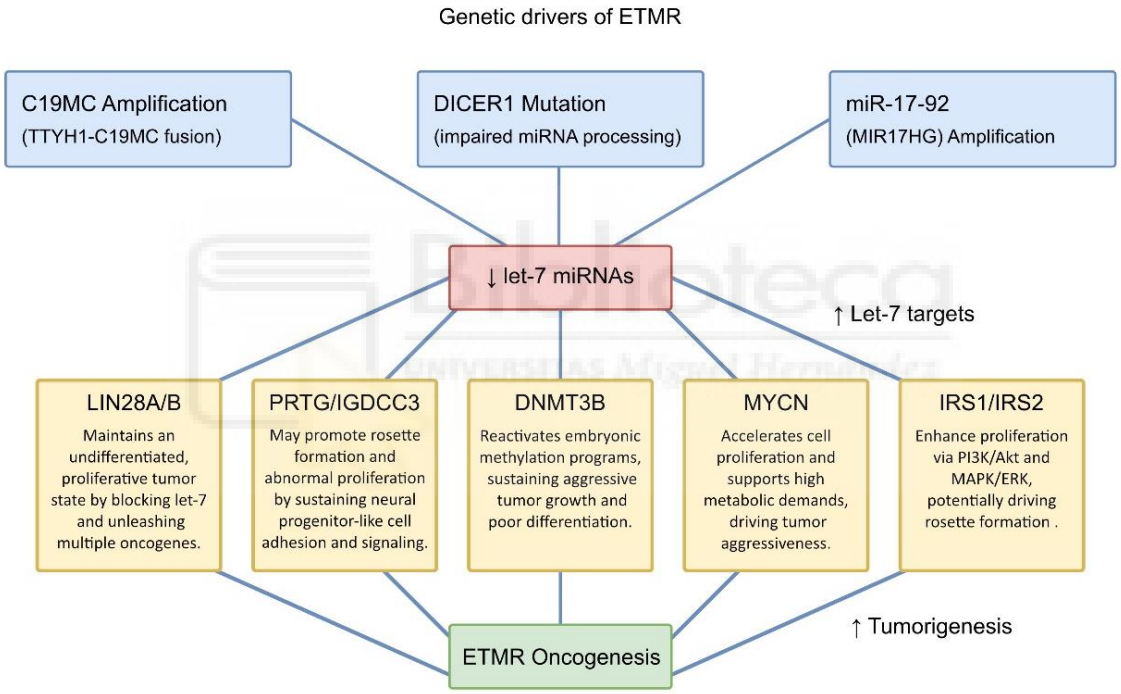


Figure 6. Summary of the ETMR Gene Network. C19MC amplification (often via the *TTYH1*-C19MC fusion), *DICER1* mutation (resulting in impaired miRNA processing), and *miR-17-92* (*MIR17HG*) amplification converge on reduced *let-7* miRNAs. Decreased *let-7* activity lifts repression on several oncogenes, including *LIN28A/B* (sustaining an undifferentiated tumor state), *PRTG/IGDCC3* (promoting rosette formation and abnormal proliferation), *DNMT3B* (reactivating embryonic methylation programs), *MYCN* (enhancing cell cycle progression and metabolic demands), and *IRS1*/*IRS2* (boosting *PI3K/Akt* and *MAPK/ERK* pathways). Together, these genetic alterations drive ETMR oncogenesis, characterized by an aggressive, poorly differentiated tumor phenotype and the distinctive multilayered rosettes seen in these pediatric brain tumors.

Following the detailed exploration of ETMR's genetic mechanisms, the *Rx-Dicer1* mouse model emerges as a promising candidate for further investigation due to its notable molecular similarities with ETMR. This model, which exhibits the formation of proliferative rosettes and shares key gene expression features with

ETMR, shows upregulation of several important genes implicated in ETMR pathogenesis, such as *Lin28a, Prtg, Greb1*, *Irs2*, and other markers associated with disrupted let-7 miRNA regulation. These similarities suggest that *Rx-Dicer1* mice could provide a valuable platform to study ETMR biology. However, further validation is required to quantify how accurately this model replicates ETMR characteristics, to assess its potential as a reliable experimental model for investigating ETMR development.

5. High-throughput study of cortical development: omics profiling of cortical progenitor cells

Cortical neurogenesis is orchestrated by neural progenitor cells whose proliferation and differentiation lay the foundation for the intricate architecture of the cerebral cortex. Recent advancements in high-throughput omics technologies have significantly expanded our understanding of these progenitor cells, revealing the molecular underpinnings of both healthy cortical development and aberrations observed in pediatric brain tumors.

In this chapter, we describe how the main omics approaches used in this thesis (bulk mRNA sequencing, single-cell RNA sequencing, and ATAC sequencing) have expanded our understanding of cortical progenitor cell biology. We highlight key discoveries emerging from these methodologies and consider how they illuminate both the evolutionary conservation and the translational potential of these findings.

5.1 Bulk mRNA sequencing

Bulk mRNA sequencing (RNA-seq) is a high-throughput method to measure the average gene expression in a bulk sample containing many cells. In bulk mRNA-seq, messenger RNA is isolated from tissue or cultured cells, converted to complementary DNA (cDNA), and then sequenced using next-generation sequencing platforms (Wang, Gerstein and Snyder, 2009). The typical workflow involves RNA extraction, mRNA enrichment or rRNA depletion, reverse transcription to cDNA, library preparation (fragmentation and addition of sequencing adapters), and high-throughput sequencing. Bioinformatics analysis then maps the sequencing reads to a reference genome to quantify transcript levels for each gene. Key analysis steps include alignment of reads, generation of a gene counts matrix, and normalization and statistical modeling to identify differentially expressed genes between conditions. Bulk RNA-seq provides a digital readout of expression, and is highly sensitive for detecting even low-abundance transcripts in a sample (Wang, Gerstein and Snyder, 2009).

Bulk RNA-seq has been widely used to chart gene expression during brain development. For example, the *BrainSpan* project generated a spatio-temporal transcriptome of the developing human brain, revealing that ~90% of analyzed genes are differentially regulated across brain regions and developmental stages (Kang *et al.*, 2011). In the embryonic and fetal cortex, bulk transcriptomic profiling

across time points has delineated major waves of gene expression that correlate with neurogenic phases. Early NECs, RGCs, basal progenitors, and newborn neurons each have characteristic gene expression signatures that can be captured by bulk RNA-seq of sorted cell populations (Mukhtar and Taylor, 2023). For instance, bulk RNA-seq of FACS-purified apical radial glia versus basal progenitors in mouse identified dozens of lineage-specific transcription factors and temporal changes in their expression, reflecting the progression from proliferative to neurogenic phases (Mukhtar *et al.*, 2022). Similarly, comparative bulk transcriptomics between species has shed light on evolutionary changes in cortical progenitors. Transcriptome comparisons of human, chimpanzee, and macaque cortices have found human-specific upregulation of genes involved in progenitor proliferation and cortical expansion (Sousa *et al.*, 2017). One striking example is the discovery of the previously mentioned human-specific gene *ARHGAP11B*, which is expressed in basal progenitors; bulk RNA analyses showed this gene is absent in non-human primates, and functional studies confirmed it can amplify progenitor pools when introduced in mice (Florio *et al.*, 2015). Thus, bulk transcriptomics has been instrumental in identifying key regulators of cortical neurogenesis and highlighting molecular innovations in the human lineage.

In neuro-oncology, bulk RNA-seq is routinely used to classify tumors and to pinpoint oncogenic pathways. Many pediatric brain tumors were initially stratified by gene expression profiling. A landmark example is medulloblastoma: bulk transcriptomic studies revealed four molecular subgroups (WNT, SHH, Group 3, Group 4) with distinct developmental signatures and outcomes (Northcott *et al.*, 2017). These subgroups were first identified via unsupervised clustering of RNA-seq data and later refined with integrated multi-omics. For instance, Group 3 medulloblastomas show an expression program with high *MYC* oncogene and photoreceptor genes, whereas Group 4 tumors have neuronal differentiation signatures (Northcott *et al.*, 2017). Such expression-based subclassification has prognostic value and informs tailored therapies. More broadly, comparative bulk transcriptomics across pediatric brain tumor types and normal brain has demonstrated that many tumors parallel specific developmental stages. For example, bulk RNA profiles of ETMR are remarkably similar to early neural progenitors, expressing stemness markers like *LIN28A/B* and *DNMT3B*; supporting the hypothesis that ETMRs arise from or freeze at an early progenitor-like state (Jessa *et al.*, 2019). In summary, bulk transcriptomic sequencing in pediatric tumors has defined molecular taxonomies and frequently traces tumor identities back to developmental cell types or aberrations of normal developmental gene networks.

5.2 Single-cell RNA sequencing

Single-cell RNA sequencing (scRNA-seq) is a transcriptomic method that profiles gene expression at the resolution of individual cells, rather than averaging over bulk tissue. This technique is crucial for resolving cellular heterogeneity. In scRNA-seq, individual cells are isolated (e.g., by microfluidics or droplet

encapsulation) and their mRNA is separately captured and barcoded, allowing cDNA from each cell to be indexed and later demultiplexed after sequencing (Macosko *et al.*, 2015). Modern scRNA-seq workflows often use droplet-based systems where thousands of single cells are partitioned into nanoliter droplets along with barcoded beads; within each droplet, the cell is lysed and its mRNAs attach to a bead carrying a unique cell barcode and molecular identifiers (UMIs) to label each transcript (Macosko *et al.*, 2015). Pooled cDNA from all droplets is then sequenced, and bioinformatic pipelines group reads by barcode to recover each cell's transcriptome. Key analysis steps include quality control (to remove low-quality cells or cell doublets), normalization of transcript counts, dimensionality reduction, and clustering of cells into transcriptomically defined groups. Marker genes for each cluster are identified to assign cell type identities (Heumos *et al.*, 2023). Unlike bulk RNA-seq, scRNA-seq can reveal discrete cell subpopulations, rare cell types, and transitional states within a complex tissue, at the cost of capturing less mRNA per cell (owing to physical and technical limitations). State-of-the-art scRNA-seq can profile tens of thousands of cells in one experiment, providing unprecedented cellular resolution of tissues.

The advent of scRNA-seq has dramatically advanced our understanding of the cellular diversity in the developing cortex. Single-cell transcriptomic atlases have been constructed for the embryonic and fetal cortex, delineating all the major progenitor and neuron types and their lineage relationships. For example, scRNA-seq of the developing mouse cortex identified distinct clusters corresponding to radial glia, intermediate basal progenitors, newborn neurons, and more mature neurons, with transitions aligning to the known histogenetic sequence (Di Bella *et al.*, 2021). In the human context, a landmark scRNA-seq study by Pollen and colleagues profiled hundreds of single cells from the mid-gestation human cortex, revealing expected cell classes but also new subtypes and temporally dynamic gene expression in progenitors (Pollen *et al.*, 2015). scRNA-seq has significantly enhanced our understanding of progenitor cell diversity during human cortical development, revealing complexities that bulk methods cannot resolve. A notable recent discovery is the identification of tripotential intermediate progenitor cells (*Tri-IPCs*), which possess the capacity to differentiate into GABAergic neurons, oligodendrocyte precursor cells, and astrocytes. This finding underscores the intricate lineage relationships among progenitor subtypes during the neurogenesis-to-gliogenesis transition (Wang *et al.*, 2025). Additionally, scRNA-seq enables trajectory analysis, allowing researchers to computationally order cells based on transcriptomic similarity and infer developmental lineages or pseudo-temporal ordering. Applied to cortical progenitors, this approach has elucidated the progression from cycling radial glia to neurogenic radial glia, intermediate progenitors, and ultimately neurons, highlighting the gradual upregulation of neuronal differentiation genes even in progenitors (Wang *et al.*, 2025). In summary, scRNA-seq provides a high-resolution census of cell types and states, identifying novel intermediate states and candidate regulators for fate transitions in cortical development.

scRNA-seq has been also transformative in cancer research by resolving intratumoral heterogeneity and identifying cell populations that bulk analysis

obscures. In pediatric brain tumors, scRNA-seq has provided direct evidence that these malignancies often contain multiple cell states, some of which recapitulate developmental progenitor programs. A compelling example comes from single-cell transcriptomic profiling of embryonal tumors and gliomas. In a pioneering study, researchers applied scRNA-seq to diffuse midline gliomas with the H3K27M mutation and found that tumor cells segregated into at least two major transcriptional states: one resembling oligodendroglial progenitor-like cells and another akin to mesenchymal lineage cells (Filbin *et al.*, 2018). Notably, the oligodendroglial-like population expressed developmental markers of oligodendrocyte precursor cells and was actively cycling, suggesting it may represent a “stem-like” compartment of the tumor, whereas the more differentiated mesenchymal-like cells expressed genes associated with astrocytic injury response (Filbin *et al.*, 2018). Importantly, scRNA-seq can identify rare but clinically significant subpopulations, such as therapy-resistant cells. For instance, in ependymoma (a pediatric glioma), scRNA-seq identified subpopulations of tumor cells with varying differentiation states and upregulation of cilium-associated genes (Wu *et al.*, 2022). Additionally, they observed enhanced interactions between malignant cells and tumor-infiltrating microglia in recurrent tumors, suggesting that immune cell crosstalk may contribute to tumor relapses (Wu *et al.*, 2022). Moreover, integrating single-cell data from tumors with developmental atlases has allowed researchers to map tumors to their developmental origins. A study projected single-cell profiles of various pediatric brain tumors onto a reference atlas of normal developing mouse brain, and they found that *WNT* medulloblastomas transcriptionally align with rhombic lip neural progenitors, while Group 3 medulloblastomas align with earlier prenatal progenitors, providing a “developmental context” for each tumor subtype (Vladoiu *et al.*, 2019). This approach confirmed that many pediatric tumors truly are aberrant executions of developmental programs, also suggested by Jessa et al. (Jessa *et al.*, 2019). Moving forward, single-cell transcriptomics is increasingly used to track how tumors respond to therapy at a cellular level and to design differentiation therapies that push cancer cells towards a non-proliferative fate (Danielli *et al.*, 2024). In summary, single-cell RNA-seq in pediatric brain tumors reveals intratumoral heterogeneity and uncovers how tumors contain cell types mirroring different stages of development. By identifying tumor subpopulations and their proportions, scRNA-seq also offers insights into tumor dynamics and potential therapeutic targets.

5.3 ATAC sequencing

Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq) is a rapid and sensitive method to profile open chromatin regions across the genome (Buenrostro *et al.*, 2013). Regions of open or accessible chromatin (lacking nucleosomes) are typically regulatory DNA elements such as promoters and enhancers. In ATAC-seq, nuclei from cells are incubated with a hyperactive Tn5 transposase loaded with sequencing adapters. The transposase preferentially inserts the adapters into open chromatin regions. Thus, after the reaction, DNA fragments from nucleosome-free regions are tagged with adapters. These

fragments (~50 bp to a few hundred bp in length) are then PCR-amplified and sequenced. The output is essentially a genome-wide map of open chromatin, often visualized as a “peak” profile. There are several steps to include, as cell lysis to yield nuclei, Tn5 transposition (which simultaneously fragments DNA and inserts adapters), purification of transposed DNA, PCR, and sequencing (Buenrostro *et al.*, 2013). Bioinformatics involves a multi-step workflow, including read quality control, adapter trimming, alignment to the reference genome, post-alignment processing (such as duplicate removal and Tn5 insertion offset correction), peak calling to identify accessible chromatin regions, generation of normalized signal tracks for visualization, differential accessibility analysis, nucleosome positioning, and transcription factor footprinting (small dips in accessibility where a TF binds and protects DNA) (Hitz *et al.*, 2023).

ATAC-seq has been used to chart developmental changes in the chromatin landscape of cortical progenitors and neurons, shedding light on gene regulatory dynamics that drive neurogenesis. For instance, during the transition of radial glia into neurons, ATAC-seq reveals a global opening of neuron-specific enhancers and closing of progenitor-specific enhancers (de la Torre-Ubieta *et al.*, 2018; Ziffra *et al.*, 2021; Jiang *et al.*, 2023). An ATAC-seq time-course in human cortical development showed that early in corticogenesis, chromatin accessibility is high at promoters of stemness and cell-cycle genes, whereas later (during neurogenesis) accessibility increases at neuronal differentiation genes and synaptic protein gene enhancers (de la Torre-Ubieta *et al.*, 2018). Such developmental epigenomic profiling has identified key temporal regulators in cortical progenitor cells. For example, an accessible enhancer near the *FGFR2* gene emerges during early corticogenesis, promoting *FGFR2* expression and thereby stimulating cortical progenitor cell proliferation. This epigenetic change supports the expansion of the progenitor pool necessary for proper cortical development (de la Torre-Ubieta *et al.*, 2018). In another foundational study, bulk ATAC-seq was performed on dissected regions of the developing human cerebral cortex, including the germinal and cortical plate zones, enabling the identification of region-specific and stage-specific patterns of open chromatin (Ziffra *et al.*, 2021). By integrating ATAC-seq with transcriptomic and chromatin interaction datasets, the researchers mapped distal regulatory elements to their target genes and uncovered regulatory programs underlying neuronal differentiation and areal identity (Ziffra *et al.*, 2021). Through such analyses, bulk ATAC-seq has proven invaluable for decoding the epigenomic architecture of the developing brain, revealing conserved and divergent features across regions, cell states, and model systems.

Many pediatric brain tumors are driven not by only by DNA mutations but by epigenetic alterations. For instance, mutations in histone-encoding genes or chromatin modifiers, or DNA methylation aberrations. ATAC-seq is a powerful tool to interrogate the regulatory landscape of tumor cells and compare it to normal developmental counterparts. One key application has been identifying super enhancers or other regulatory elements that tumors abnormally activate to drive oncogene expression. For example, ATAC-seq on patient-derived diffuse intrinsic pontine glioma cells showed that the mutant H3K27M histone leads to a

global increase in chromatin accessibility at normally Polycomb-repressed developmental genes (Lewis *et al.*, 2022). Another study also found that H3K27M glioma cells also have aberrantly open chromatin at neural lineage genes and bivalent developmental gene promoters, corresponding to the drastic reduction in H3K27me3 repression caused by the mutation (Nagaraja *et al.*, 2019). These epigenomic changes help explain why these tumors inappropriately express developmental programs. Recent integrative studies further highlight the importance of ATAC-seq in mapping tumor-specific regulatory alterations. For example, Wang et al. (2021) combined ATAC-seq with RNA-seq, ChIP-seq, and Hi-C to characterize the chromatin landscape of pediatric high-grade gliomas. They identified tumor-specific enhancers and gene co-amplification events involving oncogenes such as *JAG2* and *FLRT1*, and demonstrated that structural changes in the 3D genome contribute to enhancer hijacking and dysregulated gene expression (Wang *et al.*, 2021). Similarly, Sin-Chan et al. (2019) used ATAC-seq and chromatin profiling to uncover a hijacked super-enhancer network driving a C19MC-LIN28A-MYCN oncogenic circuit in ETMRs. These epigenetic alterations create feedforward loops that entrap early neural lineage programs and sustain tumor growth, vulnerabilities that can be targeted using bromodomain inhibitors (Sin-Chan *et al.*, 2019). Together, these studies underscore how ATAC-seq can illuminate tumor-specific regulatory mechanisms, reveal the developmental context of pediatric brain tumors, and guide therapeutic strategies.

In conclusion, these complementary omics approaches provide a multi-layered view of cortical development, bridging transcriptional and epigenetic landscapes at both population and single-cell resolution. By integrating gene expression data with chromatin accessibility profiles, we can begin to reconstruct the regulatory logic guiding progenitor cell fate decisions and understand how deviations from these programs may lead to malignancy.

6. Membrane-bound transcription factors (MTFs)

Membrane-bound transcription factors (MTFs) constitute a rare, non-canonical class of regulators that begin life as transmembrane proteins and only acquire DNA-binding activity after stimulus-dependent proteolytic release of their cytoplasmic domains (Seo, Kim and Park, 2008; Liu *et al.*, 2018; Yang *et al.*, 2020). Although a handful of mammalian examples, such as ATF6, XBP1, MYRF and the sterol-sensing SREBFs, have revealed how MTFs can couple ER stress or lipid scarcity to nuclear gene programs, the overall abundance, evolutionary conservation and developmental relevance of MTFs remain almost unexplored (Liu *et al.*, 2018). Given that membrane receptors like ROBO1 and DLL1 already play pivotal roles in cortical neurogenesis, we asked whether additional, previously unrecognized MTFs might provide a direct transcriptional conduct between extracellular cues and progenitor fate decisions. To address this possibility, we applied the deep-learning predictor *DeepTFactor* (Kim *et al.*, 2021) to the entire hu-

man and mouse transmembrane proteomes, with the aim of compiling a systematic catalogue of candidate mammalian MTFs and prioritizing those most likely to influence cortical development.

OBJECTIVES

Objective 1. Investigate the transcriptomic and epigenomic mechanisms that define different cortical neurogenesis modes across amniotes.

Objective 2. Identify and characterize mammalian membrane-bound transcription factors.

Objective 3. Evaluate transcriptomically the *Rx-Dicer1* mutant mouse as a model for ETMR.

Objective 4. Elucidate the transcriptomic and cellular mechanisms driving ETMR tumor initiation and expansion.

MATERIALS AND METHODS

1. Evolutionary analysis

1.1 Isolation of cells

- aRGC isolation for RNA-seq and ATAC-seq analyses

aRGCs were isolated from mouse (Neocortex E12, Neocortex E14, Olfactory Bulb E12), chick (Lateral and Medial Dorsal Pallium E4), and snake (all Pallial Regions E6). Cells were purified by fluorescence-activated cell sorting (FACS) using *Pard3* (PAR3), a canonical membrane marker for RGCs.

1.2 Transcriptomic analyses

- RNA-seq analyses of PAR3+ FACS-sorted cells from mouse, chick, and snake

RNA-seq reads were quality-checked with *FASTQC* v2.6.14 (Andrews, 2010) to assess overall sequencing quality and subsequently aligned to the *GRCm38.p6* (*Mus musculus*), *GRCg7b* (*Gallus gallus*) or *HLlamFul1* (*Boaedon fuliginosus*) genome assembly using *HISAT2* v2.1.0 (Kim, Langmead and Salzberg, 2015). Aligned reads were assigned to genes using *HTSeq* v2.0.2 with parameters -f bam -r pos -s no -m union --additional-attr=gene_name (Anders, Pyl and Huber, 2015). Snake genome assembly and annotation was created and annotated by Michael Hiller Lab.

Gene-level read counts were normalized and analyzed using *DESeq2* (Love, Huber and Anders, 2014) to identify differentially expressed genes (DEGs). DEGs were defined based on an adjusted p-value threshold of Padj < 0.05.

- Combined DEG analysis

To compile a unified list of DEGs associated with neurogenesis modes, a combined statistical analysis was performed. P-adjusted values from the two pairwise comparisons (NCx E12 vs. OB E12 and NCx E12 vs. NCx E14) were integrated using Fisher's method to calculate a combined p-adjusted value (Fisher, 1932). Only genes significantly differentially expressed in both comparisons were retained in the final DEG list. For each gene, the mean log2 fold change (Log2FC) across the two comparisons was calculated and used for ranking. The final list was sorted based on the combined p-adjusted values. Volcano plot was generated to visualize the DEGs between indirect neurogenesis (IN) and direct neurogenesis (DN) conditions combined.

- Gene Ontology, pathway enrichment analysis and dot plots

Functional enrichment analysis for DEGs was performed using Gene Ontology (GO) and pathway enrichment tools. GO terms were identified with *DAVID* v2021

(Sherman *et al.*, 2022) and *clusterProfiler* v4.10.0 (Yu *et al.*, 2012). Enriched GO terms were visualized using dot plots, where the x-axis represents -log10(FDR), the dot size indicates the number of genes contributing to the pathway, and the color gradient reflects fold enrichment. The dot plots were produced using *ggplot2* (Wickham, 2016) (v3.5.1), *ggrepel* (v0.9.5) (Slowikowski, 2024), and *RColorBrewer* (v1.1-3), and Gene Set Enrichment Analysis (GSEA) plots were produced using *clusterProfiler* v4.10.0 (Yu *et al.*, 2012).

- High-confidence orthologous genes across species
  - Orthology analysis and selection criteria

To identify high-confidence orthologous genes shared between mouse and chick that are associated with IN-enriched and DN-enriched samples, orthology data were extracted from the Ensembl criteria (Harrison *et al.*, 2024). The analysis focused on genes with high-confidence orthologous relationships based on the following quality control metrics provided by Ensembl:

1. Gene Order Conservation (GOC) Score: Measures the conservation of gene neighbors between orthologous pairs. Each orthologous pair was assigned a GOC score based on how many of the four closest neighbors were conserved, with a maximum score of 100%.
2. Whole Genome Alignment (WGA) Score: Evaluates the alignment coverage of orthologous gene pairs, with higher weights given to exon coverage compared to intronic regions.
3. Percentage Identity: Reflects the sequence similarity between orthologous genes.

To ensure robust identification of orthologs, the following thresholds were applied as per Ensembl's criteria for the Mammalia and Aves clades: Minimum GOC score: 75; Minimum WGA score: 75; Minimum sequence identity: 50%. Orthologous pairs meeting or exceeding these thresholds were classified as high-confidence orthologs.

  - Gene selection for IN-enriched and DN-enriched conditions

Orthologous genes associated with IN-enriched samples (NCx E12 in mouse, Lateral E4 in chick) and DN-enriched samples (OB E12 or NCx E14 in mouse, Medial E4 in chick) were identified by comparing RNA-seq datasets. DEGs in mouse were determined using *DESeq2* (as described above) with an adjusted p-value threshold of Padj < 0.05, and orthologous chick genes were identified based on expression patterns from chick transcriptomic datasets.

To create unified lists of IN- and DN-associated orthologs, p-values from multiple pairwise comparisons were combined using Fisher's method to compute a single -log10(FDR) score for each gene. Genes were ranked by their combined

significance and included in the final list if they were significantly differentially expressed in all IN-enriched or DN-enriched comparisons across both species.

- Data visualization

Bar plots of high-confidence orthologous genes for IN-enriched and DN-enriched conditions were generated using *ggplot2* v3.5.1 (Wickham, 2016), ranking genes based on their combined -log10(FDR) values.

To compare expression levels of RGC-related transcripts across chick, mouse, and snake, normalized read counts (obtained from *DESeq2*) were plotted as box plots for each embryonic stage and neurogenesis condition. Specifically, IN-enriched stages were colored in red, while DN-enriched stages were colored in blue. Box plots were generated in R (v4.2.0) using *ggplot2* (Wickham, 2016) v3.5.1.

- RNA-seq analysis of aRGCs modified with *Robo1/2* overexpression and *Delta1* knockout
    - Experimental design and RNA-seq sample preparation

To investigate the transcriptomic changes induced by the overexpression of *Robo1, Robo2*, and CRISPR-mediated *Delta1* knockout (*crDll1*) in RGCs of the developing mouse neocortex (NCx), in utero electroporation was performed at embryonic day 12.5 (E12.5). Plasmids encoding GFP, *Robo1, Robo2*, and *crDll1* were electroporated into the VZ of the NCx. Electroporated embryos were harvested 24 hours later (E13.5), and GFP-positive (GFP+) cells were isolated from the VZ) and SVZ via *FACS*. Total RNA was extracted from the sorted cells for bulk RNA-seq analysis.

- RNA-seq data processing

RNA-seq reads were quality-checked using *FASTQC* v0.11.9 (Andrews, 2010) to assess sequencing quality metrics. Reads were aligned to the *GRCm38.p6* genome assembly using *HISAT2* v2.1.0 (Kim, Langmead and Salzberg, 2015) with default parameters. Gene-level read counts were quantified using *HTSeq* v0.11.1 (Anders, Pyl and Huber, 2015) with the following parameters: -m union –no_stranded, using Ensembl gene annotations. PCA was performed to assess global transcriptomic differences, revealing clear clustering of samples based on experimental conditions.

- Differential gene expression analysis

DEGs were identified using *DESeq2* v1.28.1 (Love, Huber and Anders, 2014). Genes with an adjusted p-value of less than 0.05 were considered significantly differentially expressed. DEGs were visualized using heatmaps and volcano plots

to highlight the upregulation and downregulation of key genes. All plots were generated using *ggplot2* v3.5.1 (Wickham, 2016).

- Functional enrichment analysis

Functional annotation of DEGs was performed using GO enrichment analysis and GSEA. GO and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment were conducted using the *clusterProfiler* v4.10.0 package. Enrichment plots for GSEA were generated using *clusterProfiler* and visualized with *ggplot2* (Wickham, 2016).

- Cux2 single cell RNA-seq analyses and RShiny appliaction

The methods for analyzing the scRNA-seq data from public databases used in this study are detailed in their corresponding publications (Nano *et al.*, 2023; Del-Valle-Anton *et al.*, 2024).

The data from the paper Del-Valle-Anton *et al.*, 2024 was integrated into an *RShiny* application developed using the *ShinyCell* template (Ouyang *et al.*, 2021).

1.3 Epigenomic and regulatory analyses

- ATAC-seq analyses of PAR3+ FACS-sorted cells from mouse, chick, and Snake

  - Preprocessing and peak calling

The ATAC-seq raw data were processed following the ENCODE ATAC-seq Data Standards and Processing Pipeline (Hitz *et al.*, 2023) to ensure reproducibility and quality control.

Peaks of accessible chromatin regions were identified using the *MACS2* peak caller (Zhang *et al.*, 2008), which calculates enrichment of sequencing reads over background signal. Peak calling was performed for each replicate under treatment and control conditions separately, resulting in narrowPeak files. The peaks were subsequently converted into GRanges objects for compatibility with downstream analysis in R. To improve data reliability, a blacklist of genomic regions prone to artifacts from ENCODE for *mm10* mouse genome was applied to exclude problematic regions (Amemiya, Kundaje and Boyle, 2019). The resulting data were used to generate BAM files for subsequent analyses.

  - Differential accessibility analysis with *csaw*

The differential accessibility analysis was conducted using the *csaw* workflow in R (Lun and Smyth, 2016). This approach leverages statistical models to compare

read counts in predefined or de novo windows across experimental conditions. The workflow is outlined below:

- Peak import and GRanges conversion

Replicate peak files (narrowPeak format) were read into R and converted into GRanges objects. For each condition, peaks were combined using union operations across replicates to create condition-specific peak sets. A final consensus peak set was derived by merging peak sets from all conditions.

- Read counting and normalization:

Read counts were calculated for each consensus peak using *regionCounts*(). BAM files for each replicate were used to quantify the number of reads overlapping each peak. Peaks with low abundance were filtered out based on average logCPM thresholds (> -3) to remove noise.

The filtered peak counts were normalized using the loess-based normalization method in *csaw* (Lun and Smyth, 2016). This approach adjusts for potential biases in sequencing coverage across samples and minimizes global systematic differences. Loess normalization offsets were applied using *normOffsets(type = "loess")*.

- Window-based read counting

Sliding windows of 500 bp (adjustable) were defined across the genome to identify regions of local enrichment. Reads overlapping these windows were counted using *windowCounts()*. A local background estimator was applied to define enrichment relative to neighboring windows (2-kb neighborhood).

- Differential testing

A design matrix was constructed to model the experimental groups (e.g., mouse NCx E12 vs. NCx E14). Dispersion estimates were stabilized using empirical Bayes, and generalized linear models were fitted using *glmQLFit*() from *edgeR* (Robinson, McCarthy and Smyth, 2010). Differentially accessible regions (DARs) were identified using *glmQLFTest*(), with significance determined by FDR thresholds (< 0.05).

- Merging windows and visualization

Significant windows were merged into broader regions using *mergeWindows*(), ensuring contiguous regions were analyzed together. The resulting merged regions were annotated for gene proximity using *ChIPseeker* (Yu, Wang and He,

2015). Visualization of results was showed with volcano plots, highlighting significant peaks of differential chromatin accessibility using *ggplot2* (Wickham, 2016). *ChIPseeker* was also used to annotate DARs (5',3', Distal Intergenic > 3 kb, Promoter < 3kb, Intron, Exon).

- Motif analysis with MEME suite

To identify *SALL1* motifs in the DARs, we employed the MEME Suite (Bailey *et al.*, 2015). The analysis consisted of three major steps: (1) de novo motif discovery with *MEME* tool, (2) motif comparison with *TOMTOM*, and (3) motif scanning with *FIMO*.

- *MEME* tool analysis

We first performed de novo motif discovery using the *MEME* algorithm (Bailey and Elkan, 1994). Sequences corresponding to each DAR were extracted from the *mm10* mouse reference genome and provided as input to *MEME*. We configured *MEME* to allow zero or one occurrence of a motif per sequence (ZOOPS) and applied a default background model to reduce spurious hits. *MEME* iteratively identifies overrepresented sequence patterns and ranks them based on their E-values. From the motifs discovered, we selected one that appeared consistently across all comparisons for further analyses.

- Motif comparison with *TOMTOM*:

We next used *TOMTOM* (Gupta *et al.*, 2007) to compare the selected motif against known transcription factor motifs in JASPAR 2024 (Rauluseviciute *et al.*, 2024) and HOCOMOCO v12 (Vorontsov *et al.*, 2024). Among the candidates identified, *SALL1* emerged as statistically significant in all *TOMTOM* comparisons. In addition to its strong motif match, *SALL1* showed differential expression across species and is known to regulate modes of neurogenesis (Harrison *et al.*, 2012), making it a compelling candidate for further investigation.

- *FIMO* analysis:

To find individual *SALL1* binding sites in the mouse genome, we employed *FIMO* (Find Individual Motif Occurrences) (Grant, Bailey and Noble, 2011). Specifically, we used the SALL1.H12CORE.0.P.C position weight matrix (PWM) from the HOCOMOCO v12 database (Vorontsov *et al.*, 2024). *FIMO* computes a log-likelihood ratio for each genomic position to assess its similarity to the motif and provides a p-value for statistical significance. We applied a stringent threshold (e.g., p < 1e-4) to minimize false positives. The resulting output included genomic coordinates for predicted *SALL1* binding sites, along with corresponding scores and statistical measures. Finally, we filtered and formatted these predictions to

retain only high-confidence sites, which were used in subsequent downstream analyses.

- STRING-based protein interaction network analysis

Genes associated with DARs, DEGs in each region, and predicted to have cis-regulatory elements containing Sall1-binding motifs (*FIMO* results) were subjected to protein-protein interaction analysis using the STRING database (v12.0) (Szklarczyk, Kirsch, Koutrouli, Nastou, Mehryary, Hachilif, Annika L Gable, *et al.*, 2023). STRING integrates multiple sources of interaction data, including experimental datasets, computational predictions, text mining, co-expression data, and curated pathway information, ensuring comprehensive coverage. Each interaction was scored for confidence and functional relevance, and only interactions with a combined confidence score ≥0.7 (high confidence) were considered to ensure reliability. Gene lists for IN-enriched and DN-enriched regions were uploaded to the STRING web interface, where functional enrichment analysis was performed using STRING's enrichment module. The resulting networks were visualized with nodes representing proteins and edges denoting functional or physical interactions, with line colors indicating the sources of evidence for each interaction.

1.4 Genomic analyses

- *MIR3607* regulatory landscapes across gyrencephalic and lissencephalic species

The genomic regulatory landscape of *MIR3607* (*SNORD138*) was analyzed using publicly available data from the ENCODE project (Luo *et al.*, 2020). Candidate cis-regulatory elements (cCREs) in the human genome (*GRCh38/hg38*) were identified using the *SCREEN* tool (Moore *et al.*, 2020), which integrates chromatin accessibility, histone modifications, and TF binding data. cCREs were categorized into promoter-like, proximal enhancer-like, and distal enhancer-like regions based on ENCODE-defined signatures, and these regions were color-coded for visualization.

To investigate conservation, the genomic region encompassing -5 kb upstream to +1 kb downstream of the *MIR3607* transcription start site (TSS) was analyzed across 18 species. These species were divided into gyrencephalic (e.g., humans, chimpanzees, pigs) and lissencephalic (e.g., mice, rats, guinea pigs) groups based on cortical folding patterns. Multiple sequence alignments were performed to identify conserved regions, which were highlighted in yellow boxes. Sequence conservation and divergence patterns were visualized as percent sequence identity, plotted below the alignment for each group (red for gyrencephalic, green for lissencephalic species). *Jalview* was utilized for the visualization and analysis of multiple sequence alignments (Waterhouse *et al.*, 2009).

Putative transcription factor (TF) binding sites within the conserved regions were identified using the MEME Suite (Bailey *et al.*, 2015). First, motif discovery was performed with the *MEME* algorithm, which scanned the conserved sequences for statistically enriched motifs without prior motif definitions. The analysis was conducted using default parameters with a focus on identifying both short and medium-length motifs, commonly associated with TF binding (Bailey and Elkan, 1994). To predict which transcription factors might bind to the discovered motifs, we used *TOMTOM*, a motif comparison tool within the MEME Suite. *TOMTOM* compares the input motifs against a database of known TF binding motifs (e.g., JASPAR, HOCOMOCO) and returns likely TF matches along with statistical confidence scores (Gupta *et al.*, 2007).

- RNA secondary structure prediction of pre-*MIR3607* in humans and mice

The secondary structure of pre-*MIR3607* was predicted and compared between humans (*Homo sapiens*) and mice (*Mus musculus*) to investigate how sequence differences may impact miRNA maturation. Accurate predictions of RNA secondary structures were achieved using *MXfold2*, an advanced algorithm that combines deep learning with Turner's nearest-neighbor free energy parameters (Sato, Akiyama and Sakakibara, 2021).

2. Proteomics and membrane transcription factors

2.1 Tools and data sources

The amino acid sequences of mouse and human proteins were downloaded from UniProt (Release 2022_05) (The UniProt Consortium, 2024). In silico prediction of TF activity in transmembrane proteins was performed using *DeepTFactor* (Kim *et al.*, 2021), and feature attribution analysis was conducted with *Captum* v0.5.0 (Kokhlikyan *et al.*, 2020). Subcellular localization data for selected canonical TFs and newly predicted membrane-bound TF (MTF) candidates were obtained from the Human Protein Atlas (HPA) v21.1 (Thul *et al.*, 2017).

To identify canonical MTFs in humans and mice, we selected proteins from UniProt meeting the following criteria:

- Reviewed (*Swiss-Prot*)
- Containing an annotated cytoplasmic domain
- Bearing the GO term GO:0003700 ("DNA-binding transcription factor activity")

The final intersection of canonical MTFs across human and mouse included 13 (ATF6, ATF6B, CREB3, CREB3L1, CREB3L2, CREB3L3, CREB3L4, JPH2, MYRF, PLSCR1, SREBF1, SREBF2, and XBP1).

2.2 Structural annotation of transmembrane proteins

To explore novel MTFs in mammals, we focused on 14,160 transmembrane proteins (combined total for human and mouse) annotated in UniProt (The UniProt Consortium, 2021). Each transmembrane protein was segmented into distinct domains according to UniProt annotations:

- Cytoplasmic Domain (CD): Intracellular side facing the cytoplasm.
- Extracellular Domain (ED): Facing the extracellular space.
- Lumenal Domain (LD): Within the lumen of intracellular organelles (e.g., ER, Golgi).
- Transmembrane Domain (TMD): Hydrophobic region(s) spanning the lipid bilayer.

Proteins were categorized as either single-pass or multi-pass, depending on the number of TMDs, with multi-pass proteins having alternating intracellular and extracellular (or lumenal) loops.

Because cleavage at the TMD can release a cytoplasmic fragment capable of entering the nucleus (Lichtenthaler, Lemberg and Fluhrer, 2018), we examined how full, half, or no TMD inclusion affected TF prediction scores.

2.3 Data filtering and retrieval

1. Filtering in UniProt:
    - Species set to “Human” or “Mouse.”
    - Domain annotated as “cytoplasmic,” “extracellular,” or “lumenal.”
    - Reviewed sequences (*Swiss-Prot*) only.
2. Extraction of Protein Domains:
    - For each protein, the coordinates of annotated cytoplasmic, extracellular, and lumenal domains were extracted.
    - Where multiple TMDs existed, loop domains were individually parsed.
    - For each domain, three conditions were generated: (i) without TMD, (ii) with full TMD, (iii) with half TMD.
3. Sequence Retrieval:
    - The corresponding FASTA sequences were downloaded using *UniProt’s Retrieve/ID* mapping tool.

2.4 MTF prediction with *DeepTFactor*

All domain variants (no TMD, half TMD, full TMD) were fed into *DeepTFactor* (Kim et al., 2021), which assigns a TF prediction score between 0 and 1. For downstream analyses, scores ≥ 0.5 were arbitrarily considered “putative MTFs”.

The final list of putative MTFs was annotated by gene name using UniProt *Retrieve/ID mapping*.

2.5 Domain-level analysis of validated MTFs

A detailed domain-level validation was performed on four experimentally characterized MTFs—human XBP1, ATF6A, MYRF, and PLSCR1. In these proteins, the *DeepTFactor* score corresponds to the probability that each transmembrane domain configuration functions as a TF. Experimental studies (Dj *et al.*, 2002; Zhou *et al.*, 2005; Cubillos-Ruiz *et al.*, 2015; Choi *et al.*, 2018; Tsai and Lee, 2018; Fan *et al.*, 2021) have demonstrated that removing or mutating specific domains—such as the DNA-binding domain (DBD), the transactivation domain (TD), or palmitoylation motifs (PM)—abolishes TF activity, so we use these experiments to create in-silico validations of our predictions.

2.6 Human Protein Atlas immunostaining

The analysis of the subcellular localization of canonical TFs, cytoplasmic proteins, and new candidate proteins identified in our study was performed on confocal image datasets downloaded from the Human Protein Atlas public database. Images analyzed were from single confocal planes, each representing a single optical section of the cells. Cell lines analyzed were: SOX2 [Antibody: HPA045725, Cell Line: U-251 MG], ZIC1 [Antibody: HPA004098, Cell Line: AF22], ITGB2 [Antibody: HPA016894, Cell Line: EFO-21], PABPC1 [Antibody: HPA045423, Cell Line: U-2 OS], ADGRG5 [Antibody: HPA007133, Cell Line: CACO-2], ERAP1 [Antibody: HPA042317, Cell Line: A-431], FNDC4 [Antibody: HPA063581, Cell line: SH-SY5Y], NGFR [Antibody: HPA004765, Cell Line: A-431], PCDHGB4 [Antibody: HPA010580, Cell line: SK-MEL-30], RTP4 [Antibody: HPA064887, Cell line: SK-MEL-30], SCNN1G [Antibody: HPA071194, Cell Line: RT4], TMEM131L [Antibody: HPA043472, Cell Line: U-2 OS], TNFRSF9 [Antibody: HPA071425, Cell Line: U-2 OS], CACNA1I [Antibody: HPA066992, Cell Line: MCF7], EPHB1 [Antibody: HPA067740, Cell line: U-251 MG], KCNQ1 [Antibody: HPA071107, Cell line: OE19], LSR [Antibody: HPA075309, Cell Line: U-2 OS], PCDH18 [Antibody: HPA017976, Cell Line: U-2 OS], SLC5A13 [Antibody: HPA036449, Cell Line: SK-MEL-30]. We only analyzed immunofluorescence images obtained using antibodies that recognize the protein domain predicted as MTF.

3. Analysis of *Dicer1* mutant and ETMR

3.1 Microdissections

- Bulk mRNA and miRNA sequencing

Bulk mRNA sequencing was performed using E17.5 and E11.5 mouse brains. Under sterile conditions, tissue was microdissected into defined cortical and septal/ventral regions, then snap-frozen for RNA isolation. After verifying genotypes, samples were pooled by sex, genotype, and brain region. Only RNA preparations with a RIN above 8 were chosen for library preparation. Poly(A)-selected mRNA libraries were constructed for each pooled sample and sequenced on a NextSeq 2000 platform, producing approximately 30–40 million reads per library.

- Single cell RNA sequencing

For single-cell RNA sequencing, E17.5 and E12.5 brains were microdissected into cortical and septal regions under sterile conditions (only E12 NCx was analyze due to time restrictions of the thesis period). Following parallel genotyping, fresh tissue from wild-type and mutant animals was pooled. The samples were enzymatically and mechanically dissociated to obtain single-cell suspensions. After flow sorting, a defined number of cells (17,000) was loaded onto a *10X Genomics* chip following the manufacturer's instructions for single-cell library preparation, enabling detailed single-cell transcriptomic profiling.

3.2 Transcriptomic analyses

- RNA-seq of *Rx-Cre-Dicer*$^{F/F}$ mutant mice
    - Data processing, differential expression analysis, and visualization of DEGs

RNA-seq reads were quality-checked using *FASTQC* v2.6.14 (Andrews, 2010) and aligned to the *GRCm39* genome assembly using *HISAT2* v2.1.0 (Kim, Langmead and Salzberg, 2015). Gene-level read counts were computed with *HTSeq* v2.0.2, using the parameters -f bam -r pos -s no -m union --additional-attr=gene_name (Anders, Pyl and Huber, 2015). DEGs were identified using *DESeq2* (Love, Huber and Anders, 2014), with significance thresholds set at adjusted p-value (Padj < 0.05). To eliminate potential biases caused by sex-specific expression differences, genes associated with sexual dimorphism (*Erdr1, Erdr1_1, G530011O06Rik, Kdm6a, Kdm5c, Eif2s3x, Uty, Xist, Ddx3y, Eif2s3y, Kdm5d, Ppp1r1b*) were excluded from the analysis. Additionally, in both developmental stages multi-factor designs were applied in *DESeq2* to account for the effects of Zone (NCx-Septum) and Sex (Male-Female) to more accurately identify DEGs while controlling for these variables.

Volcano plots were generated to display DEGs identified in *Rx-Cre-Dicer*$^{F/F}$ *(Rx-Dicer1)* mutants at both E11.5 and E17.5. Upregulated genes are shown in red, while downregulated genes are shown in blue. Genes highlighted in dark red or dark blue represent match the ETMR (embryonal tumor with multilayered rosettes) gene signature. Genes marked with an asterisk (*) are targets of the *let-7* miRNA family using *miRBD* (Liu and Wang, 2019; Chen and Wang, 2020).

- Gene Ontology and pathway enrichment analysis

Functional enrichment analysis for DEGs was conducted using GO. This analysis was performed using the *DAVID* online tool (v2021) (Sherman *et al.*, 2022) and *clusterProfiler* (v4.10.0) (Yu *et al.*, 2012) for both E11.5 and E17.5. Dot plots were used to visualize these enrichment results. The x-axis represents -log10(FDR), while the size of each dot indicates the number of genes contributing to the pathway. The color gradient reflects fold enrichment. Dot plots were generated using *ggplot2* (v3.5.1) (Wickham, 2016) and *RColorBrewer* (v1.1-3).

- Gene set enrichment analysis plots and dot plots

GSEA method was used for enrichment analysis of the DEGs in *Rx-Dicer1* at both ages (Subramanian *et al.*, 2005) using *MSigDB* v2023 and default conditions. The dot plots were produced using *ggplot2* (v3.5.1) (Wickham, 2016), *ggrepel* (v0.9.5) (Slowikowski, 2024), and *RColorBrewer* (v1.1-3), and GSEA plots were produced using *clusterProfiler* v4.10.0 (Yu *et al.*, 2012).

GSEA enrichment plots with multiple terms were created to display enrichment scores (ES) along ranked DEGs. The x-axis shows ranked genes based on Log2FC, while the y-axis represents cumulative ES. Vertical bars indicate gene set members. GSEA plots were generated using *clusterProfiler* (Yu *et al.*, 2012) and styled with *ggplot2* (Wickham, 2016). The header terms for each of the five selected biological pathways were annotated manually to ensure clarity and relevance to the biological context of the analysis.

- Brain tumor signature creation

Brain tumor gene signatures were established using bulk RNA-seq data published by Jessa et al. (2019), encompassing a wide range of pediatric and adult brain tumor types (Jessa *et al.*, 2019). These included HGNET (high-grade neuroepithelial tumor with *BCOR* alteration), ETMR (embryonal tumor with multilayered rosettes), Other_EBT (other embryonal brain tumors), MB (medulloblastoma), ATRT (atypical teratoid/rhabdoid tumor), and various molecular subtypes of HGG (high-grade gliomas), such as HGG-IDH (IDH-mutant high-grade gliomas), HGG-WT (wild-type high-grade gliomas without histone or IDH mutations), HGG H3K27M pontine, HGG H3K27M non-pontine, and other HGG subtypes. Additionally, LGG (low-grade gliomas) subtypes 1 and 2 were included. The raw RNA-seq data were reanalyzed using the same standardized pipeline applied to the *Rx-Dicer1* muntat bulk RNA-seq samples, aligning to the *GRCh38.p14* human genome assembly and using consistent parameters for accurate cross-comparison.

Two differential expression analyses were conducted for each tumor type:

1. Tumor vs healthy control samples, to identify DEGs associated with tumor biology.
2. Tumor vs all other tumor types, to identify tumor-specific DEGs defining the unique molecular signature of each tumor type.

These analyses provided two sets of DEGs for each tumor type: (1) DEGs relative to healthy tissue and (2) DEGs distinguishing each tumor from others. To integrate the 2nd comparison into a unified gene signature, p-adjusted values from all comparisons were combined using Fisher's test, ensuring statistical robustness, while Log2FC values were averaged across comparisons. Genes were then ranked based on ascending p-adjusted values to prioritize those most strongly associated with the respective tumor type. This process was conducted separately for upregulated and downregulated DEGs, generating distinct signatures for each.

The resulting ETMR gene signature (consisting of both upregulated and downregulated genes) was applied to annotate DEGs in the *Rx-Dicer1* mutant volcano plots. Genes present in the ETMR signature were highlighted in dark red (upregulated) or dark blue (downregulated).

- GBS mouse model analyses

Data from the GBS mouse model, an ETMR model published by Neumann et al., 2017, was reanalyzed to further investigate gene expression profiles (Neumann *et al.*, 2017). The publicly available dataset was retrieved from the Gene Expression Omnibus (GEO) and analyzed using *GEO2R*, a web-based tool for differential gene expression analysis (Clough *et al.*, 2024). Log2FC values for all genes were calculated, enabling the identification of transcriptional changes across conditions in the GBS mouse model. This analysis provided a comparative framework for integrating findings from the *Rx-Dicer1* mutant mouse model, facilitating cross-validation of ETMR-associated gene expression patterns.

- Heatmap analysis of comparative transcriptomics
  - Full transcriptome

To compare the transcriptomes of *Rx-Dicer1* mutant mice, the GBS mouse model, and various human brain tumors, including ETMR, high-grade gliomas (HGG), and other subtypes, a comprehensive heatmap was generated using Log2FC values relative to their respective control samples. The human brain tumor data were derived from published bulk RNA-seq studies (Jessa *et al.*, 2019), while the mouse transcriptomes were obtained from *Rx-Dicer1* mutant RNA-seq (E11.5 and E17.5) and the GBS model (Neumann et al., 2017). Only confident orthologous genes between human and mouse, as identified by Ensembl's ortholog mapping criteria for mammalian species, were included to ensure robust cross-species comparison (Harrison et al., 2024).

Hierarchical clustering was performed in an unsupervised manner using *Euclidean* distances (default parameters) to group tumor and mouse model samples based on global transcriptomic similarities. Columns in the heatmaps represent these samples, while rows depict orthologous genes. The heatmaps were generated from Log2FC scaled between -1 and 1, with red indicating upregulation and blue indicating downregulation relative to wild-type (mouse) or healthy controls (human). For this full transcriptome comparison, various human brain tumors and different ETMR mouse models were included, and only "confident orthologous genes," defined according to Ensembl criteria for species comparisons (Harrison *et al.*, 2024), were used to align mouse and human data. All heatmaps were generated using *pheatmap* (v1.0.12) (Kolde, 2018).

- Top 500 differentially expressed genes

Heatmaps were created to compare gene set enrichment profiles of brain tumor types using custom gene sets comprising the top 500 upregulated (URG) and downregulated (DRG) DEGs for each tumor type compared to controls. GSEA was performed using Log2FC values from the full transcriptomes of the *Rx-Dicer1* mutant mouse model at E11.5 and E17.5 and the GBS mouse model, comparing them to wild-type (WT) controls. The top 500 upregulated (URG) and downregulated (DRG) DEGs heatmap represents the combined Normalized Enrichment Scores (NES) for upregulated and downregulated gene signatures (URG + DRG), calculated as the mean NES by combining the absolute value of downregulated NES and upregulated NES. Only statistically significant NES values (p-value < 0.05) are shown, highlighting enriched pathways for each tumor type. Hierarchical clustering was performed using *Euclidean* distances to group samples by similarity.

- Signature genes

The signature genes heatmap was generated using custom gene sets comprising the upregulated and downregulated gene signatures for each brain tumor type. NES for upregulated and downregulated gene signatures were combined by calculating the mean NES, using the absolute value of the downregulated NES and the upregulated NES. GSEA was performed on the full transcriptomes of mouse models, comparing them to controls based on Log2FCs. Only heatmap cells with a NES different from 0 are shown, indicating statistically significant enrichment (p-value < 0.05). Fewer tumor samples were included in this analysis compared to previous heatmaps, as tumor types were regrouped to derive robust and representative gene signatures.

- Bar plots and pie charts

Gene signature bar plots were generated to represent the ETMR upregulated signature using the combined p-adjusted values from Fisher's exact test. These

values were ranked from most statistically significant (left) to least significant (right) within the signature. Bars were color-coded to indicate whether the confident orthologous genes in mice were also DEGs between mouse models and controls. The pie charts accompanying the bar plots display the proportion of each category (colored bars representing DEGs) relative to the total upregulated ETMR signature, providing an overview of the contribution of DEGs across mouse models.

The similarity bar plot illustrates the comparison of DEGs from various mouse models (e.g., *Rx-Dicer1*, E11.5, E17.5, GBS) with ETMR molecular profiles. DEGs from each model were analyzed against two datasets: the ETMR-positive and ETMR-negative tumor signatures, and the DEGs from ETMR tumor samples compared to WT controls. The bar plot employs dual y-axes to represent: (1) the percentage similarity between mouse model DEGs and the combined ETMR-positive and negative tumor signatures (left y-axis) and (2) the percentage similarity between mouse model DEGs and the combined ETMR tumor vs. WT DEGs (right y-axis). Separate comparisons were performed for upregulated and downregulated genes: downregulated genes from each mouse model were matched to downregulated ETMR signatures and DEGs, while upregulated genes were compared similarly. The combined percentages for upregulated and downregulated genes were combined and visualized in this dual-axis bar plot.

- scRNA-seq of *Rx-Cre-Dicer*$^{F/F}$ mutant mice
    - Data generation and preprocessing

Single-cell RNA sequencing (scRNA-seq) data was generated using the *10x Genomics Cell Ranger* (v7.1.0) pipeline with the *mm10* mouse reference genome (Zheng et al., 2017). *Cell Ranger* was run with default parameters for alignment, barcode assignment, and UMI counting, ensuring high-quality preprocessing of raw sequencing data. After alignment, several quality control steps were performed to remove artifacts and low-quality cells. Ambient RNA contamination was addressed using the *SoupX* package (Young and Behjati, 2020), which corrected for background RNA to ensure accurate quantification of cell-specific transcripts. To identify and exclude empty droplets or high-nuclear read droplets, the *DropletQC* package was employed (Muskovic and Powell, 2021).

Further filtering criteria were applied using *Seurat v5* to retain only high-quality cells for downstream analysis (Hao *et al.*, 2024). Specifically, cells with more than 3,000 unique features, a nuclear fraction less than 0.5, and mitochondrial RNA content below 10% were included in the final dataset. This rigorous prefiltering ensured that the data used for clustering and downstream analysis was robust and representative of true biological variability.

    - Dimensionality reduction and clustering

Following preprocessing, the filtered datasets were normalized and scaled using *Seurat*'s standard workflow. Highly variable genes were identified using the *vst* method with the top 3,000 genes retained for subsequent analysis. PCA was performed to reduce dimensionality, and the top 30 principal components were selected based on an elbow plot. These components served as the input for Uniform Manifold Approximation and Projection (UMAP), which was used to visualize the datasets in two dimensions.

Cell clustering was performed using the *Seurat* functions *FindNeighbors* and *FindClusters*. For all datasets, including spinal cord and cortex samples, the *FindNeighbors* function was executed with dims = 1:30, followed by clustering with a resolution of 0.8. These parameters were consistently applied across WT and *Dicer1* knockout samples, resulting in the identification of 36 distinct cellular clusters.

- Cluster annotation and visualization

To assign specific cell types to the identified clusters, we utilized canonical gene markers for cortical development, as established in previous studies (Di Bella *et al.*, 2021; Ruan *et al.*, 2021; Dave *et al.*, 2023). NECs were identified using markers such as *Wnt8b*, *Crabp2*, *Hmga2*, *Dlk1*, *Prtg*, *Ccnd1*, *Igf2bp1*, and *Lrrn1*. aRGCs were annotated using markers including *Hes1*, *Hes5*, *Pax6*, *Sox2*, *Nes*, and *Aldoc*. IPCs were characterized by the expression of *Eomes*, *Neurog2*, *Gadd45g*, *Btg2*, and *Neurog1*, while early-born neurons were identified with markers such as *Neurod1*, *Tubb3*, *Nrp1*, *Gap43*, and *Dcx*. Finally, interneurons were annotated using *Dlx2* and *Dlx1*. Cycling markers used were *Top2a* and *Mki67*.

The expression patterns of these canonical markers are visualized in a dot plot, organized by unsupervised hierarchical clustering. The plot highlights the strong co-expression of genes within specific clusters. Additionally, we grouped the clusters into higher-level cell types for simplicity. The dot size represents the percentage of cells expressing a gene, and the color intensity reflects the average expression level. These dot plots were generated using the *ggplot2* package in R (Wickham, 2016).

To visualize the spatial distribution of annotated clusters, UMAP plots were generated. Clusters were aggregated into the previously mentioned broader categories, NECs, RGCs, IPCs, neurons, and cycling progenitors, to simplify the complex cellular landscape and highlight overarching developmental relationships. Among the identified cell populations, the cluster labeled *NEC_4_Lin28a* was named based on its strong expression of known NEC markers coupled with a notably high expression of *Lin28a*, which was not observed in other NEC clusters. In a similar way, the *Diff_IPCs_to_Interneurons_Lin28a* and *Diff_IPCs-Neurons_Lin28a* clusters exhibited gene signatures indicative of IPCs differentiating toward interneurons or neurons, with an additional marked expression of *Lin28a*.

To simplify downstream analyses, all NEC sub-clusters (e.g., *NEC_1, NEC_2, NEC_3*, etc.) were merged into a single group referred to collectively as NECs, with other cell types similarly aggregated under unified cluster labels.

- Cell type proportion analysis

To identify cell types with differential abundance between *Dicer1* KO and WT conditions, we utilized the *propeller* method from the *speckle* R package (Phipson *et al.*, 2022), designed specifically for testing differences in cell type proportions in single-cell RNA-seq data. *Propeller* considers both the number of cells in each cluster and the biological replicates across experimental conditions. It performs statistical tests to evaluate whether the observed differences in cell type proportions between KO and WT groups are significant. False discovery rates (FDRs) were calculated to correct for multiple testing, and FDR-adjusted p-values were used to determine the significance of the differences.

The statistical output from *propeller* includes p-values and FDRs for each cell type, which were used to annotate their significance levels. For the scatter plot of Log2FC in cell type proportions between KO and WT conditions, each point represents a cell type, with the x-axis denoting the cell type baseline proportion frequency and the y-axis representing the Log2FC. The point size reflects the baseline proportion frequency, and colors indicate significance levels: Red: Very significant (FDR < 1e-50), Orange: Significant (FDR < 0.05) and Black: Not significant (FDR > 0.05). The dashed horizontal line separates cell types enriched in KO (positive values) from those enriched in WT (negative values).

For the Log2FCs in cell proportions across the clusters UMAP, cell clusters enriched in KO were highlighted in red, while those enriched in WT were shown in blue, with intermediate fold change values visualized in a gradient. This spatial visualization was generated using *Seurat* and customized with *ggplot2* (Wickham, 2016).

- RNA velocity analysis using *Velocyto* and *scVelo*

To investigate transcriptional dynamics and predict cell fate trajectories, RNA velocity analysis was performed using *Velocyto* and *scVelo* (La Manno *et al.*, 2018; Bergen *et al.*, 2020). These tools allow tracking the transition of immature to mature RNA, estimating the velocity of individual cells, and inferring transcriptional trajectories in single-cell RNA-seq data.

The *Seurat* object was converted into an *AnnData* format using the *SeuratDisk* and *Matrix* R packages (La Manno *et al.*, 2018). The cell metadata, gene expression matrix, and dimensionality reductions were exported, and the *AnnData* object was created in Python. Spliced and unspliced transcript counts were generated using the *Velocyto* command-line tool with default parameters. The inputs included BAM files from single-cell sequencing aligned to the *mm10*

genome, along with GTF annotations. Repeat-masking was applied using the *mm10* repeat annotation to avoid erroneous signal from repetitive regions.

The spliced and unspliced counts matrices were integrated into the *AnnData* object using *scVelo* (Bergen et al., 2020). Cell barcodes were aligned, and duplicate genes were handled by making variable names unique. Dimensionality reduction embeddings and metadata from *Seurat* were added to the *AnnData* object for consistency with previously defined clusters.

RNA velocity was computed using the stochastic model of *scVelo* (Bergen *et al.*, 2020). Data was preprocessed by filtering and normalizing the spliced and unspliced matrices, followed by calculation of moments for transcriptional kinetics. A velocity graph was constructed to predict directional transitions between cell states based on splicing kinetics. RNA velocity was overlaid on the UMAP embedding to visualize transcriptional dynamics across clusters.

PAGA (Partition-based Graph Abstraction) was used to construct lineage trajectories and highlight the flow of RNA velocity vectors (Wolf *et al.*, 2019). The connectivity of transcriptional states was computed using velocity *pseudotime*.

- Feature plots and violin plots

To visualize the expression of ETMR marker genes across cell types and conditions, we used the *Seurat* v5 (Hao *et al.*, 2024) package to generate feature plots and violin plots.

Feature plots were created to display the spatial distribution of gene expression for key ETMR-associated genes, including *Lin28a, Sp8, Prtg, Trim71, Greb1*, and *Igf2bp1*, across all cell types. These plots overlay the normalized expression levels on the UMAP embedding of cells, allowing for a clear visualization of cell-type-specific expression patterns. The color gradient (blue to red) represents normalized expression values, with red indicating higher expression.

Violin plots were generated to compare the expression levels of two key ETMR markers, *Lin28a* and *Prtg*, between KO and WT samples. The *Seurat v5* violin plot function was used to present the distribution of normalized expression levels across cell types, sorted by their mean expression values.

RESULTS

Evolution of cortical neurogenesis mechanisms in amniotes

1. Mechanisms determining the modes of neurogenesis

To investigate the mechanisms regulating modes of neurogenesis in aRGCs across phylogeny, we performed bulk RNA-seq and bulk ATAC-seq analyses on purified aRGCs. We focused on developmental stages and cortical regions with known and different rates of direct vs indirect neurogenesis, across our amniote species: mouse, chick and snake. We selected specific conditions based on our lab's previous studies measuring basal progenitor (BP) levels using phosphohistone H3 (PH3). Histone H3, a core nuclear protein found in DNA chromatin, plays a significant role in chromosome condensation and cell-cycle progression during mitosis and meiosis. PH3 staining is highly specific for mitotically active cells, making it a marker for cell proliferation (Kim *et al.*, 2017).

These previous studies showed that neocortical aRGCs from E12 produce dividing BPs massively observed at E13-E14 in mouse, while neocortical aRGCs from E14 reduce the amount of dividing BPs production observed by E15-E16 (Figure 1A). In addition, at E12 there is more direct neurogenesis in the olfactory bulb (OB) compared to the Neocortex (NCx), so we separated OB aRGCs from E12 as well (Cárdenas *et al.*, 2018). In the chick dorsal pallium, our lab also previously discovered that aRGCs from E4 produce BPs peak between embryonic days (E) 5-6, having more BP mitoses in the lateral dorsal pallium (lDP) than in the medial dorsal pallium (mDP) (Figure 1B).

Knowing these BP mitotic ratios, we sequenced three conditions in mouse aRGCs: NCx E12 with the highest BP divisions (indicating higher indirect neurogenesis), OB E12 (same age, different area), and NCx E14 (same area, different age) with fewer BP divisions (indicating higher direct neurogenesis). In chick, we sequenced these two conditions in chick aRGCs: lateral E4 with the highest BP divisions (i.e. higher indirect neurogenesis), and medial E4 with fewer BP divisions (higher direct neurogenesis). In snakes, BP mitoses are absent throughout the dorsal pallium across development, so we took aRGCs from all pallial regions at E6.

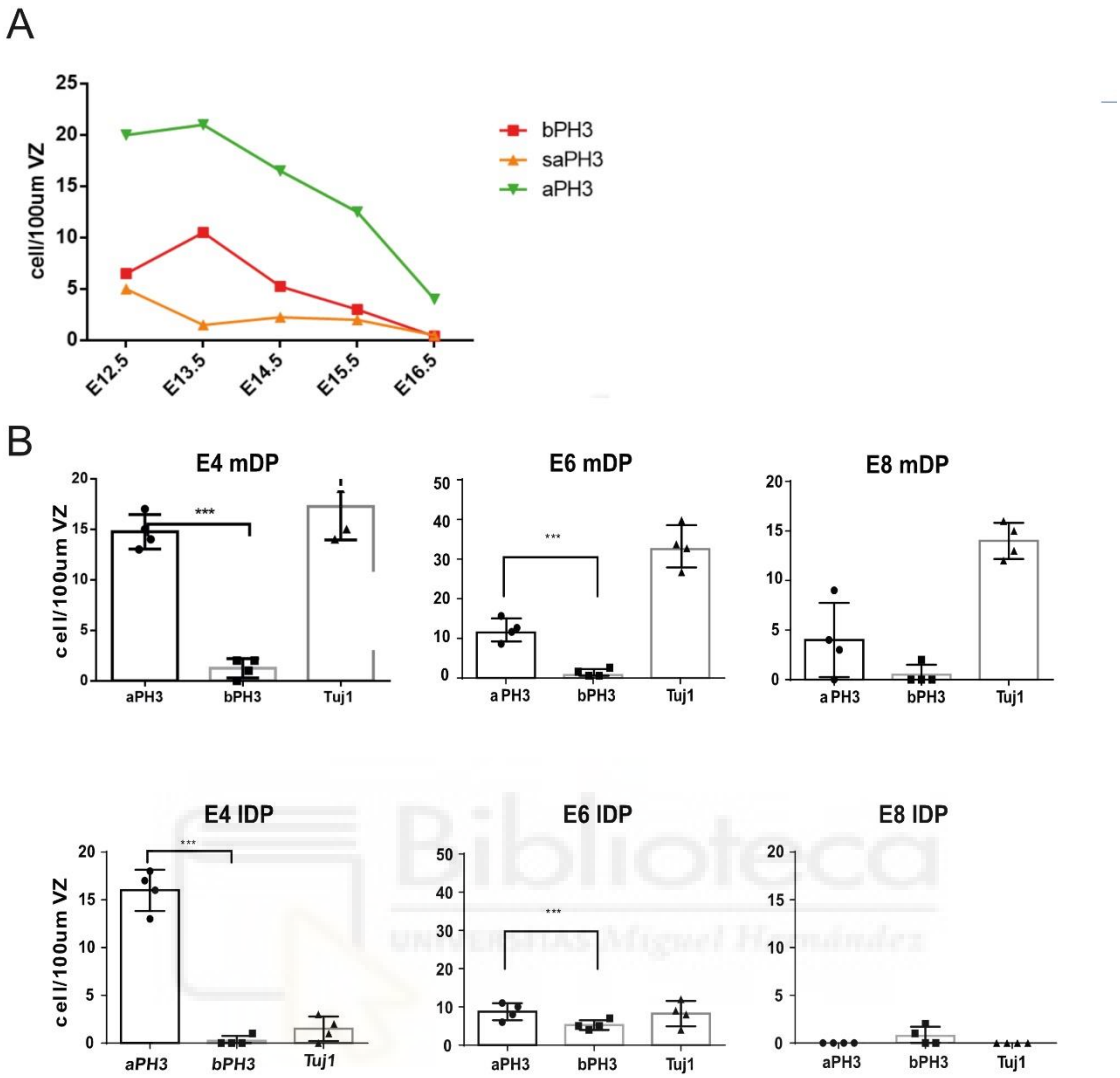


Figure 1. Characterization of basal mitoses in the developing neocortex and pallium across species. A) Quantification of mitotic cells in the ventricular zone (VZ) across embryonic stages (E12.5–E16.5) in the developing neocortex. Data points represent the number of PH3+ mitotic cells per 100 µm of VZ. Three mitotic subtypes are shown: basal PH3+ (bPH3; red), sub-apical PH3+ (saPH3; orange), and apical PH3+ (aPH3; green). **B**) Quantification of mitotic cells in the developing chick pallium at E4, E6, and E8. Boxplots depict the distribution of apical (aPH3) and basal (bPH3) PH3+ cells per 100 µm of VZ across developmental stages. Notably, basal mitoses peak at E6 in the lateral dorsal pallium (lDP) compared to the medial dorsal pallium (mDP). Statistical significance is indicated: ***p < 0.001.

To separate and isolate aRGCs, we used Immuno Fluorescence-Activated Cell Sorting (FACS), selecting the cells of interest with cell type-specific antibodies. To specifically select aRGCs we used anti-PAR3 (*Pard3*) antibodies, a member of the apical complex proteins present only in the apical adherent junction of aRGCs, which hence allows us to separate this cell type from each species specifically (Chou, Li and Wang, 2018; Tsuboi and Gotoh, 2020; Xu *et al.*, 2023).

For each condition, we performed bulk RNA-seq and bulk ATAC-seq on the purified aRGCs. The following results section details the analyses performed to

decipher the mechanisms involved in determining the modes of cortical neurogenesis across amniote phylogeny.

1.1. Transcriptomic mechanisms

1.1.1. Mouse

To unravel the mechanisms involved in determining the modes of neurogenesis, we first compared the transcriptomes of mouse aRGCs from regions and developmental stages that are enriched for indirect neurogenesis (IN-enriched; NCx E12) versus those enriched for direct neurogenesis (DN-enriched; OB E12, NCx E14). While both modes occur at each stage, these sample sets are biased toward one mode, providing a useful contrast consistent with—but not exclusive to—indirect and direct neurogenesis. RNA-seq analysis of protein-coding genes revealed 5.401 differentially expressed genes (DEGs) between NCx E12 and OB E12 littermates (Padj < 0.05 and 8.882 DEGs between NCx E12 and NCx E14 (Padj < 0.05).

To compile a unique list of DEGs between DN-enriched and IN-enriched samples, we combined the p-adjusted values of the two comparisons using Fisher's test method and calculated the mean log2 fold change (Log2FC). Only genes that were DEGs in both conditions were included in the final list, ordered by the combined p-adjusted value of Fisher's test. The volcano plot in Figure 2A illustrates DEGs more expressed in IN-enriched or DN-enriched conditions. Among the top significantly upregulated DEGs in the IN-enriched condition, we found *Ccnd1/2*, which promotes progenitor proliferation and self-renewal by inhibiting *Bcl6* (Bonnefont *et al.*, 2019). Specifically, Tsunekawa et al. (2012) observed that CCND2 protein can remain in the basal end-foot of one daughter cell after an aRGC division and hypothesized, on the basis of this correlative observation, that its retention might bias that daughter toward self-renewal, whereas its absence might allow terminal neuronal differentiation. To date, no direct gain- or loss-of-function experiments have tested this model (Tsunekawa *et al.*, 2012; D'Arcy and Silver, 2020; D'Arcy *et al.*, 2023).

Another interesting DEG was *Dmrta1*, which is activated by *Pax6*, regulate the expression of *Neurog2* (Noack *et al.*, 2022), necessary for the expression of *Eomes* (*Tbr2*) and the rewiring of the 3D genome structure to produce IPCs (Kikkawa *et al.*, 2013; Shimojo, Masaki and Kageyama, 2024). *Hmga2* maintains and expands neurogenic progenitors, particularly *Tbr2*-positive intermediate neuronal progenitors, by upregulating cell cycle genes (e.g., *Ki67*, *Cyclins*, *Cdks*) and neurogenic genes (e.g., *Eomes/Tbr2*, *Neurod1*, *Neurod6*), while inhibiting astrocyte marker expression (*Gfap*, *Aldh1l1*, *Sox9*, *Fabp7*) (Kuwayama et al., 2023). These findings suggest that the genes upregulated in the NCx E12 (IN-enriched) condition are prime candidates for promoting the IN neurogenic mode, given their roles in progenitor proliferation and BP production.

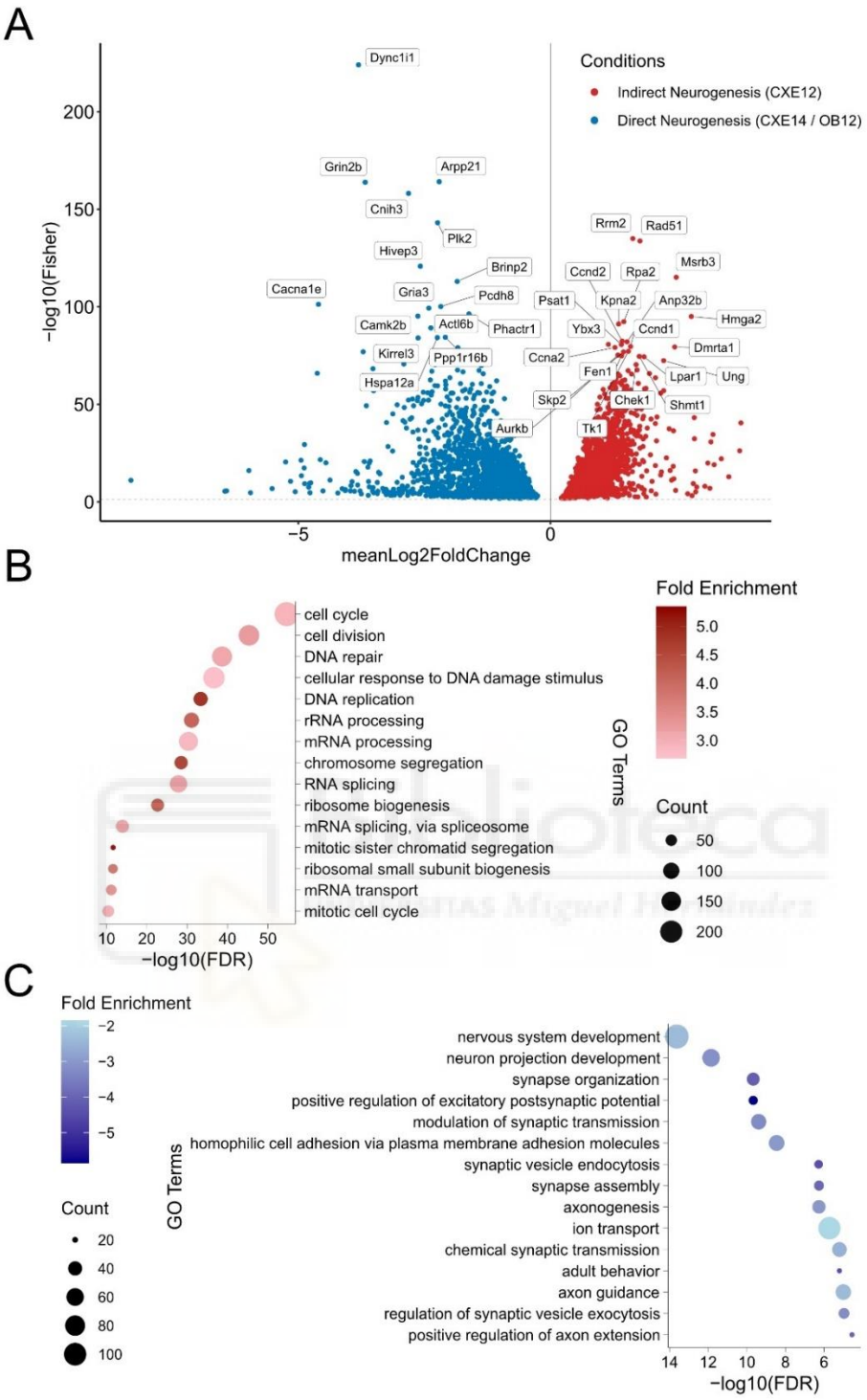


Figure 2. Comparative Transcriptomic Analysis of Indirect-Enriched and Direct-Enriched Neurogenesis in Mice. **A)** Combined analysis of DEGs between IN-enriched and DN-enriched conditions using Fisher's test. The volcano plot highlights significant DEGs more expressed in the IN-enriched or DN-enriched conditions, with critical proliferation-related genes such as *Ccnd1*, *Siva1*, and *Rrm2* marked as top significantly altered between the conditions. **D)** Gene Ontology (GO) term enrichment analysis for the IN-enriched condition (NCx E12), highlighting significant terms related to cell cycle, DNA repair, and RNA processing. Enrichment scores and p-values are represented, with cell cycle-related processes being the most enriched in the IN-enriched condition. **E)** GO term enrichment analysis for the DN-enriched condition, emphasizing terms associated with axonogenesis, synapse organization, and neuron projection development. The

color gradient indicates fold enrichment, with higher enrichment for processes involved in neuronal development and synapse formation.

In contrast, genes such as Glutamate ionotropic receptors (*Gria3*, *Grin2b*), calcium channels (*Cacna1e*), and Down Syndrome Cell Adhesion Molecules (*Dscaml1*) were more expressed in the DN-enriched condition. These genes are associated with neural maturation and cortical developmental syndromes (Platzer *et al.*, 2017; Mayo *et al.*, 2023). For example, *Grin2b* variants are linked to polymicrogyria (PMG) and the cortical appearance of tubulopathies (Mitsogiannis *et al.*, 2021), and *Gria3* is associated with intellectual disability (Karaca *et al.*, 2015). In conclusion, the genes upregulated in the OB 12.5 and NCx E14 conditions may act as key modulators of the DN neurogenic mode, promoting neuronal differentiation and maturation processes.

Gene Ontology (GO) analyses of DEGs revealed that the most highly enriched terms in IN-enriched samples were related to cell cycle and cell division (Figure 2B), whereas in DN-enriched samples, the terms were axonogenesis, synapse organization, and neuron projection development (Figure 2C). These results indicate that aRGCs from IN-enriched regions express higher levels of genes related to cell proliferation just before producing IPCs, which will go on to proliferate. In contrast, aRGCs from DN-enriched regions already express genes related to neuronal development, synapse organization, and axon guidance, typical of differentiating neurons.

1.1.2. Chick

To better understand the conserved and specific mechanisms driving neurogenesis across amniotes, we compared the transcriptomes of the E4 lateral Dorsal Pallium (E4 lateral) and E4 medial Dorsal Pallium (E4 medial) in chick embryos. RNA-seq analysis of protein-coding genes identified 2.042 DEGs between E4 lateral and E4 medial littermates (Padj < 0.05, Figure 3A).

Among the top significantly altered DEGs in the IN-enriched condition, several genes with interesting functions were identified, including *NR2E1*, *NR2F1*, *GADD45G*, *CTNNA1*, *EPHA3*, *DUSP16*, *CXCR7* and *LMO3* (Figure 3A). In fact, Tyrosine-Protein Kinase Receptor A3 (EPHA3), a RGC marker, is notably expressed in tumor-initiating cell populations in glioma, playing a crucial role in maintaining tumor cells in a less differentiated state through the activation of the mitogen-activated protein kinase (*MAPK*) pathway (Day *et al.*, 2013). On the other hand, dual-specificity phosphatase 16 (*DUSP16*), known to negatively regulate the *MAPK* pathway, promotes neural differentiation. Mice deficient in *Dusp16* (*Dusp16*−/−) develop fully penetrant hydrocephalus due to delayed cell cycle exit of neural progenitors, leading to overproliferation and an increased number of neurons, which is associated with increased brain volume (Zega *et al.*, 2017). Another gene, atypical chemokine receptor 3 (*CXCR7*), facilitates NPC migration via ERK1/2 activation (Chen *et al.*, 2015), while *LMO3* (+HEN2) expression diminishes *HES1* function, resulting in a phenotype similar to *Robo1/2*−/− mouse mutants (Borrell *et al.*, 2012), characterized by an abnormal

increase in IPCs and a hydrocephalic phenotype (Isogai *et al.*, 2015). Taken together, these findings indicate that in the chick, genes more highly expressed under the E4 Lateral condition exhibit a dual role, simultaneously promoting progenitor proliferation and driving neuronal differentiation. This contrasts with the mouse system, where the majority of genes upregulated in the IN-enriched condition appear predominantly to enhance proliferative mechanisms.

In the DN-enriched condition, several genes with interesting functions were identified, including *JAG2, ZNRF2, GRIA2/3, WNT7b, RSPO2/3* and *NFIA/NFIB*. The most significantly expressed gene in this condition was Wnt Family Member 7B (*WNT7B*), which inhibits IPCs formation and neuronal differentiation. Consistent with this upregulation, experimental overexpression of *WNT7B* in neural progenitor cells nearly eliminates the production of Tbr2+ cells (IPCs) and markedly reduces the number of Tbr1+ or Tubb3+ cells (early born neurons), without affecting the number of Pax6+ cells (aRGCs) (Papachristou *et al.*, 2014). Other *Wnt* DEGs, such as *WNT8B* and *WNT9A*, increases the aRGC population, hinders their maturation, and disrupts cortical organization by forming overproliferative rosettes (Zhu *et al.*, 2024). This overproliferative phenotype is also observed with two other genes more expressed in the DN-enriched condition, *NFIA* and *NFIB*. Double homozygous knockout mice for *NFIA* and *NFIB* exhibit severe ventricular enlargement and reduced numbers of differentiated neurons (Bunt *et al.*, 2017). The *RSPO* family members, *RSPO2* and *RSPO3*, which are roof plate-specific spondins, are more highly expressed in the DN-enriched condition. These *WNT* signaling regulators are crucial for neurodevelopmental decisions related to the dorso-ventral axis and the production of glutamatergic neurons (Zhang *et al.*, 2020; Prasad *et al.*, 2021). Although RSPOs' roles in cortical development are not well understood, *RSPO2* has been shown to promote the differentiation of midbrain dopaminergic neuroblasts into dopaminergic neurons in mouse primary cultures and embryonic stem cells, indicating a link to neuronal maturation (Gyllborg *et al.*, 2018). In conclusion, these findings reveal that in the chick, the genes upregulated under the E4 Medial (Direct Neurogenesis) condition simultaneously drive both progenitor proliferation and neuronal differentiation. This pattern contrasts with the mouse, where the genes more highly expressed in the DN-enriched condition primarily promote neural differentiation and maturation.

GO analysis of DEGs revealed that IN-enriched regions were most strongly associated with terms related to nervous system development and axon guidance (Figure 3B), while DN-enriched regions were linked to terms involving extracellular matrix organization, BMP signaling pathway, and collagen fibril organization (Figure 3C). These findings suggest that aRGCs from IN-enriched regions express more neurogenesis-related genes just prior to produce IPCs, whereas aRGCs from DN-enriched regions already express genes related to extracellular matrix and BMP signaling.

In conclusion, aRGCs from both IN-enriched and DN-enriched conditions in chick express genes associated with the proliferation and expansion of apical

progenitors, as well as with IPCs production, neuronal differentiation and maturation.

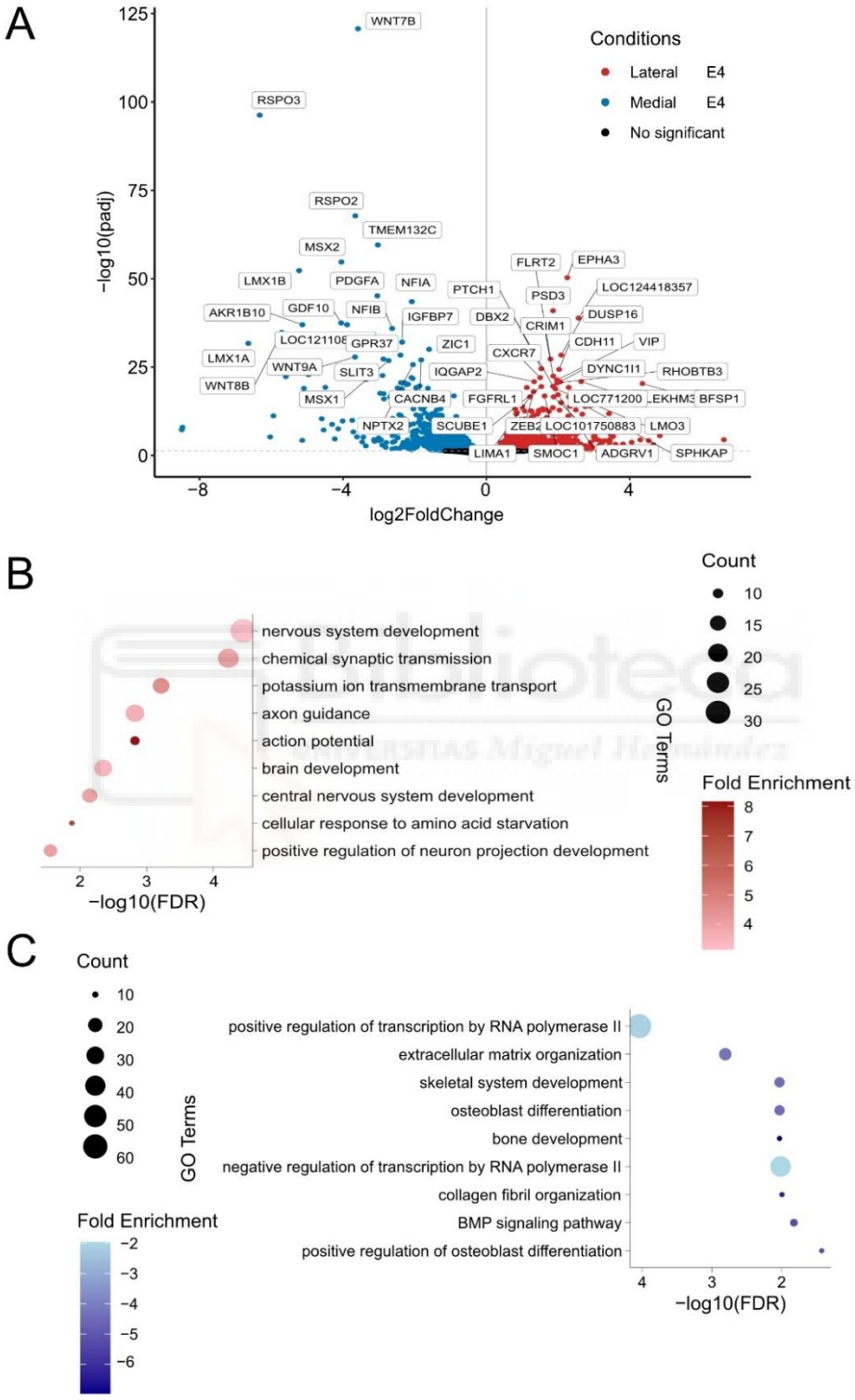


Figure 3. Comparative Transcriptomic Analysis of Indirect and Direct Neurogenesis in Chick. **A)** Volcano plot displaying DEGs between the lateral (red) and medial (blue) dorsal pallium regions in E4 chick embryos. Top genes such as *WNT7B, RSPO3*, and *EPHA3* are highlighted, indicating their differential expression in lateral versus medial regions. **B)** Gene Ontology (GO)

enrichment analysis for DEGs from the lateral dorsal pallium, revealing terms associated with nervous system development, synaptic transmission, and axon guidance. **C)** GO enrichment analysis for DEGs from the medial dorsal pallium, highlighting enriched terms related to extracellular matrix organization, osteoblast differentiation, and BMP signaling pathways.

1.1.3. High confidence orthologous genes

The mechanisms by which aRGCs generate neurons either through DN, or indirectly through the production of IPCs (IN), are likely complex and tightly regulated. While the transcriptomic landscapes of these processes have been studied separately in mouse and chick models, identifying shared genetic pathways across species is crucial for understanding the conserved and divergent mechanisms of neurogenesis in amniotes.

In this section, we aimed to analyze genes that are confident orthologous coding genes between chick and mouse, and identify those that are consistently more expressed in either IN-enriched or DN-enriched conditions. Orthologous genes are those that have evolved from a common ancestral gene and typically retain the same function across different species. By selecting these genes, we can uncover conserved molecular mechanisms that govern neurogenesis, offering insights into the evolutionary processes that shape brain development across phylogeny.

To ensure the accuracy of our orthologous gene identification, we used the criteria established by Ensembl to determine high confidence orthologs. According to Ensembl, a high confidence ortholog is defined based on criteria such as sequence identity, synteny, and phylogenetic support. Phylogenetic support is also considered, ensuring that the orthologs are derived from a common ancestral gene and are not paralogs that have arisen from gene duplication events within the same species (Harrison *et al.*, 2024). These stringent criteria provide a reliable framework for identifying true orthologs, minimizing the likelihood of false positives.

Our approach involves comparing the transcriptomes of DN-enriched and IN-enriched conditions in both mouse and chick, as detailed in the previous sections. By focusing on these high confidence orthologous genes that are DEGs in both species, we will compile a list of genes that consistently exhibit higher expression in either IN-enriched or DN-enriched across species. This analysis will allow us to discern not only the shared genetic basis of neurogenesis but also the species-specific adaptations that may contribute to differences in neurogenic modes between chick and mouse.

1.1.3.1. Indirect neurogenesis

Figure 4 highlights high-confidence orthologous genes that are significantly overexpressed across all IN-enriched conditions in both mouse and chick. Among the most differently expressed we found members of the Nuclear Receptor Subfamily 2, including *NR2E1* (*Nr2e1*) and *NR2F1* (*Nr2f1*), which comprises

prominent TFs family involved in cortical development. *Nr2e1*, expressed in aRGCs during cortical development, is critical for promoting the IN-enriched pathway: its absence leads to a marked reduction in IPCs throughout cortical development and disproportionate reduction in size of specific cortical areas (Drill, 2009; Sobhan and Funa, 2017). Similarly, *Nr2f1* overexpression in the mouse NCx significantly increases the number of *Tbr2*-expressing cells, whereas *Nr2f1* knockout (*Nr2f1*-/-) mutants show reduced *Tbr2*-expressing cell populations, underscoring its role in IPC formation (Faedo *et al.*, 2008).

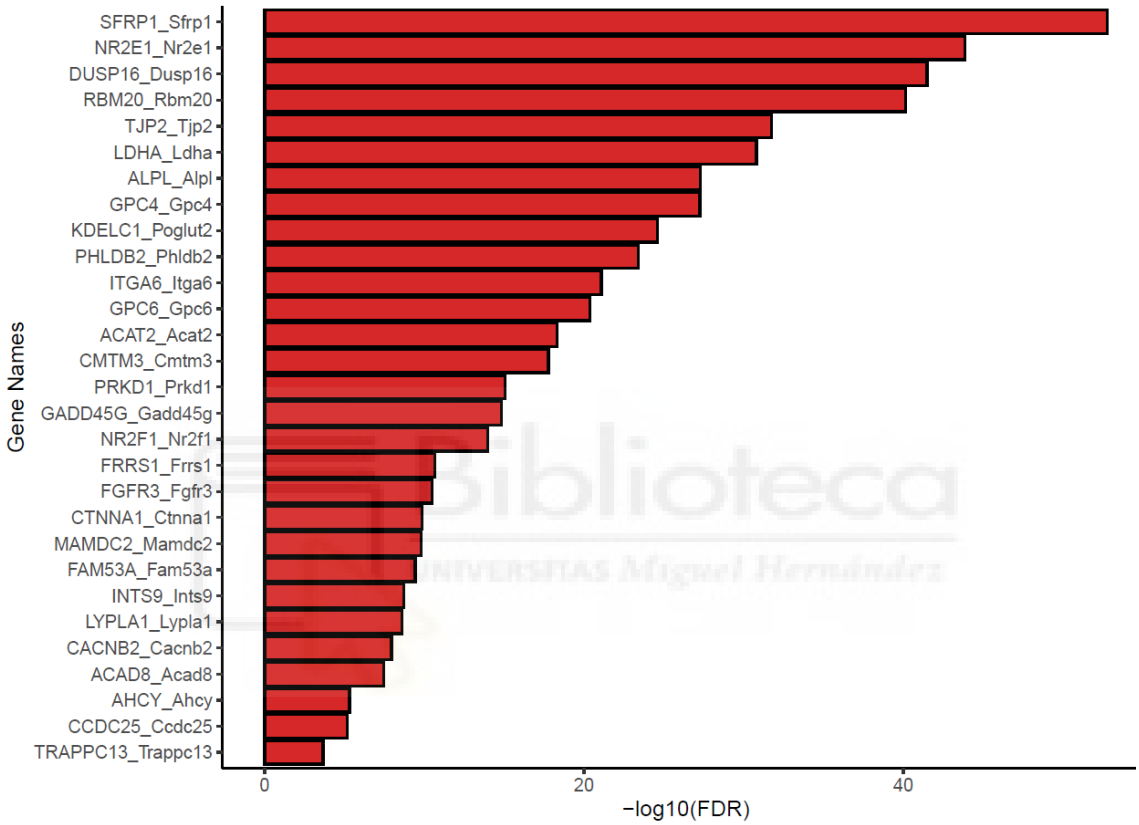


Figure 4. High-Confidence Orthologous Genes Associated with Indirect Neurogenesis in Mouse and Chick. This bar plot shows genes significantly overexpressed across all indirect neurogenesis (IN) conditions in both mouse and chick models, representing high-confidence orthologous genes. Genes are ordered based on their combined significance across multiple comparisons, using the Fisher -log10(FDR) for ranking. This selection includes key transcription factors (TFs), cell cycle regulators, and signaling molecules critical for neurogenesis, highlighting conserved regulatory mechanisms across species.

Another significant gene linked to IPCs is *GADD45G* (*Gadd45g*). In studies utilizing the *Gadd45g*-d4Venus mouse model, which visualizes *Gadd45g* expression with green fluorescent labeling during cortical development, most *Gadd45g*-d4Venus+ mitotic cells (pH3+) in the SVZ correspond to Tbr2+ IPCs. Moreover, cerebral tissues labeled with FM4-64, which highlights cell boundaries, show that apical processes of aRGCs retract and disappear when *Gadd45g* is expressed, indicating a role for *Gadd45g* in the commitment of aRGCs to the IPC lineage (Kawaue *et al.*, 2014).

DEGs overexpressed in IN-enriched samples also included genes related to cell proliferation, such as *CTNNA1* (*Ctnna1*, also known as αE-catenin). *Ctnna1* expression activates Hedgehog signaling and sustains β-catenin signaling, contributing to hyperproliferation of aRGCs in the VZ (Lien *et al.*, 2006; Stocker and Chenn, 2009). Interestingly, the α-catenin isoform transitions from *Ctnna1* in aRGCs to *Ctnna2* (αN-catenin) in IPCs and differentiating post-mitotic neurons, suggesting distinct roles in neurogenesis (Bedogni and Hevner, 2021).

Despite its inclusion in the IN-enriched gene list, *SFRP1* (*Sfrp1*, Secreted frizzled-related protein 1) has been shown to limit proliferation. Overexpression of *Sfrp1* decreases Pax6+ BrdU+ and Tbr2+ BrdU+ cell numbers, while *Sfrp1* knockout (*Sfrp1*-/-) mutants exhibit increased populations of these cells, indicating that *Sfrp1* restricts the early-stage proliferation of RGPs and IPCs (Miao et al., 2018). Since aRGCs in NCx E12 are more proliferative than those in NCx E14, and the expression of *Sfrp1* decreases with development, the presence of *Sfrp1* on the IN-enriched gene list is paradoxical and requires further exploration.

1.1.3.2. Direct neurogenesis

We also found a number of genes significantly overexpressed across all DN-enriched conditions in both mouse and chick models, representing high-confidence orthologous genes (Figure 5). Among these, *JAG2* (*Jag2*, Jagged2) stands out, supported by previous research from our lab showing that elevated levels of *Robo1/2* inhibit *Dll1* expression. This repression promotes the expression of *Jag1* and *Jag2*, favoring direct neurogenesis, as observed in the OB and contrasting with the NCx (Cárdenas *et al.*, 2018). This observation strongly reinforces the inclusion of *Jag2* among genes associated with DN.

Another notable gene on this list is *ZNRF2* (*Znrf2*), which has been shown to be significantly enriched in aRGCs that produce neurons directly rather than IPCs. This finding aligns with the proposed hypothesis that *Znrf2* may play a role in promoting direct neurogenesis pathways (Pinto *et al.*, 2008).

The list also includes several genes related to synaptogenesis and axon guidance, although their specific roles in regulating neurogenic modes are yet to be elucidated.

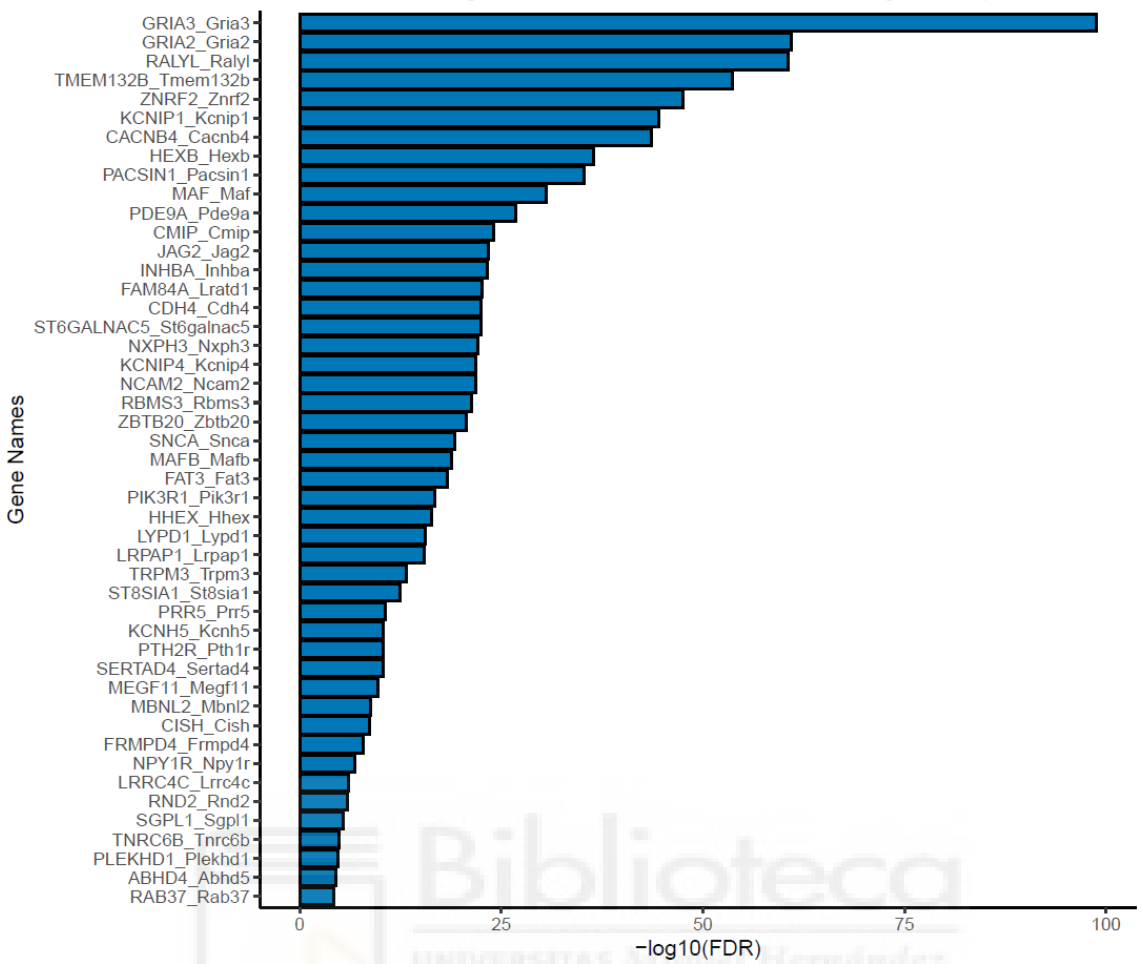


Figure 5. High-Confidence Orthologous Genes Associated with Direct Neurogenesis in Mouse and Chick. This bar plot shows genes significantly overexpressed across all DN-enriched conditions in both mouse and chick models, representing high-confidence orthologous genes. Genes are ordered based on their combined significance across multiple comparisons, using the Fisher -log10(FDR) for ranking. This selection highlights conserved genes associated with direct neurogenesis, including transcriptional regulators and molecules involved in synapse formation and axonal pathfinding.

1.2. Epigenomic mechanisms

1.2.1. Mouse

To investigate chromatin accessibility changes underlying differential gene expression and the different modes of neurogenesis, we performed ATAC-seq on aRGC samples from IN-enriched conditions (NCx E12) and DN-enriched conditions (OB E12, NCx E14) in mice. ATAC-seq analysis revealed 8.270 Differentially Accessible Regions (DARs) between NCx E12 and NCx E14 (Padj < 0.05; Figure 6A) and 1.724 DARs between NCx E12 and OB E12 littermates (Padj < 0.05; Figure 6B). These DARs provide insights into regulatory elements that potentially influence neurogenic pathways under different conditions.

Among the top significantly altered DARs in the IN-enriched condition that are also DEGs, we identified regions associated with *Dmrta1* [NCx E12 vs OB E12 -

chr4:89677974-89680145, padj=1,60E-06; NCx E12 vs NCx E14 - chr4:89650307-89652359, padj=2,52E-25], Jun [NCx E12 vs OB E12 - chr4:95184354-95186546, padj=4,06E-08; NCx E12 vs NCx E14 - chr4:95134664-95135223, padj=9,06E-17] or *Fzd8* [NCx E12 vs OB E12 - chr18:8719003-8720051, padj=8,62E-10; NCx E12 vs NCx E14 - chr18:9026181-9027771, padj= 5,19E-07]. *Dmrta1*, as previously discussed, regulates IN mode genes such as *Pax6* and *Neurog2*. *Jun*, a component of the Activator Protein 1 (*AP-1*) complex along with *Fos* and related factors, plays a crucial role in cell cycle regulation by controlling genes like *Cyclin D1, p53, p21cip1/waf1, p19ARF, and p16*. *Jun* promotes cell proliferation by repressing tumor suppressor gene functions (Shaulian and Karin, 2001). Additionally, blocking neuro stem cell proliferation is achievable through inhibitors of the *FOS*/*JUN* complex (Pagin *et al.*, 2021).

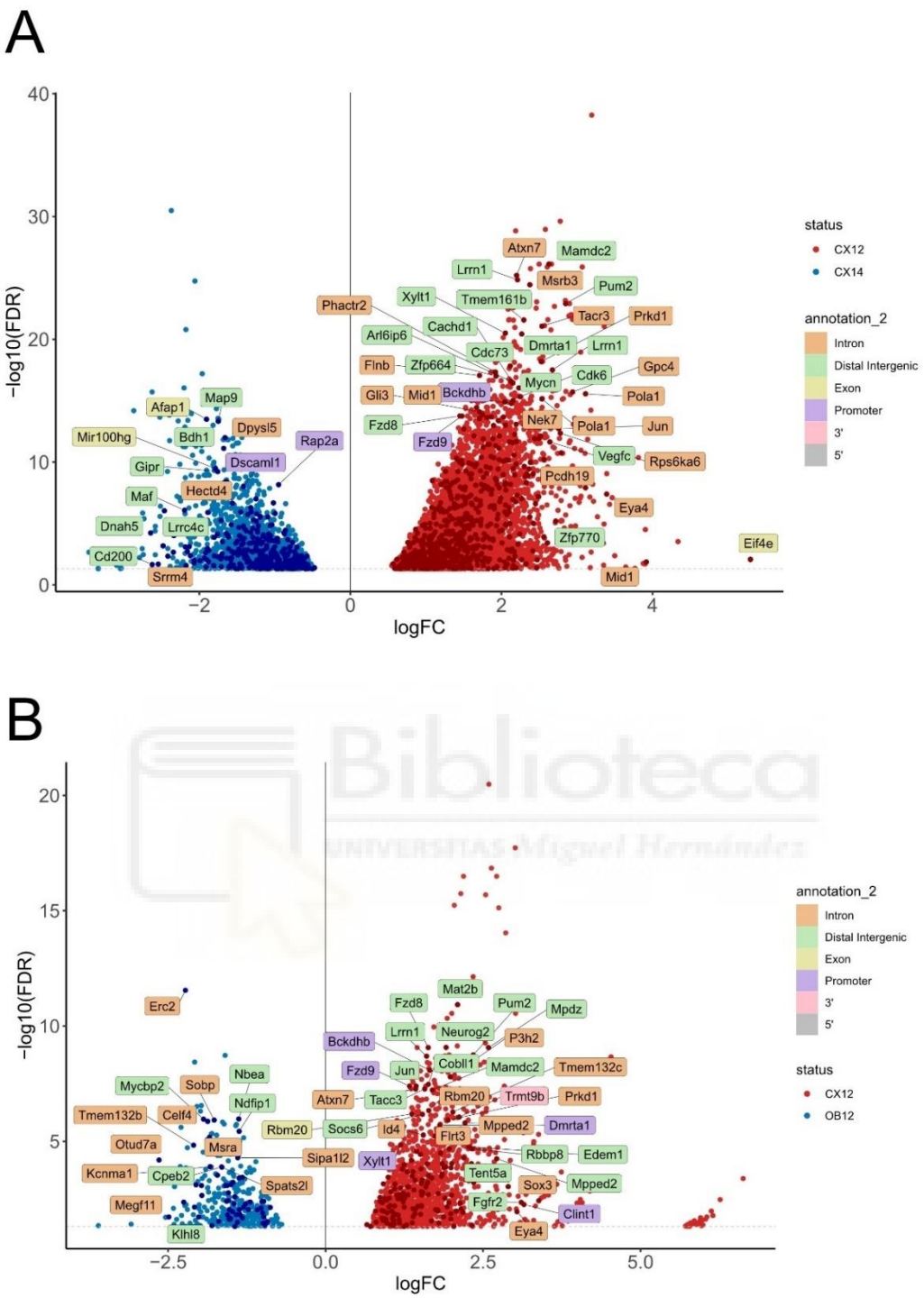


Figure 6. Differential Chromatin Accessibility Associated with Indirect and Direct Neurogenesis in Mice. A) Volcano plot showing differentially accessible regions (DARs) between NCx E12 (indirect neurogenesis, CX12) NCx E14 (direct neurogenesis CX14) conditions. **B)** Volcano plot showing DARs between NCx E12 (indirect neurogenesis, CX12) and OB E12 (direct neurogenesis, OB12) conditions. Red points indicate regions more open in IN-enriched regions and blue points representing regions more open in DN-enriched regions. Dark red or dark blue points highlight DARs that also correspond to DEGs from RNA-seq analysis. Notable genes associated with these DEGs are annotated in the plot, highlighting potential regulatory elements that influence neurogenic pathways. Annotation labels (color-coded) indicate the genomic features associated with each DAR, including introns, exons, promoters, and distal intergenic regions. Non-significant regions dots were not included to enhance the visualization of significant DARs and maintain clarity.

Overexpression of *Jun* in hippocampal neural stem cells leads to increased expression of *Tbr2* (*Eomes*) and *Foxg1*, directly influencing the IPC population (Siegenthaler, Tremper-Wells and Miller, 2008; Scopa *et al.*, 2023). Furthermore, enhanced expression of *Fzd8* through the HARE5 human enhancer shortens cell-cycle duration, reduces G1/G2/M phases, and expands the cortical progenitor pool, resulting in a larger and thicker cortex (Boyd *et al.*, 2015; Silver, 2016).

In the DN-enriched condition, notable DARs that are also DEGs include regions associated with *Mir100hg* [NCx E12 vs OB E12 - chr9:41376046-41379383, padj=7,55E-07; NCx E12 vs NCx E14 - chr9:41664873-41665898, padj=3,95E-13]. *Mir100hg* is a microRNA host gene that encodes m*ir-100, mir-125b-1, and let-7a-2*. Interestingly, overexpression of Notch1-ICD in the dorsal telencephalon (E10.5) substantially increases *Mir100hg* levels, subsequently elevating *mir-100, mir-125, and let-7* levels (Lu *et al.*, 2017). This upregulation completely depletes proliferative IPCs (Ph3+ Eomes/Tbr2+ BPs) in the mouse SVZ, suggesting a role for *Mir100hg* in inhibiting the IN pathway (Han *et al.*, 2023).

1.2.2. Chick

To investigate the chromatin accessibility changes underlying different modes of neurogenesis in chick, we performed ATAC-seq on samples from lateral and medial regions of the dorsal pallium of E4 embryos. ATAC-seq analysis revealed 2.397 DARs between lateral E4 and medial E4 (Padj < 0.05; Figure 7) in the chick model.

Among the top significantly altered DARs in the lateral E4 condition that are also DEGs, we identified regions associated with genes such as Immunoglobulin Superfamily DCC Subclass Member 3 (*IGDCC3* [chr10:17893914-17894754, padj=3,64E-19]), which has been previously implicated in neural development pathways and is notably overexpressed in *Dicer1* downregulation, leading to hyperproliferative rosettes (Goggolidou *et al.*, 2013; Fernández *et al.*, 2020). Additionally, *DUSP16* [chr1:71987135-71988409, padj=6,35E-10] also emerged as a significant region, whose deficiency in mice has been linked to fully penetrant hydrocephalus due to delayed cell cycle exit of neural progenitors, as mentioned before (Zega *et al.*, 2017). Another key enhancer identified was associated to *FGFR3 [chr4:82967351-82968278*, padj=7,07E-05] known to control cortical development by regulating the proliferation and apoptosis of cortical progenitors (Inglis-Broadgate *et al.*, 2005).

Genomic regions with increased accessibility in the medial dorsal pallium included genes such as *NFIB* [chrZ:32077624-32078914, padj=7,73E-12] and *HMGA2* [chr1:34254731-34255370, padj=8,35E-08], both of which are critical in aRGC proliferation and the formation of BPs (Kishi *et al.*, 2012; Betancourt, Katzman and Chen, 2014; Kuwayama *et al.*, 2023). However, we observed increased accessibility in genomic regions associated with genes crucial for neuronal maturation, such as *CACNB4* [chr7:34974435-34975054,

padj=6,86648E-05], KCNIP4 [chr4:74117457-74118162, padj=3,06E-05], and *KCNK7* [chr3:38354726-38355266, padj=4,60E-04] which have roles in synaptic transmission, ion channel regulation, and neuronal excitability (Pruunsild and Timmusk, 2005; Yost *et al.*, 2008; Ronjat *et al.*, 2013).

These findings suggest that both lateral and medial E4 regions exhibit distinct but overlapping regulatory landscapes. Notably, genes related to both DN-enriched and IN-enriched samples are found in both conditions, indicating that the regulatory patterns are not strictly segregated between the two regions like in mouse. This overlap points to a complex interplay between chromatin accessibility and gene expression in determining neurogenic modes, making the results less definitive in distinguishing between direct and indirect modes of neurogenesis.

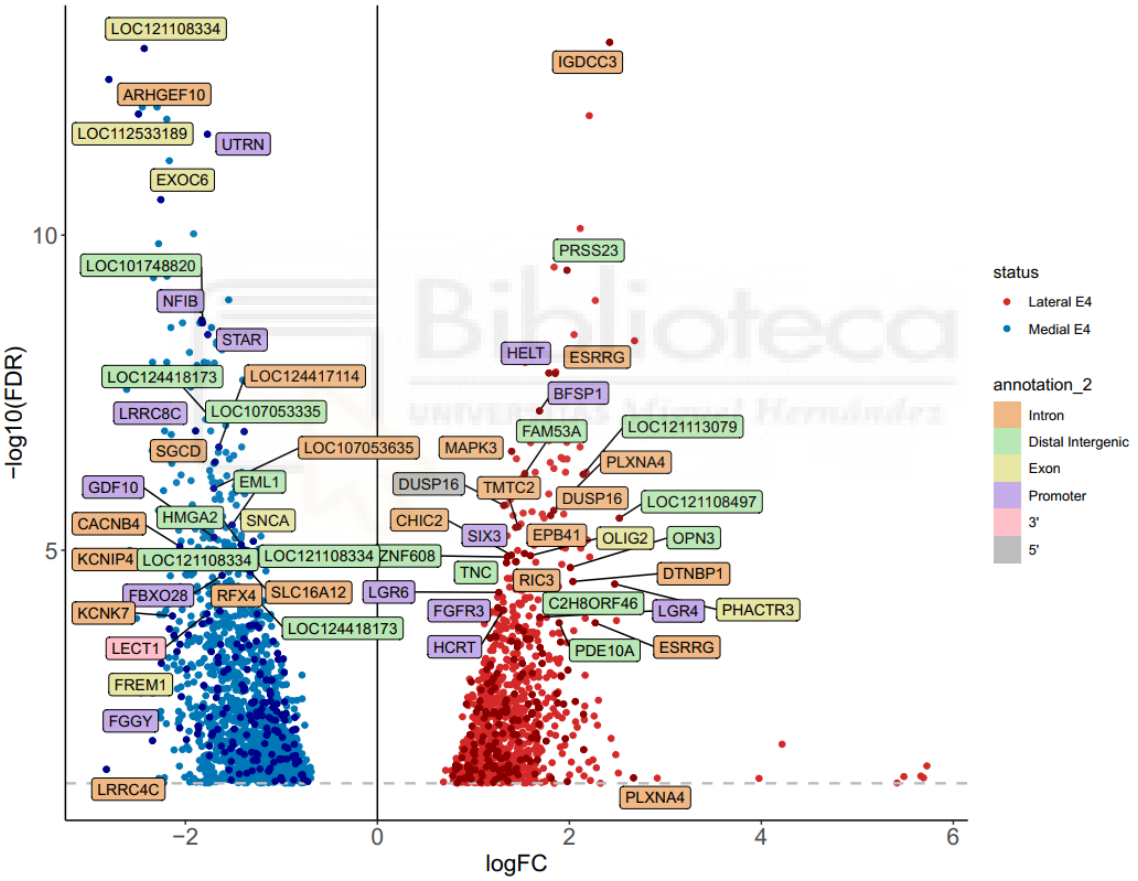


Figure 7. Differential Chromatin Accessibility in Lateral and Medial E4 Regions of Chick Embryos. Volcano plot showing the regions of differential chromatin accessibility (DARs) between lateral E4 and medial E4 regions. Red points represent DARs with increased accessibility in the lateral E4 regions, while blue points indicate increased accessibility in the medial E4 regions. Dark red and dark blue points correspond to regions that are also differentially expressed genes (DEGs) in RNA-seq comparisons. Notable genes associated with these DEGs are annotated in the plot, highlighting potential regulatory elements that influence neurogenic pathways. Annotation labels (color-coded) indicate the genomic features associated with each DAR, including introns, exons, promoters, and distal intergenic regions. Non-significant regions dots were not included to enhance the visualization of significant DARs and maintain clarity.

1.3. Candidate genes

1.3.1. aRGC self-amplification: *Mir3607*, critical regulator of aRGC expansion and cortical size in mammal

As we have seen, miRNAs like *Mir100hg, Mir-100, Mir-125,* and *let-7* have been well established as influential regulators of neurogenesis, particularly in the modulation of progenitor cell proliferation and differentiation. Building on this understanding of miRNA functions, our lab identified *Mir3607* (also known as *SNORD138*) as another miRNA playing key roles in the expansion of aRGCs in an independent study.

Research into the evolutionary expansion of the mammalian cerebral cortex has highlighted that the proliferation of progenitor cells during embryonic development was a key factor in the increased brain size and complexity seen in species such as primates. However, in rodents, this process reversed, resulting in smaller, smoother brains. The genetic mechanisms underlying this secondary loss in rodents have been largely unknown until our investigation into *Mir3607*.

We discovered that *Mir3607* is expressed embryonically in the large cortex of primates and ferrets—species distant from the rodent lineage—but not in mice. Experimental expression of *Mir3607* in embryonic mouse cortex led to a marked increase in *Wnt*/β-catenin signaling, which in turn amplified aRGC and IPCs proliferation and expanded the VZ. This effect is mediated by the inhibition of APC (adenomatous polyposis coli), a key negative regulator of β-catenin signaling. Loss of endogenous *Mir3607* in ferrets was found to reduce RGC proliferation and increase in migration, while its overexpression in human cerebral organoids promoted VZ expansion.

Although *Mir3607* is conserved in mice, ferrets, and humans, it is not expressed during cortical development in mice. To investigate the underlying reasons for this difference in expression, we explored the proximal regulatory sequences to identify which are conserved or specific between gyrencephalic (e.g., human, ferret) and lissencephalic (e.g., mouse) animals. This comparison could help explain the observed differences in gene expression.

In Figure 8A and Figure 8B, we performed a comprehensive view of the proximal enhancers and promoters for *Mir3607* using *SCREEN* (Moore *et al.*, 2020), highlighting differences in TFs predicted to bind specifically to either gyrencephalic or lissencephalic sequences. These differences likely arise from the lack of sequence conservation between these two groups of animals. Notably, TFs such as *Stat1, Insm1, Foxb1, Tead1, and Nkx2-1*, which are critical for regulating cortical neurogenesis, show distinct binding patterns that could contribute to the divergent expression profiles observed (Farkas *et al.*, 2008;

Zhao *et al.*, 2008; Delgado and Lim, 2015; Mukhtar *et al.*, 2020; Imitola *et al.*, 2023).

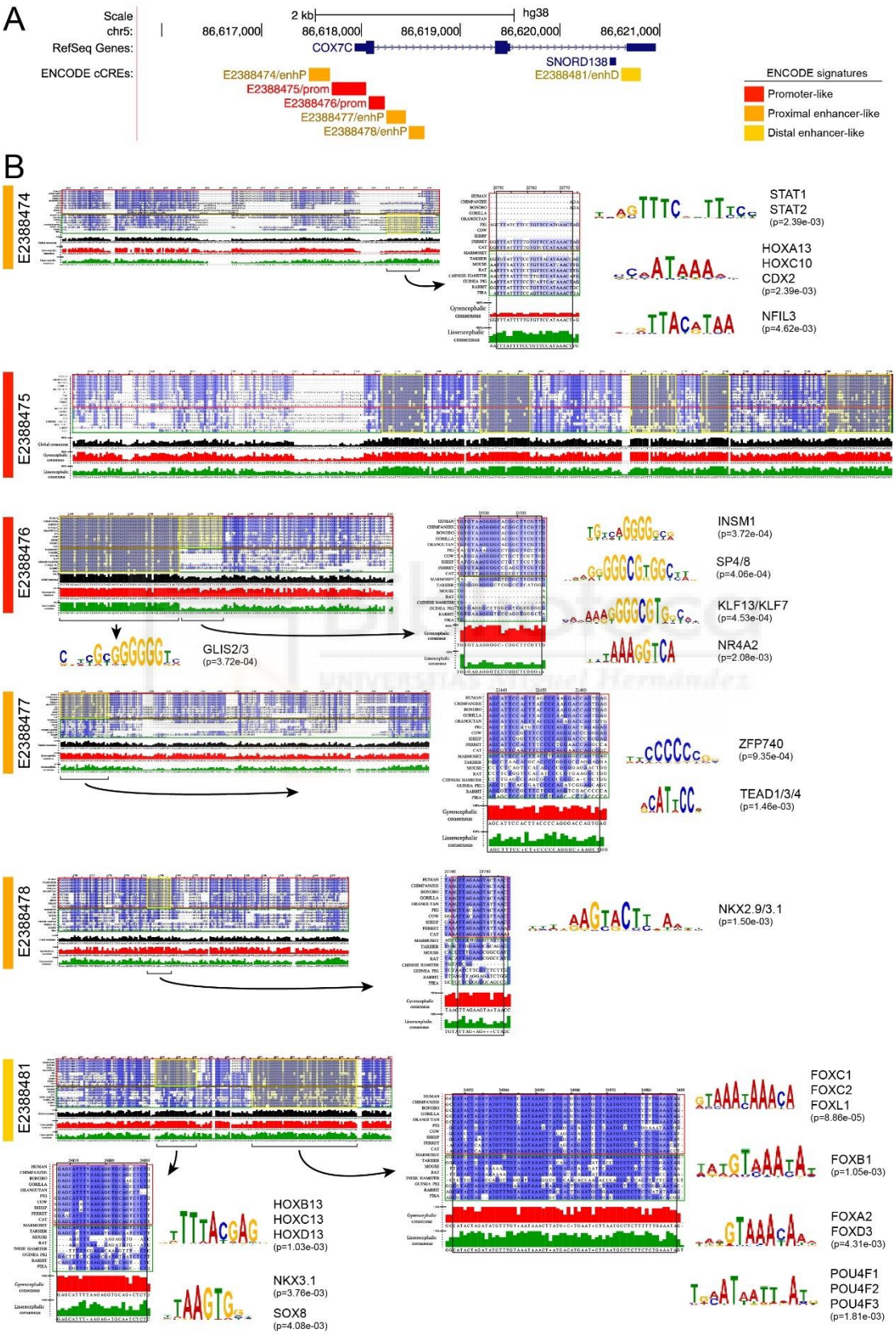


Figure 8. Conservation and Divergence of *MIR3607* (*SNORD138*) Expression Regulation Across Species. A) Genomic region of *MIR3607* (*SNORD138*) in the human genome (*GRCh38/hg38*), showing candidate cis-regulatory elements (cCREs) by *SCREEN* (Moore *et al.*, 2020) with promoter-like, proximal enhancer-like, and distal enhancer-like signatures. These regions were identified based on ENCODE data and are color-coded accordingly.
B) DNA sequence alignments for the region spanning -5 kb to +1 kb downstream of the *MIR3607* transcription start site (TSS) across 18 species. The top portion (red boxes) includes gyrencephalic species such as humans, chimpanzees, and pigs, while the lower portion (green boxes) includes lissencephalic species such as mice, rats, and guinea pigs. Conserved regions (yellow boxes) were selected based on sequence alignments, and putative TF binding sites were identified using MEME. The identified motifs and their associated TFs are shown along with p-values for binding significance. The percent sequence identity for gyrencephalic (red) and lissencephalic (green) species is plotted below each alignment overview, highlighting conservation patterns across these species.

An additional possibility is that the miRNA is expressed but subsequently degraded due to incorrect folding, which prevents *Dicer1* from processing and maturing it. In Figure 9, we present a schematic comparison of the predicted secondary structure of pre-*MIR3607* in humans (*Homo sapiens*) and mice (*Mus musculus*) using *MXfold2* (Sato, Akiyama and Sakakibara, 2021). The sequence differences between these species are predicted to impact the structure of the loop region (highlighted by red brackets), which is shorter in mice (indicated by an arrow) (Gu *et al.*, 2012). This structural variation may influence how *Dicer1* matures *Mir3607* differently between humans and mice.

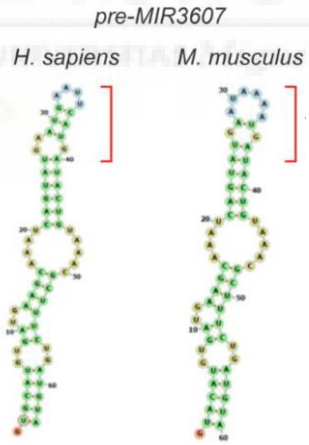


Figure 9. Comparative Predicted Secondary Structure of pre-*MIR3607* in Humans and Mice. This schematic depicts the predicted secondary structure of pre-*MIR3607* in humans (*Homo sapiens*) and mice (*Mus musculus*), based on (Sato, Akiyama and Sakakibara, 2021). Sequence differences between the species are predicted to affect the loop region (highlighted by red brackets), which is shorter in mice, as indicated by the arrow.

These findings highlight *Mir3607* as a pivotal miRNA that underwent secondary loss during mammalian evolution, playing a crucial role in limiting aRGC amplification and contributing to the reduced cortex size observed in rodents. Understanding the function of *Mir3607*, along with other key miRNAs such as *Mir100hg, mir-100, mir-125, and let-7*, provides valuable insights into the molecular mechanisms underlying brain development across different mammalian species.

1.3.2. Indirect Neurogenesis: *Sall1*, master regulator of modes of neurogenesis in mice

ATAC-seq analyses in mice comparing NCx E12 vs NCx E14 (CX12 – CX14) and NCx E12 vs OB E12 (CX12 – OB12) revealed distinct DARs. Using MEME (Bailey *et al.*, 2015), we identified ungapped TF binding motifs that are enriched within these DARs. Remarkably, a similar motif appeared in all conditions with a highly significant p-value. Using *TOMTOM* (Gupta *et al.*, 2007) with TF database HOmo sapiens COmprehensive MOdel COllection v12 (*HOCOMOCO*) (Vorontsov *et al.*, 2024), we found that SALL1.H12CORE.0.P.C is significantly predicted to bind to all these motifs (Figure 10).

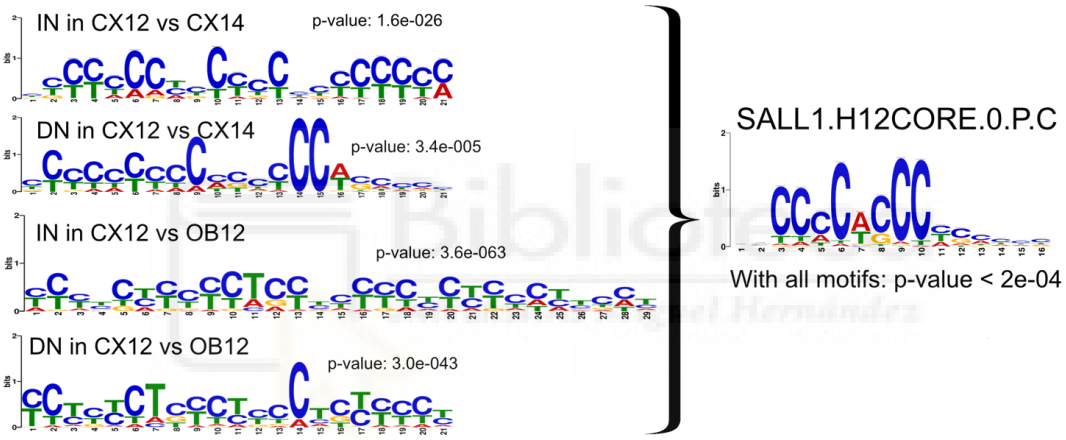


Figure 10. Identification of Enriched Motifs in Differentially Accessible Regions (DARs) from ATAC-seq Analysis in Mice. ATAC-seq analysis comparing neocortex E12 – neocortex E14 (CX12 – CX14) and neocortex E12 – olfactory bulb E12 (CX12 – OB12) conditions revealed distinct differentially accessible regions (DARs). Using the MEME algorithm, we identified ungapped motifs enriched within these DARs, with recurring patterns significantly appearing across all conditions. Further motif scanning with *FIMO* revealed that SALL1.H12CORE.0.P.C is significantly predicted to bind to these motifs (p-value < 2e-04). This TF binding prediction was made using the HOmo sapiens Comprehensive Model Collection v12 (*HOCOMOCO*) database.

We became very interested in *Sall1* because it stands out as a strong candidate for a master TF regulator of neurogenesis for several reasons. First, it is significantly predicted to bind to all the selected enriched motifs in mouse. Second, *Sall1* is consistently more highly expressed in mouse IN-enriched regions compared to DN-enriched regions (p.adj: NCx E12 vs NCx E14 = 1.06E-20, NCx E12 vs OB E12 = 5.41E-05), and almost in Chick (p.adj: Lateral E4 vs Medial E4 = 0.08; Figure 11). Third, MEME *FIMO* prediction of *Sall1* (Grant, Bailey and Noble, 2011) predicts its binding to the *Robo1* intron that is differentially more accessible in NCx E12 than NCx E14 (chr16:72699073-72700674, Log2FC = 1.07, Padj = 1.62E-06), suggesting its role in regulating this

key gene, which is crucial for balancing neurogenic modes (Cárdenas *et al.*, 2018). Furthermore, it has been previously reported that the absence of *Sall1* in mice leads to a reduction in both the surface area and depth of the cerebral cortex at E18.5, a consequence linked to altered progenitor cell properties during development. In early cortical progenitor cells, *Sall1* promotes proliferative divisions over neurogenic ones, while at later developmental stages, it regulates the production and differentiation of IPCs. These structural changes are associated with a shift in progenitor cell behavior, where *Sall1* appears to transition progenitor cells from a proliferative state to a neurogenic one as development progresses (Harrison *et al.*, 2012). This evidence suggests that *Sall1* plays a critical role not only in regulating the balance between different modes of neurogenesis but also in determining the timing of cell differentiation.

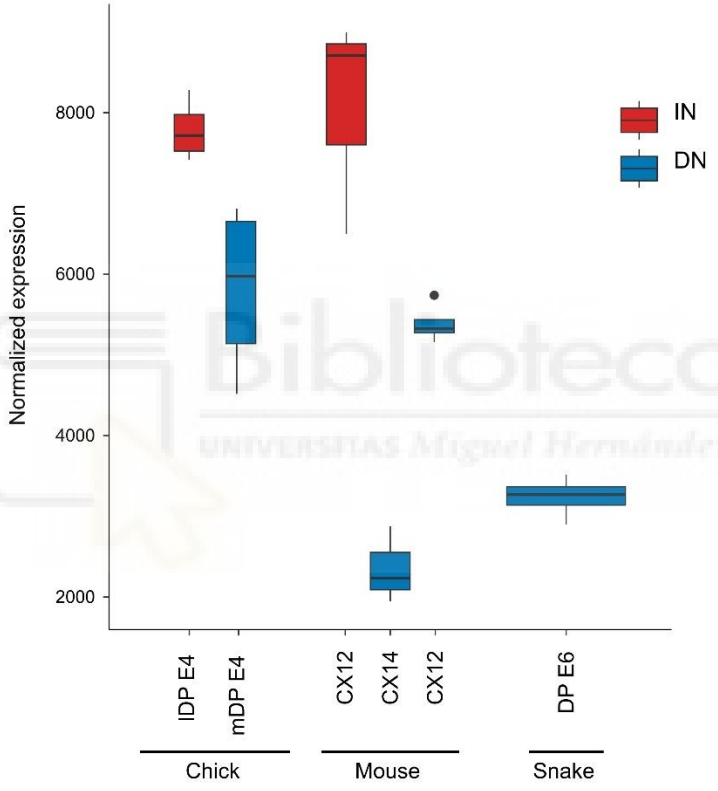


Figure 11. Comparative Expression of *Sall1* in Indirect Neurogenesis (IN) and Direct Neurogenesis (DN) Across Species. This box plot displays the normalized expression levels of *Sall1* in various developmental conditions across species, including chick, mouse, and snake. Red boxes represent *Sall1* expression in indirect neurogenesis (IN) conditions, while blue boxes represent direct neurogenesis (DN) conditions. *Sall1* is significantly more highly expressed in mouse IN-enriched regions (CX12) compared to DN-enriched regions (CX14 and OB12), with p.adjust values of 1.06E-20 and 5.41E-05, respectively. A similar, though less pronounced, pattern is observed in chick (p.adjust: Lateral E4 vs. Medial E4 = 0.08).

Finally, we constructed two network diagrams using the STRING database (Szklarczyk, Kirsch, Koutrouli, Nastou, Mehryary, Hachilif, Annika L. Gable, *et al.*, 2023). The STRING network diagrams were constructed to visualize the functional relationships between genes that meet specific criteria related to gene expression, chromatin accessibility, and regulatory binding sites. By integrating

information from the STRING database, which predicts protein-protein interactions based on known and computationally inferred associations, this approach allows for the identification of potential co-regulatory mechanisms and signaling pathways influenced by *Sall1* in distinct neurogenic regions. Specifically, genes included in the network diagrams were selected based on the following criteria: they are differently highly expressed in either IN-enriched or DN-enriched regions, are associated with DARs that are more open in the corresponding neurogenic context, and contain predicted *Sall1* binding sites within their cis-regulatory sequences in *mm10* identified by *FIMO* (Grant, Bailey and Noble, 2011).

The genes related to IN-enriched regions, which are potentially regulated by *Sall1*, are significantly enriched in those involved in the *WNT* pathway (Figure 12A), such as *Sfrp1, Fzd8/9,* and *Jun*. Crucially, WNT signalling in corticogenesis operates through both the canonical (β-catenin–dependent) and non-canonical (planar-cell-polarity and WNT/Ca^{2+}) branches; accumulating evidence indicates that canonical WNT activity favours aRGC self-renewal, whereas non-canonical pathways are more often linked to cell-cycle exit and neuronal differentiation (Liu *et al.*, 2025). Additionally, other notable genes include *Dmrta1*, a regulator of *Neurog2* and *Pax6* (Noack *et al.*, 2022), and *Irs1*, a key gene in the regulation of the *PI3K/Akt* and *MAPK* pathways (Zheng and Wang, 2021). Another compelling candidate gene that *Sall1* may activate is *Mycn*, which is expressed in aRGCs during cortical development. Overexpression of *Mycn* has been associated with excessive proliferation of IPCs and delayed neuronal migration, leading to postnatal megalencephaly with an increased number of neurons, thus promoting IN mode (Nishio *et al.*, 2023). Furthermore, *Mycn* has been implicated in the decision-making process between asymmetric and symmetric division in neuroblastoma cells, making it a strong candidate for regulating the balance between different modes of neurogenesis (Izumi and Kaneko, 2012).

On the other hand, the genes associated with DN-enriched regions that *Sall1* might regulate are significantly enriched in those involved in axon guidance and synapse formation (Figure 12B), such as N-type calcium channels (*Cacna1b, Cacna1a, Cacna1d*), *Gap43*, and *Grin2b*. While these genes are well-known for their expression in post-mitotic neurons, their roles in aRGCs remain largely unexplored. However, some of these genes have known impacts on cortical progenitor differentiation. For instance, the failure to express *Gap43* during cortical neurogenesis alters cell cycle regulation and differentiation of neural precursors (Mani *et al.*, 2001). These two final gene lists are strong candidates for controlling the balance between the two modes of neurogenesis, all of which may be downstream targets of the master TF regulator *Sall1*.

Further validation of *Sall*1's role as a regulator of neurogenesis is necessary, and applying similar motif discovery approaches in chick DARs will provide crucial insights into its function across species. Since we already have ATAC-seq and RNA-seq data from the chick Lateral (IN) and Medial (DN) Dorsal Pallium at E4, we can leverage this information to determine whether *Sall1* binding motifs are

differentially enriched in these regions. By integrating motif discovery with our existing chromatin accessibility and gene expression data, we can assess whether *Sall1* is differentially regulating neurogenic modes between the Lateral (IN) and Medial (DN) Dorsal Pallium at E4. This would clarify whether its role in neurogenesis is conserved or context-dependent, reinforcing its potential as a master TF governing neurogenic divergence across species.

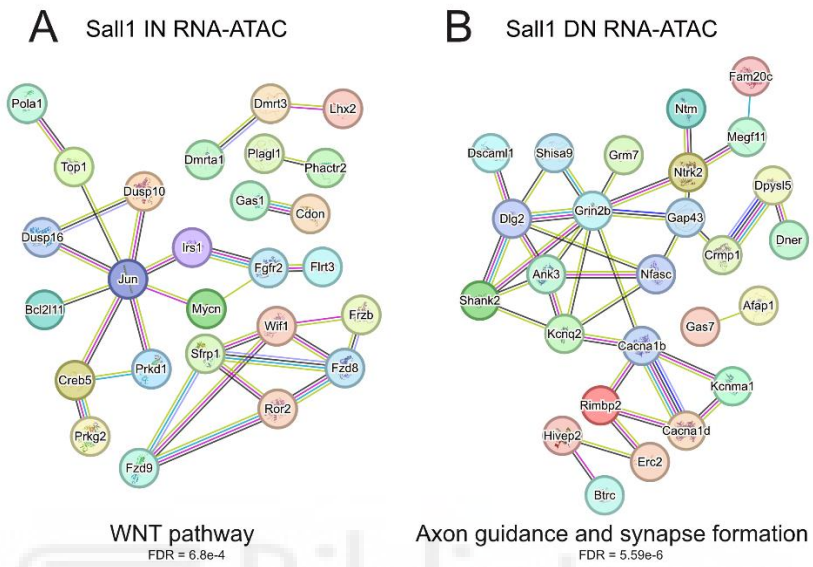


Figure 12. Network Diagrams of *Sall1*-Associated Genes in Indirect and Direct Neurogenesis Regions. A) Network diagram of genes highly expressed in indirect neurogenesis (IN) regions, showing associated differentially accessible regions (DARs) that are more open in IN-enriched regions and containing *Sall1* binding sites in their cis-regulatory sequences. The network reveals significant enrichment in the *WNT* signaling pathway (FDR = 6.8e-4), including genes such as *Sfrp1*, *Fzd8*, *Fzd9*, and *Jun*. **B)** Network diagram of genes highly expressed in direct neurogenesis (DN) regions, with associated DARs more open in DN-enriched regions and featuring *Sall1* binding sites. These genes are enriched in pathways related to axon guidance and synapse formation (FDR = 5.59e-6), including *Cacna1b, Gap43,* and *Grin2b*.

1.3.3. Direct Neurogenesis: *Cux2*, a key regulator of cortical folding and neurogenesis in radial glial cells

A gene traditionally known for its expression in post-mitotic neurons, but whose role in progenitors was hotly debated, is Cut-Like Homeobox 2 (*Cux2*). *Cux2* is a TF commonly associated with dendrite and spine development, as well as synapse formation in projection neurons located in the upper neocortical layers of the mouse brain (Cubelos *et al.*, 2010; Weiss and Nieto, 2019; Miškić *et al.*, 2021). As a result, *Cux2* is frequently used as a marker for these upper cortical layers. However, studies utilizing genetic fate-mapping with *Cux2*-Cre and *Cux2*-CreERT2 mice have revealed that *Cux2* mRNA is also expressed in the VZ and SVZ in a salt-and-pepper pattern, particularly in RGCs with restricted fate potentials (Franco *et al.*, 2012; Gil-Sanz *et al.*, 2015). Strikingly, contrary results were obtained by other labs, making *Cux2* expression in aRGCs a highly controversial issue.

One of the main aims of our laboratory is to understand the cellular and genetic mechanisms leading to primary cortical folding, especially its highly stereotyped patterns that are predefined by transcriptomic protomaps in the cortical germinal zones. To achieve this, in collaboration with Tiwari laboratory, we characterized the spatiotemporal dynamics of gene expression and the active epigenetic landscape (H3K27ac) across prospective folds and fissures during ferret cortical development in an independent research. The analysis identified *Cux2* as one of the top genes with differential mRNA expression between the splenial gyrus (SG) and the lateral sulcus (LS), with also differential H3K27ac patterns. Based on these findings, we hypothesized that *Cux2* plays a crucial role in influencing cortical folding patterns in ferrets by modulating progenitor cell fate (Singh *et al.*, 2024).

To further elucidate the role of *Cux2* in RGC fate and its impact on cortical folding, we studied its expression across specific cell types in the developing ferret cortex. We used a comprehensive scRNA-seq dataset generated from the germinal zones of ferret cortex at two developmental stages (Del-Valle-Anton *et al.*, 2024). Our results showed that *Cux2* was most highly expressed in newborn neurons, but also in IPCs and RGCs (Figure 13A and Figure 13B). Among RGCs, *Cux2* expression was particularly enriched in a recently identified, highly proliferative subtype of RGC named RGα, as well as in its associated mitotic clusters (9 and 22; Figure 13A). Other RGC types, such as truncated RG (tRG), did not express *Cux2*. Expression of *Cux2* was notably restricted to the LS and the OSVZ, particularly in the basal RGα (bRGα), which is critical for cortical folding (Reillo *et al.*, 2011; Nonaka-Kinoshita *et al.*, 2013).

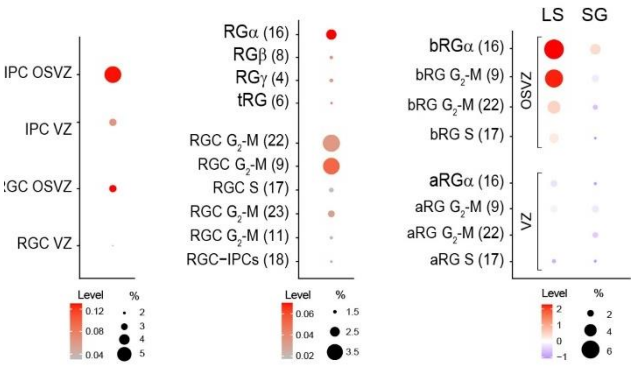


Figure 13 Expression and Role of Cux2 in Radial Glial Cells (RGCs) and Cortical Folding. Dot plot showing *Cux2* expression levels in various progenitor cell populations, including IPCs and RGCs, highlighting the significant expression of *Cux2* in basal radial glial cells (bRGα) within the outer subventricular zone (OSVZ). Expression is particularly prominent in the lateral sulcus (LS) region, linking *Cux2* to cortical folding.

Mining public scRNA-seq datasets from the fetal human cortex (Nano *et al.*, 2023) revealed remarkably similar patterns: *Cux2* is primarily expressed in immature neurons and specifically in a subset of bRG cells (Figure 14).

To determine whether the cortical folding program driven by *Cux2* involves significant transcriptional and cell fate changes, another researcher conducted a single-cell transcriptome analysis on FACS-sorted *Cux2* overexpressing cells from the developing ferret cortex. The results revealed that *Cux2* overexpression led to a reduction in the most immature RGC population and a decrease in the production of IPCs. Concurrently, *Cux2* induced a more committed RGC subtype to proliferate and undergo direct neurogenesis (DN) (Del-Valle-Anton *et al.*, 2024). This interpretation is supported by the previously mentioned scRNA-seq data from ferrets, which showed that *Cux2* is most highly expressed in bRGα and specific mitotic RGC clusters involved in direct neurogenesis (Figure 13B) (Del-Valle-Anton *et al.*, 2024).

Direct neurogenesis initially accelerates neuron production but ultimately results in a reduced overall neuron count and smaller cortex size. The preferential occurrence of DN in the LS aligns with the development of this region into a smaller sulcus rather than a large gyrus. These results confirm the role of *Cux2*, a canonical neuronal gene, in promoting the DN mode by being expressed in RGCs early in cortical development (Del-Valle-Anton *et al.*, 2024).

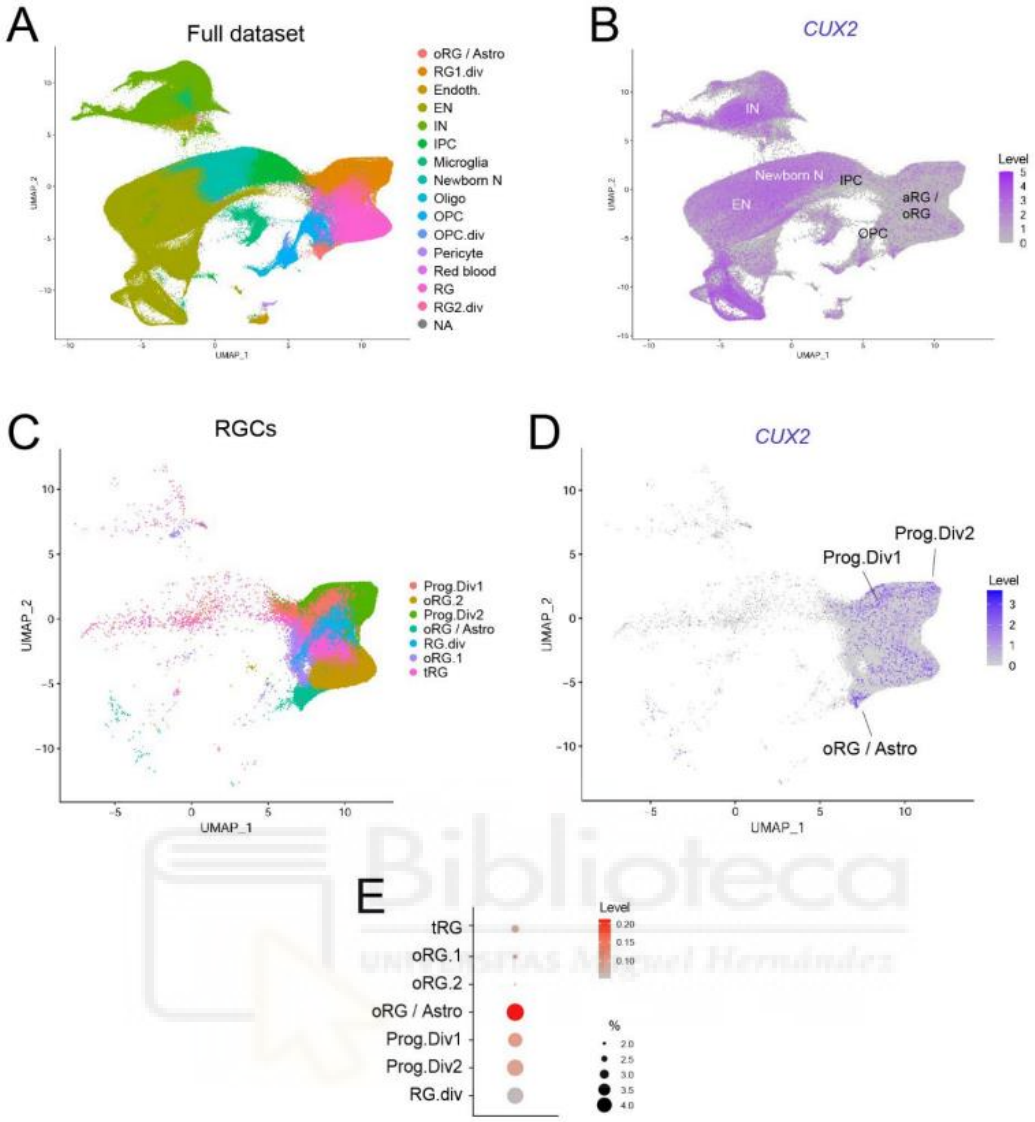


Figure 14. Cux2 Expression in Human Cortical Progenitors During Fetal Development. A) UMAP projection showing cell types in the human fetal cortex, including immature neurons, RGCs, and IPCs. *Cux2* expression is most prominent in immature neurons and a specific subset of basal radial glial cells (bRGs). **B)** Magnified view of RGC populations, highlighting distinct subtypes. *Cux2* express ion is enriched in the proliferative bRG subtypes that are crucial for direct neurogenesis (DN) and cortical folding. **C-E)** Further analyses of *Cux2* expression patterns show that its highest levels are found in specific progenitor subtypes involved in cortical expansion and folding, mirroring findings from the ferret cortex.

1.3.4. Direct Neurogenesis: *Robo* and *Dll1*

Robo1 and *Robo2* are also genes that are traditionally expressed in neurons, specifically to control the guidance of major axonal tracts in the mammalian forebrain (Jaworski, Long and Tessier-Lavigne, 2010). However, previous research demonstrated that in utero electroporation of mouse *Robo1/2* alone in the NCx RGCs results in a mild promotion of direct neurogenesis. This effect is significantly enhanced by concurrently reducing *Dll1* levels (Cárdenas *et al.*, 2018). Interestingly, reducing *Dll1* alone does not impact direct neurogenesis, suggesting a crucial synergistic interaction between *Dll1* and *Robo* signaling. To

further explore the transcriptional changes in RGCs associated with *Robo*-driven direct neurogenesis, we performed in utero electroporation to express mouse *Robo1/2, crDll1*, and GFP in RGCs of E12.5 mouse NCx, and analyzed the bulk transcriptome of these cells one day later to measure the effect of the genetic modification in neural progenitors (Figure 15A, Figure 15B).

The transcriptomic profile of GFP+ cells from embryos expressing *Robo1/2+crDll1* was significantly altered compared to control GFP embryos, with 1,533 differentially expressed genes (DEGs; padj<0.05; Figure 15C, Figure 15D). Although the majority of DEGs were downregulated, the most significantly altered and highest fold-change genes were upregulated (Figure 15D, Figure 15E). GO analysis of the upregulated genes indicated an enrichment in categories related to cell cycle, cell division, and negative regulation of proliferation, which aligns with increased direct neurogenesis (Figure 15F). In contrast, downregulated genes were enriched in categories associated with neuron maturation, axon development, neurogenesis, and *WNT* signaling pathway, the latter suggesting a reduction in progenitor cell self-renewal (Chenn and Walsh, 2002).

Gene set enrichment analysis (GSEA) (Mootha *et al.*, 2003; Subramanian *et al.*, 2005) revealed similar patterns. Upregulated genes were particularly enriched in terms related to mitotic checkpoints and cell cycle arrest, including *Cdkn1a*, *Rps27l*, *Ccnd1*, and *Phlda3* (Galderisi, Jori and Giordano, 2003; Kawase *et al.*, 2009; Xiong *et al.*, 2014; Karimian, Ahmadi and Yousefi, 2016). Additionally, genes involved in mitochondrial function, specifically oxidative phosphorylation (OxPhos), such as *Dglucy*, *Sesn2*, *Cox6b2*, and *Ndufa12*, were also enriched (Mckenzie and Ryan, 2010; Katane *et al.*, 2018; Kovaleva *et al.*, 2020; Nie *et al.*, 2020; Saha *et al.*, 2022) (Figure 15F, Figure 15G, Figure 15H). The mitochondrial metabolic state of neural progenitor cells is crucial for determining their fate, with increased oxidative phosphorylation being linked to cell cycle exit and neurogenesis (Arrázola *et al.*, 2019; Khacho, Harris and Slack, 2019; Iwata and Vanderhaeghen, 2021; Andrews and Pearson, 2023). Conversely, downregulated genes were enriched in GSEA terms associated with synaptic function and axon guidance (Figure 15G). Notably, the top upregulated gene, *Cdkn1a* (*p21*), is known for its role in arresting cell cycle progression by inhibiting cyclin-dependent kinases (Karimian, Ahmadi and Yousefi, 2016). In fact, *p21* is directly repressed by oscillatory *Hes1* expression, suggesting that depending on its expression dynamics, *Hes1* differentially controls neuro steam cell proliferation via *p21* cell cycle exit (Maeda *et al.*, 2023). Altogether, the overexpression of *Robo1/2* and *crDll1* in the embryonic mouse NCx induced widespread transcriptomic changes that elevated mitochondrial OXPHOS activity and promoted cell cycle arrest, ultimately leading to increased direct neurogenesis.

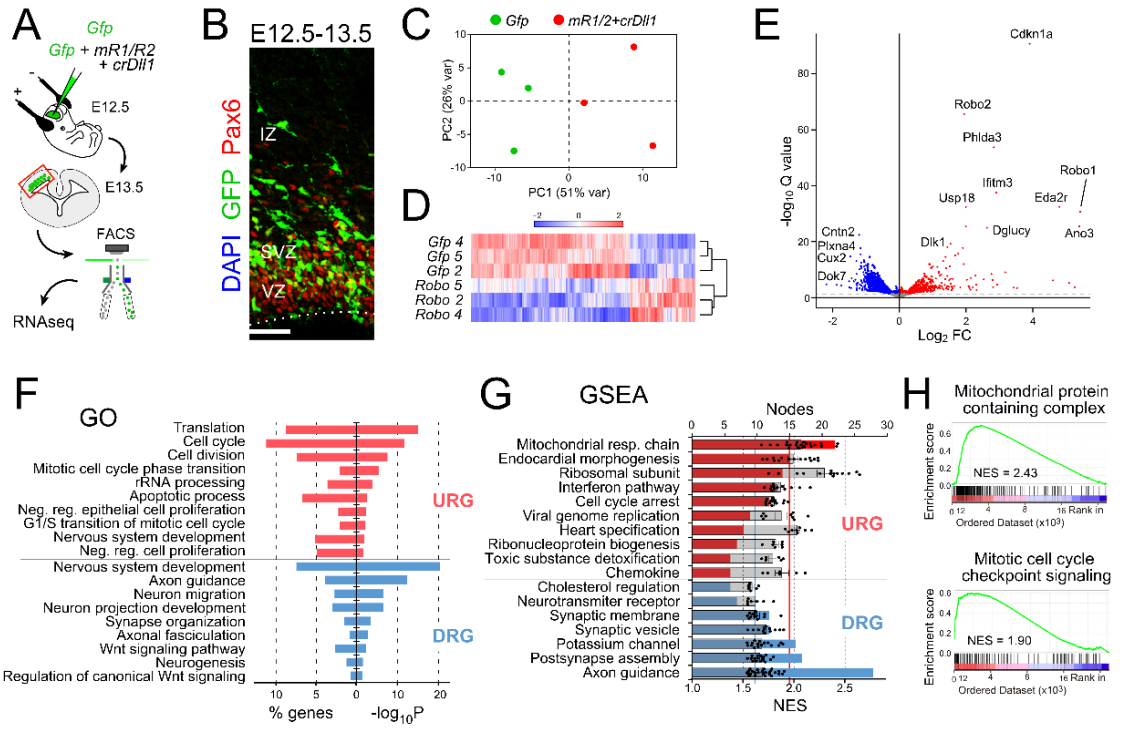


Figure 15. Transcriptomic Changes in Cortical Radial Glial Cells (RGCs) Following Overexpression of *Robo1/2* and *crDll1*. A) Schematic of in utero electroporation experiment to express *Robo1/2*, *crDll1*, and GFP in radial glial cells (RGCs) of E12.5 mouse neocortex (NCx), followed by FACS sorting and RNA-seq analysis one day later. **B)** Immunostaining of the NCx at E12.5–13.5, showing GFP+ cells in the ventricular zone (VZ) and subventricular zone (SVZ). *Pax6* expression highlights the progenitor regions. **C)** Principal component analysis (PCA) of transcriptomic profiles from GFP and *Robo1/2+crDll1* expressing RGCs, showing distinct separation between the conditions. **D)** Heatmap of differentially expressed genes (DEGs) in GFP+ cells following *Robo1/2+crDll1* expression, with significant upregulation of several cell cycle-related genes. **E)** Volcano plot showing the distribution of DEGs, with significant upregulation of genes like *Cdkn1a*, *Phlda3*, and *Dglevy*, and downregulation of axon guidance-related genes such as *Robo1* and *Robo2*. **F)** Gene Ontology (GO) analysis of upregulated (URGs) and downregulated genes (DRGs), highlighting enrichment in cell cycle, negative regulation of proliferation, neuron migration, and axon guidance categories. **G)** Gene set enrichment analysis (GSEA) showing upregulated terms related to mitochondrial function (OXPHOS) and mitotic checkpoints, while downregulated terms include axon guidance and synaptic function. **H)** GSEA plots illustrate the enrichment of genes related to mitochondrial protein-containing complexes and mitotic cell cycle checkpoint signaling in *Robo1/2+crDll1* expressing RGCs.

1.3.5. Public web portal to explore scRNA-seq datasets

The previously mentioned scRNA-seq data from the developing ferret brain cortex has proven to be a valuable resource in *Cux2* research. This dataset has enabled us to precisely identify the specific layers and cell types expressing *Cux2*, leading to the hypothesis that *Cux2* RGCs are involved in DN. Given the significant utility of this dataset, we developed a public web resource using *RShiny* (Ouyang *et al.*, 2021; Chang *et al.*, 2024). This resource is designed to make data accessible to researchers who may not have expertise in

bioinformatics, allowing them to explore candidate genes associated with different modes of neurogenesis.

This *RShiny* application is the foundation of a larger project we initiated, called *CortEvo* (Del-Valle-Anton et al., 2024). *CortEvo* is envisioned as a comprehensive repository where we will integrate both proprietary and publicly available scRNA-seq data on cortical development across various species. The primary goal is to enable users to explore gene expression patterns across different cell types, developmental stages, and species, providing a broader perspective on cortical development through an evolutionary lens.

Prediction of new mammalian membrane transcription factors

2. Unveiling a non-canonical class of transcription factors in mammals

We have seen that several transmembrane receptor proteins such as *Robo* or *Dll1* regulate the modes of cortical neurogenesis across phylogeny. However, how these membrane proteins regulate the expression of genes remains unclear. Recent studies have identified an infrequent non-canonical class of TFs, so-called membrane-bound TFs (MTFs) (Seo, Kim and Park, 2008; Liu *et al.*, 2018; Yang *et al.*, 2020). MTFs are proteolytically cleaved at the transmembrane domain (TMD), shedding a cytoplasmic domain (CD). This CD is then shuttled to the nucleus where it exerts TF activity, binding DNA and regulating gene expression (Liu *et al.*, 2018; Greb-Markiewicz *et al.*, 2019). MTFs have been observed in a wide variety of organisms including procaryotes, plants and animals, but very few have been described in mammals so far. This raises the question of the evolutionary conservation, potential abundance and overall transcriptional relevance of MTFs in mammals, especially humans (Seo, Kim and Park, 2008; He *et al.*, 2020; Leng and Zhao, 2020; Yang *et al.*, 2020).

To investigate the potential role and prevalence of membrane-bound transcription factors (MTFs) in mammals, we applied a computational approach using *DeepTFactor* (Kim *et al.*, 2021), a deep learning tool designed to predict TFs. Given that MTFs are an infrequent and non-canonical class of TFs, we aimed to systematically assess their presence among mammalian transmembrane proteins. We analyzed 14,160 transmembrane proteins from human and mouse (UniProt), focusing separately on their cytoplasmic, extracellular, and lumenal domains. In addition, we examined key predictive features associated with MTF identity and assessed the subcellular localization of selected transmembrane proteins in cell lines. This approach allowed us to explore the evolutionary conservation, abundance, and transcriptional relevance of MTFs in mammals.

2.1. Structure and analysis of transmembrane proteins

In order to identify novel candidate MTFs, we analyzed all 14,160 mouse and human transmembrane proteins as identified in Uniprot, a freely accessible database of protein sequences (The UniProt Consortium, 2021). Transmembrane proteins are integral membrane proteins that span the lipid bilayer, and they typically consist of distinct structural regions known as domains, each with specific properties and functions. In our analysis, we classified these domains as follows:

- **Cytoplasmic Domain (CD):** This region is located on the intracellular side of the membrane, facing the cytoplasm.
- **Extracellular Domain (ED):** This portion extends outside the cell into the extracellular environment, where it can interact with ligands, other proteins, or extracellular signaling molecules.
- **Lumenal Domain (LD):** Found in transmembrane proteins of organelles, this domain resides within the lumen of intracellular compartments such as the endoplasmic reticulum (ER) or Golgi apparatus.
- **Transmembrane Domain (TMD):** This hydrophobic region spans the lipid bilayer and anchors the protein within the membrane.

We further distinguished between single-span transmembrane proteins, which contain a single TMD that connects one extracellular/lumenal domain to one cytoplasmic domain, and multi-span transmembrane proteins, which contain multiple TMDs, resulting in a more complex arrangement with alternating intracellular and extracellular/lumenal domains. For multi-span transmembrane proteins, we also considered intermediate domains, which exist between multiple transmembrane regions and may have distinct functional roles (Figure 16A).

These transmembrane proteins are cleaved at the TMD, and this cleavage may have different outcomes regarding the inheritance of this TMD: it may be inherited in full, only part of it, or none (Figure 16B). These different options critically determine their biological function (Lichtenthaler, Lemberg and Fluhrer, 2018). To determine whether retention of the TMD affected the prediction of being a TF, we analyzed the cytoplasmic, extracellular, and lumenal domains together with the full TMD, half of it, or in its absence, using *DeepTFactor* (Figure 16B).

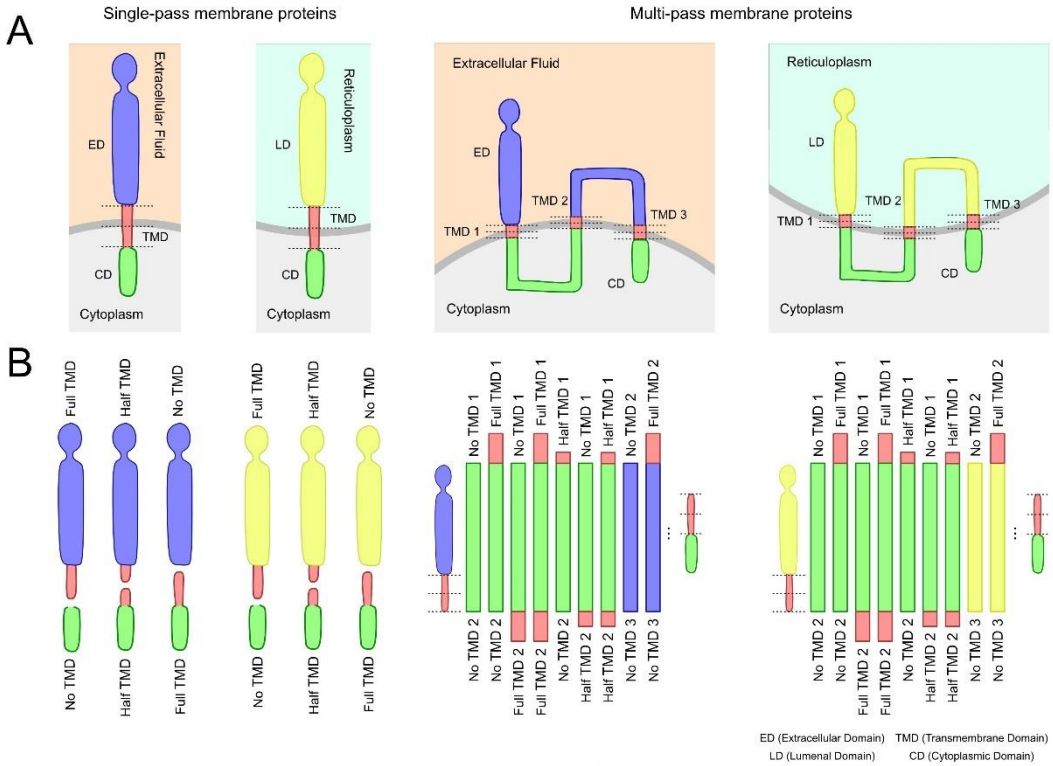


Figure 16. Definition of membrane protein domains tested in this study. Schematic drawings of protein domains downloaded from UniProt and analyzed. **A)** Four types of proteins analyzed: single-pass cell membrane, single-pass endoplasmic reticulum, and multipass cell membrane and endoplasmic reticulum. **B)** Transmembrane domain (TMD) conditions analyzed.

2.2. Hundreds of potential MTFs exist in human and mouse

To maximize the reliability of our analysis, we filtered the UniProt database to select proteins that fulfilled three simultaneous criteria: 1) to have a high quality manually annotated aminoacid sequence (*UniProtKB/Swiss-Prot*); 2) to contain at least one of the domains of interest (cytoplasmic, extracellular, or lumenal); 3) to be a mouse or human protein. This filtering process reduced the number of proteins from a starting total of 20,404 human and 17,137 mouse proteins, down to 7,633 human and 6,527 mouse transmembrane proteins (Figure 17A). The list of filtered proteins was then downloaded as *GFF* file, from which we extracted for each protein the aminoacid coordinates corresponding to the domains of interest: cytoplasmic, extracellular and lumenal. For multi-span proteins, we also extracted the coordinates of the loop domains. Then, we considered three different versions of each of these protein domains: 1) without TMD; 2) with the full TMD; 3) with the proximal half of TMD. For multi-span transmembrane proteins, the loop domains were considered as follows: 1) alone; 2) with the full TMD of both ends, or only one of the ends; 3) with the proximal half of the TMD of both ends, or only one of the ends (Figure 17B). Starting with our above set of transmembrane proteins, this yielded 3,819 human and 3,295 mouse proteins with cytoplasmic domain; 2,929 human and 2,438 mouse proteins with extracellular domain; 660

human and 638 mouse proteins with lumenal domain (Figure 17B). Unfortunately, some proteins do not have domain annotations in Uniprot, which caused a loss of 3807 transmembrane proteins in human, and 3217 transmembrane proteins in mouse.

Next, we downloaded the aminoacid sequences of *UniProtKB IDs* using the *Retrieve/ID mapping* tool in UniProt (Figure 17C), and used *DeepTFactor* (Kim *et al.*, 2021) to calculate, for each protein, the probability score of being a TF (Figure 17D). Proteins were ranked by prediction score, and only those above 0.5 were arbitrarily considered as potential MTFs for further analyses (Figure 17E). Potential MTFs were finally identified by the *UniProt ID* and original gene name (using *Retrieve/ID mapping*) followed by the aminoacid coordinates (Figure 17F). We predicted 144 membrane-tethered transcription factors (MTFs) in humans (about 9% of the ≈1 600 transcription factors encoded by the human genome) and an identical 144 MTFs in mice (likewise ~9% of the ≈1 600 murine TF) (Pratt *et al.*, 2022). Of these, 77 candidates are conserved between the two species, whereas the remainder appear species-specific.

Figure 17. Step-by-step workflow for protein domain analysis. A) Filtering in UniProt to obtain the membrane proteins that contain a domain of interest by species. Starting sample size: 20,401 human, 17,131 mouse proteins. **B)** Filtering to obtain only the domains of interest and transform them into *UniProtKB AC/ID* format. The corresponding transmembrane domain was added. **C)** Download UniProt fasta files. D) Prediction with *DeepTFactor* and interpretation of the result with Integrated Gradients. E) Filtering to obtain only positive predictions ordered by score. F) Annotation of positive proteins with *GeneNames*.

2.3. Impact of TMD on MTF predictions

Our pipeline of MTF prediction revealed that only a small number of transmembrane protein domains are putative ("predicted" from hereon) MTFs (0.31%; Figure 18A), suggestive that *DeepTFactor* is relatively conservative in predicting MTFs. Next, we examined the TF score distribution among predicted MTFs in human and mouse, considering the protein domain composition. Both in human and mouse, the vast majority of predicted MTFs corresponded to cytoplasmic domains, in contrast to a very low abundance among extracellular or lumenal domains. This was true both in terms of absolute and relative abundance (Figure 18B-D). Additionally, their median TF score was not dramatically different between human and mouse (Figure 18E).

Next, we tested the effect of the presence of the TMD on the TF prediction score. Overall, incorporation of either half or the entire TMD caused a reduction in the number of predicted MTFs among cytoplasmic, extracellular, and lumenal domains, although in a minority of cases the median TF score increased (Figure 18B). These findings indicate that the TF score is conditioned by the existence or absence of the TMD. In fact, if the TMD is not cleaved and shed it may critically affect the subcellular localization of the resulting peptide, and hence its function (Lichtenthaler, Lemberg and Fluhrer, 2018). To test this notion, we performed a detailed prediction analysis focused on Cytoplasmic Domains, as we had found these to represent the vast majority of predicted MTFs (Figure 18B, Figure 18C). We observed that the inclusion of the TMD systematically had an impact on the MTF score, which most frequently became smaller (lower likelihood of being an MTF). This was particularly the case with the inclusion of the full TMD, which caused MTF values to often fall below the 0.5 threshold (Figure 18F-H). Intriguingly, examination of MTFs with high TF score revealed that this deleterious effect of TMD inclusion was minimal on proteins with top scores (>0.9), which include MTFs that have been experimentally validated (Figure 18F, G). Remarkably, whereas the inclusion of the TMD (half or full) reduced the TF score for most MTFs (~71% half TMD, ~95% full TMD), for a minority it had the opposite effect, increasing this score (~29% half, ~5% full; Figure 18G). In fact, in a handful of cases the addition of half or full TMD had such a positive effect on the MTF score (>0.5) that led to their inclusion as candidate MTFs, whereas they were not considered in the absence of TMD (Figure 18H).

Predicted MTFs with top scores corresponded to experimentally validated MTFs (see below), confirming the value of our prediction analysis pipeline. For example, the family of ATF6/CREB3 are canonical MTFs, activated by proteolysis upon ER stress thus generating a functional TF that is transported to the cell nucleus. Similarly, MYRF is also a MTF activated by proteolysis in the ER, which promotes central nervous system (CNS) myelination and may regulate oligodendrocyte differentiation (Bujalka *et al.*, 2013; Milan *et al.*, 2020; Fan *et al.*, 2021; Huang *et*

al., 2021). Another validated MTF is XBP1, which plays a vital role in the unfolded protein response (UPR). Under ER stress, unspliced Xbp1 mRNA is cleaved by the activated stress sensor IRE1α and converted to a mature form encoding

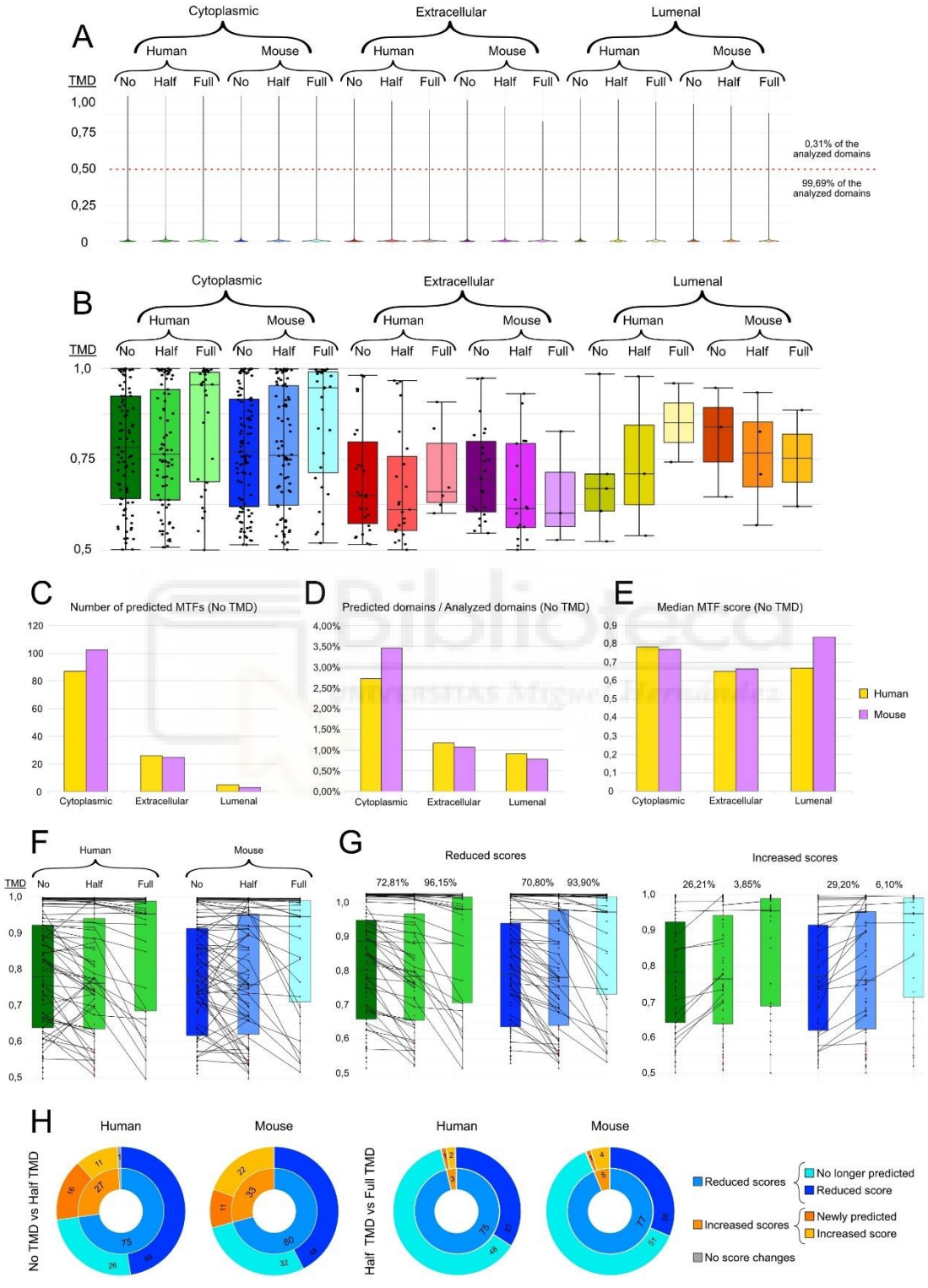


Figure 18. Comparison between conditions in MTF predictions. A) Violin plots with scores for all human and mouse proteins, separated by domain, species, and transmembrane domain. Scores higher than 0.5 are considered MTF candidate. **B)** Box plots of MTF scores for human and mouse proteins >0.5, separated by domain, species, and transmembrane domain as

indicated. **C-E)** Bar plots of the absolute number (C), proportion (D) and median MTF score (E) of predicted MTFs separated by domain and species (only no TMD condition). **F,G)** Box plots of MTF score distribution as in (B), where lines link the scores of the same protein domains depending on the TMD included. In (G), proteins where MTF score decreased with TMD inclusion (left), or increased (right) are distinguished. Percentages indicate the proportion of proteins changed between options. **H)** Amount of protein domains that: 1) reduced or increased their score; 2) were no longer predicted; or 3) were newly predicted as MTF, between the conditions indicated.

spliced XBP1 (XBP1s), which regulates the innate immune response in macrophages (Yoshida *et al.*, 2001; Uemura *et al.*, 2009; Shi *et al.*, 2019; Milan *et al.*, 2020). Other examples of correctly predicted canonical MTFs are Junctophilin 2 (JPH2), Phospholipid Scramblase 1 (PLSCR1) or Sterol Regulatory Element Binding Transcription Factor 1 and 2 (SREBF1, SREBF2) (The UniProt Consortium, 2021). Finally, our analysis accurately predicted both canonical MTFs and recently validated ones, such as Teneurin Transmembrane Protein 1 (TENM1), Insulin Like Growth Factor 1 Receptor (IGF1R), Phospholipid Scramblase 2 (PLSCR2) or Neogenin 1 (NEO1), although these have not yet been annotated in public databases as TFs (Schöler *et al.*, 2015; Kaneko *et al.*, 2016; Aleksic *et al.*, 2018; Tsai and Lee, 2018).

2.4. Protein domain analysis

To characterize which specific amino acid sequences were recognized by *DeepTFactor* to predict transmembrane proteins as TFs, we performed a more detailed analysis on validated MTFs. We focused on four MTFs where structure-function analyses have been performed experimentally. Human XBP1, ATF6A and MYRF proteins have an intracellular domain (ICD) containing a DNA-binding domain (Yoshida *et al.*, 2001; Uemura *et al.*, 2009; Gupta, Read and Gupta, 2015; Choi *et al.*, 2018; Shi *et al.*, 2019; Stauffer *et al.*, 2020; Fan *et al.*, 2021). Functional analyses of truncated versions of these ICDs demonstrate that the DNA-binding domain is most critical for the TF activity of these proteins, as also predicted by *DeepTFactor* (Figure 19A-C). In contrast, the loss of the N-terminus or the link domain in XBP1 had little effect. ATF6A and MYRF have an N-terminus trans-activation domain, and in both cases experimental testing showed that this domain as also key for their TF activity, whereas *DeepTFactor* considers that this is the case for MYRF but not for ATF6A (Figure 19B, C). Finally, we analyzed human PLSCR1, a protein with an ICD that contains a DNA-binding domain plus other relevant regions, such as palmitoylation, proline-rich domain (PRD), and nuclear localization signal (NLS). Experimental studies have demonstrated that the DNA-binding domain of PLSCR1 is important for its TF activity (Zhou *et al.*, 2005; Huang *et al.*, 2020). However, *DeepTFactor* considers this not to be critical, but rather mostly the palmitoylation motif domain, which is key for the nuclear transport (Dal Col *et al.*, 2022) (Figure 19D). Remarkably, the TF activity of the paralog protein PLSCR2 does not depend on the DNA-binding domain but on the palmitoylation domain (Tsai and Lee, 2018).

Our analyses identified the DNA-binding domain as critical for a majority of these MTFs (5/8), whereas for some others the key domains were the transactivation domain, a palmitoylation motif or a MORN (Membrane Occupation and Recognition Nexus) domain (Figure 19E).

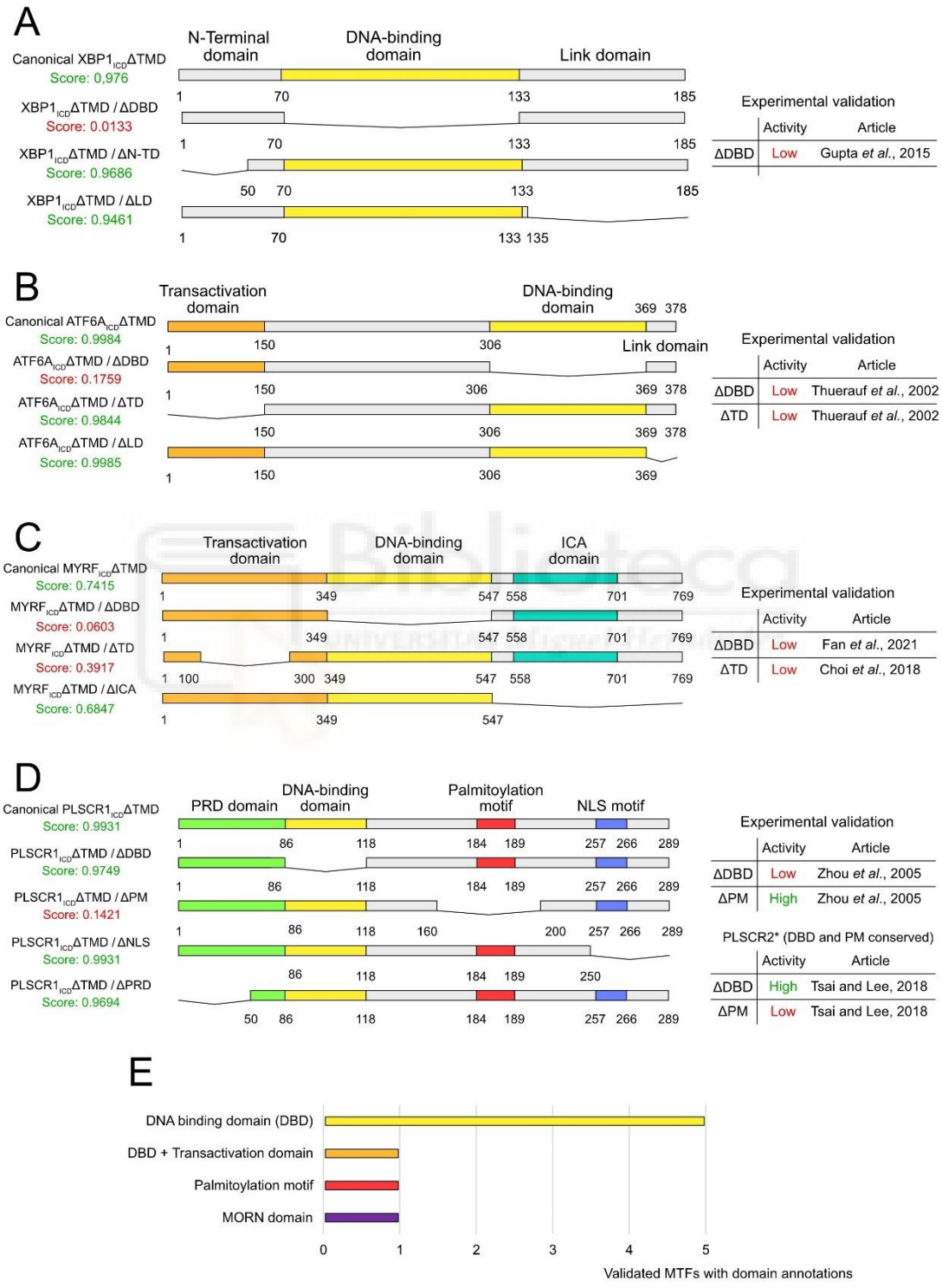


Figure 19. Structure-function analysis of validated MTFs. In silico validation of our analysis pipeline to identify domains critical for TF function for the indicated proteins. **A-D)** Schematic drawings and TF score for each version (full-length, truncation, deletion) of the indicated proteins,

as indicated. Green indicates predicted TF; red, not a TF. Numbers below colored boxes indicate amino acids. The results of the experimental validation of each prediction, indicated as level of TF activity, are shown on the right. **E)** Frequency in which each of the indicated protein domains was considered key for the identification of MTFs by *DeepTFactor*.

2.5. Predicted MTFs localize at the cell nucleus

Our prediction pipeline yielded 211 MTF candidates, most of which were previously unknown. For a protein to exert a function as TF, it must translocate to the cell nucleus to interact with DNA. To increase our confidence on the prediction of novel MTFs, we searched the Human Protein Atlas (HPA) database (Thul *et al.*, 2017) to interrogate the subcellular location of our predicted MTFs, based on immunocytochemistry. We only considered the stains that used antibodies specifically recognizing the domain predicted as MTF (n = 55/144; 38.19% of MTFs). Proteins were considered to have a nuclear localization only if the antibody stain was positive inside the cell nucleus in single confocal plane. As shown in Figure 20A, the canonical TFs SOX2 and ZIC1 clearly localize in the cell nucleus, while ITGB2 and PABPC1, canonical cytoplasmic proteins, are exclusively found in the cytoplasm. The annotation of HPA for our predicted MTFs indicated that 41.81% (n = 23/55) localize to the nucleus (Figure 20B). These results were quite remarkable considering that these were initially produced as membrane proteins. Aside from the proteins annotated as nuclear, we identified an additional number of our predicted MTFs stained in HPA that localize to the nucleus but had not been annotated as such (Figure 20C), corresponding to another 43,64% (n = 24/55) of predicted MTFs. Overall, we determined that ~84% (n = 47/55) of the predicted MTFs examined have a clear nuclear localization, strongly supportive of their potential interaction with nuclear DNA and function as TFs.

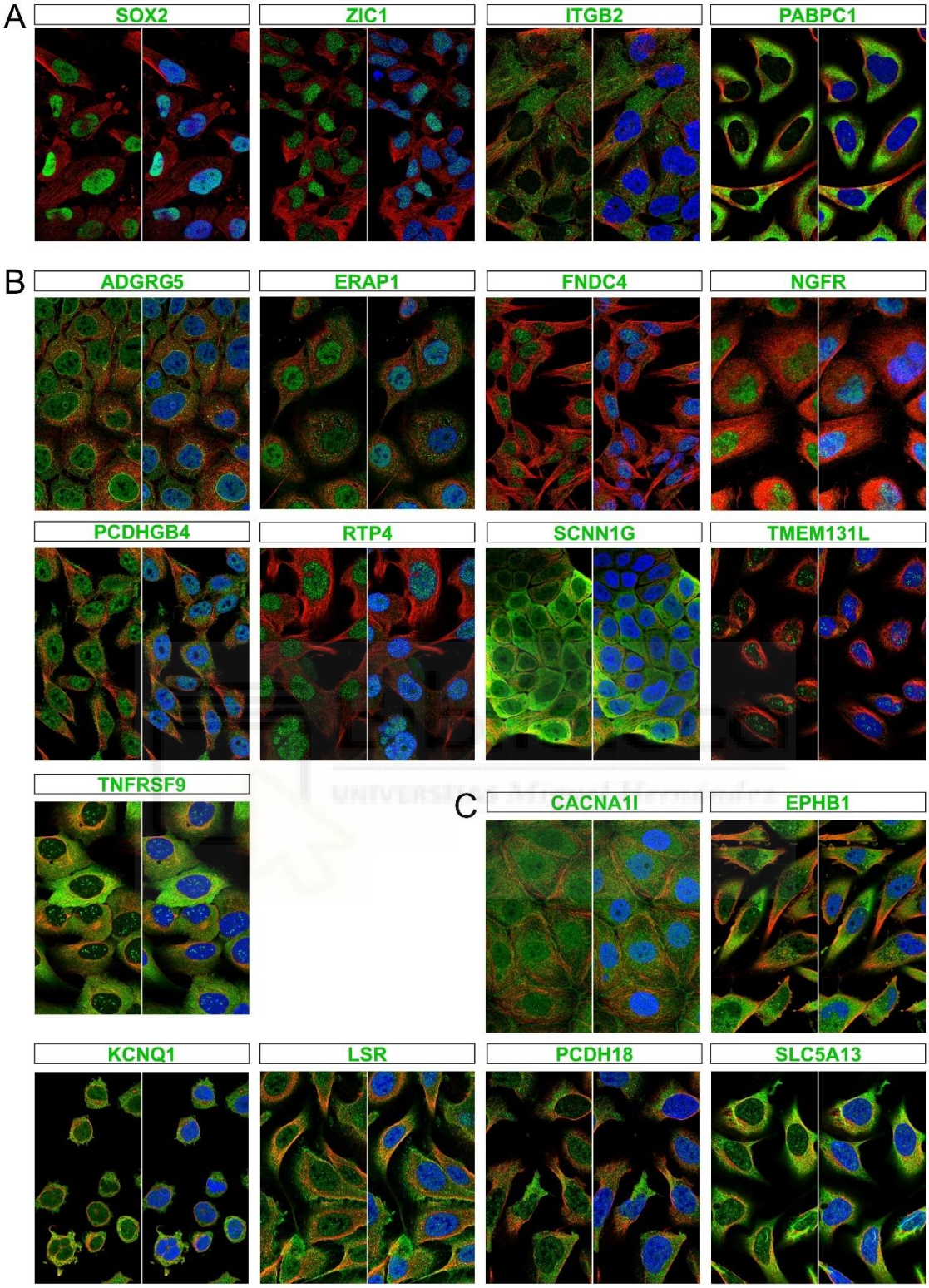


Figure 20. Nuclear localization of predicted MTFs. Single plane confocal images of cell lines stained for the detection of the indicated proteins (green), together with microtubules (anti-tubulin; red) and DNA (DAPI; blue). For each protein, the same image is shown twice to illustrate nuclear localization of the green channel: without, and with DAPI. **A)** Canonical TFs (Sox2, Zic1) and canonical cytoplasmic proteins (Itgb2, Pabpc1). **B)** Predicted MTFs with nuclear localization annotated by HPA. **C)** Predicted MTFs with manual validation of nuclear localization.

2.6. ROBO1 emerges as a strong MTF candidate

We next applied our *DeepTFactor*-based pipeline to the intracellular domains (ICDs) of various transmembrane proteins implicated in neurogenesis, including members of the ROBO receptor and DLL1, as well as other known or putative MTFs (Figure 21). Notably, ROBO1 and ROBO3 ICDs both exceeded the 0.5 threshold in multiple species—human, mouse, and ferret—consistent with the possibility that they bind to the DNA regulating gene expression. Although ROBO1's TF score in human was somewhat lower than in mouse and ferret, it still surpassed the 0.5 cutoff, while ROBO3 showed high score in human and ferret, but low score in mouse. In contrast, ROBO2 and ROBO4 remained near zero in all three species, suggesting limited potential for these paralogs to function as MTFs.

We also evaluated additional transmembrane proteins (DLL1, NOTCH1, NOTCH2, PTPRU, EPHB1, NCAM1, APP, DCC, and NUP107), alongside canonical transcription factors (ZIC2, PAX6, TBR1, SOX2, MYCN, MEIS2). As expected, the known TFs all scored near or above 0.9, illustrating the reliability of *DeepTFactor* in recognizing conventional DNA-binding proteins. Among the transmembrane cohort, DLL1 displayed a consistently low TF score (< 0.1), reinforcing our earlier conclusion that it is unlikely to be cleaved into a nuclear-acting factor. Together, these findings highlight ROBO1 and ROBO3 as strong MTF candidates and further corroborate the specificity of *DeepTFactor* in distinguishing membrane proteins that genuinely harbor nuclear regulatory potential.

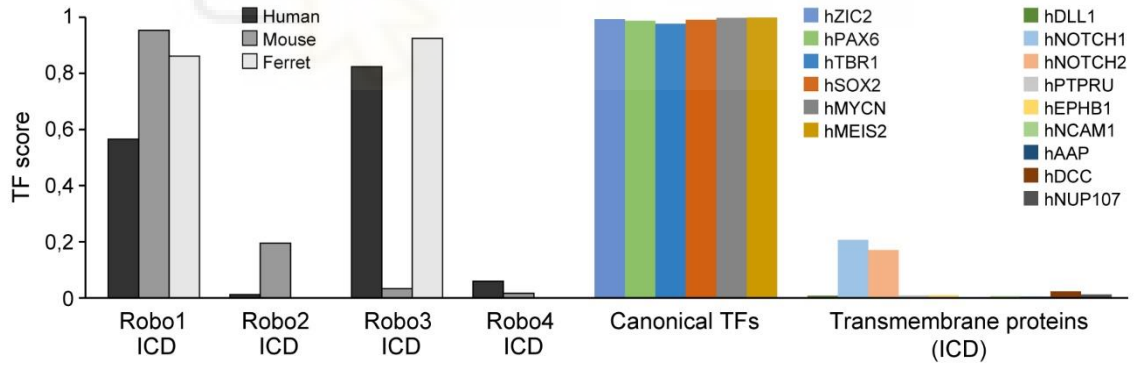


Figure 21. ROBO1 and ROBO3 Are Robustly Predicted to Function as MTFs. Bar plots depict *DeepTFactor* transcription factor (TF) scores for the intracellular domains (ICDs) of ROBO1, ROBO2, ROBO3, ROBO4, and DLL1 across human, mouse, and ferret. ROBO1 and ROBO3 surpass the 0.5 threshold in all three species, indicating strong potential for nuclear activity upon proteolytic cleavage. In contrast, DLL1 remains below 0.1, suggesting it does not function as a membrane-bound transcription factor.

Evaluating the *Rx-Dicer1* mutant mouse as a model for Embryonal Tumor with Multilayered Rosettes

3. Evaluation of *Rx-Dicer1* mutant mouse as a preclinical model for ETMR

Embryonal Tumor with Multilayered Rosettes (ETMR) is a type of brain tumor characterized by the over proliferation of neural progenitor cells, leading to the formation of distinctive rosettes. This abnormal self-amplification of progenitor cells plays a crucial role in the aggressive nature of ETMR. In previous research from our lab, the *Dicer1* mutant mouse was identified as a promising candidate for studying ETMR tumors. To leverage this potential, it is essential first to evaluate the *Rx-Dicer1* mutant mouse as a model for ETMR. This assessment is critical to confirm the model's validity and reliability in accurately representing the human disease. By doing so, we can better understand the mechanisms through which neural progenitor cells transform into tumor cells, ultimately facilitating the development of efficacious therapeutic strategies.

3.1. Transcriptomic changes in *Rx-Dicer1* mutant mouse embryos

Parallel studies in our laboratory have demonstrated that *Rx-Dicer1* mutant mice faithfully recapitulate many of the key histological and molecular features of human ETMRs. Notably, these mutants form hyperproliferative rosettes during neocortex (NCx) and septum (SP) development, which emerge around E12.5 and persist until birth (Fernández *et al.*, 2020). To elucidate the genes involved in tumor rosette formation in these mutants, and potentially in pediatric ETMRs, we performed transcriptomic analyses of *Rx-Dicer1* mutant embryos and wild-type (WT) controls at E11.5 and at E17.5 from both the NCx and the SP.

E11.5 represents a developmental stage that precedes the morphological manifestation of rosettes, providing a critical window for investigating the earliest molecular alterations that may act as triggers for tumor rosette initiation. By analyzing gene expression profiles at this early timepoint, we aimed to capture upstream regulatory events, signaling cues, and transcriptional changes that set the stage for the later emergence of hyperproliferative structures. In contrast, E17.5 corresponds to a more advanced stage of brain development in which rosettes are already well-established and histologically mature. Transcriptomic analysis at this later timepoint allows for the characterization of gene expression programs that sustain or define the rosette phenotype once it is fully developed.

3.1.1. Onset of rosettes (E11.5)

RNA-Seq analysis of protein-coding genes at E11 identified 4,641 DEGs between mutant and control embryos, combining NCx and SP samples using multifactor design (Love, Huber and Anders, 2014) (Padj < 0.05; Figure 22A). Among the most significantly altered DEGs, we identified *Lin28a/b, Prtg, Nr6a1*, and *Greb1*,

which are known positive regulators of cell proliferation and are upregulated in various cancers (Cheng *et al.*, 2016; Sin-Chan *et al.*, 2019; Xiang *et al.*, 2021; Cotino-Nájera *et al.*, 2024).

To analyze the similarity between different tumor types and the *Rx-Dicer1* mouse model, we generated tumor-specific signatures for each brain tumor type using publicly available RNA-Seq data from embryonic and adult brain tumors (Jessa *et al.*, 2019). These signatures consist of DEGs uniquely upregulated or downregulated in each tumor type compared to the remaining brain tumor samples, yielding a distinct set of signature genes. Then, the ETMR gene signature contains genes that are overexpressed or downregulated exclusively in ETMRs relative to all other brain tumors. Among these ETMR-signature genes, several (*Lin28a/b, Prtg, Nr6a1, Greb1*) also show enrichment in *Rx-Dicer1* mutants (Figure 22A) and are known to promote proliferation and stemness (Cheng *et al.*, 2016; Sin-Chan *et al.*, 2019; Xiang *et al.*, 2021; Cotino-Nájera *et al.*, 2024).

Prior research has demonstrated that *Rx-Dicer1* mutants have reduced levels of mature *let-7* miRNAs, leading to overexpression of their target genes (Fernández *et al.*, 2020). A similar pattern emerges in ETMRs, where *let-7* targets remain consistently overexpressed regardless of the patient's genetic mutation/alteration. Importantly, several *let-7* targets (*Lin28a/b, Trim71, Nr6a1, and Greb1*) are overexpressed in *Rx-Dicer1* mutants and belong to the ETMR gene signature, highlighting how *Dicer1* mutation and consequent *let-7* reduction directly contribute to upregulation of these ETMR-associated genes (Cheng *et al.*, 2016; Sin-Chan *et al.*, 2019; Xiang *et al.*, 2021; Cotino-Nájera *et al.*, 2024).

Among the downregulated genes, *Dicer1* itself was significantly suppressed, likely due to the mutation in *Dicer1* that disrupts the maturation of the *let-7* miRNA family, thereby increasing *Lin28a/b* levels and creating a feedback loop that further represses *Dicer1* (Leone *et al.*, 2008). Additionally, *Cdkn2c* (*p18Ink4c*), a cyclin-dependent kinase inhibitor and a critical promoter of cell cycle exit and differentiation (Tourigny *et al.*, 2002; Uziel *et al.*, 2006; Leone *et al.*, 2008; Kirsch *et al.*, 2009), was downregulated. *Olig2*, another downregulated gene part of the ETMR signature, plays a key role in the differentiation of neural precursors into neurons, oligodendrocytes, and astrocytes (Koyama, 2014). Its downregulation has been linked to a shift in tumor phenotype in diffuse intrinsic pontine glioma cells, a highly aggressive pediatric brainstem tumor (Liao et al., 2021).

GO analysis revealed that downregulated genes were primarily associated with neural differentiation and synaptic function. In contrast, the upregulated genes were significantly enriched for functions related to cell proliferation, cell migration, apoptosis, and immune response (Figure 22B). GSEA confirmed an overall activation of immune system pathways, alongside a widespread reduction in neuronal differentiation (Figure 22C). One of the most intriguing pathways identified as upregulated in both the GO and GSEA analyses is the *interleukin*-6 (*IL-6*) signaling pathway. Secreted *IL-6* enhances *AKT* activity in glioblastoma,

forming a positive feedback loop with the *MSI1* (Musashi-1, an RNA-binding protein) and *AKT* pathway (Chen *et al.*, 2016). This mechanism has also been observed in other cancers, such as prostate cancer, where *IL-6* activates the *PI3K/AKT* pathway and regulates *CCNA1*, promoting tumor cell survival (Wegiel *et al.*, 2008). In lymphoma, where both the *IL-6* and *AKT* pathways are similarly upregulated, *IL-6* knockdown in cell lines with acquired resistance to *PI3K/AKT/mTOR* inhibitors like copanlisib or duvelisib has been shown to resensitize these cells, underscoring the therapeutic potential of targeting *IL-6* in cancers resistant to these inhibitors (Kim, Kim and Park, 2019).

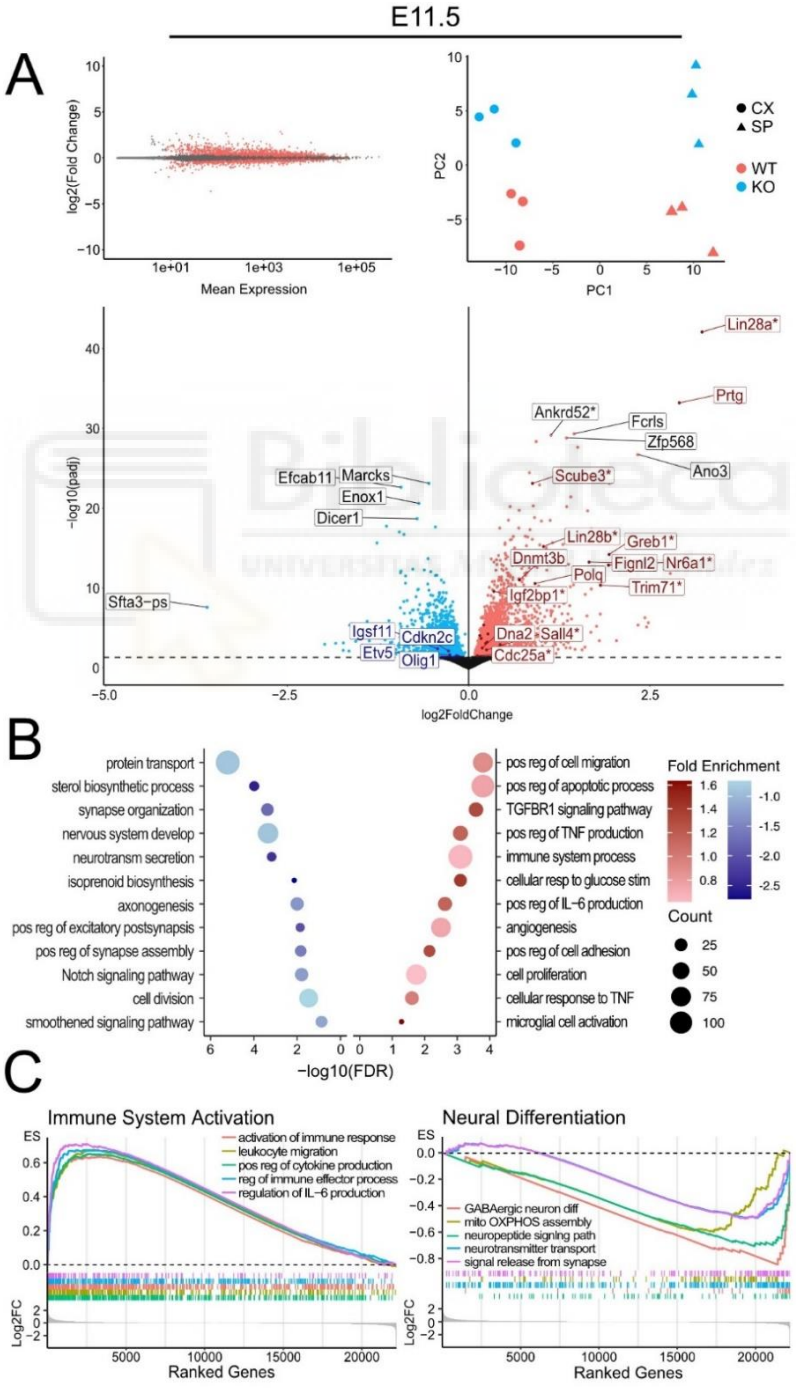


Figure 22. Differential Gene Expression and Functional Analysis in *Rx-Dicer1* Mutants at E11.5. A) RNA-seq analysis comparing *Rx-Dicer1* mutant embryos and wildtype (WT) controls at embryonic day 11.5 (E11.5) revealed 4,642 differentially expressed genes (DEGs; Padj < 0.05). The volcano plot highlights genes that are significantly upregulated (red) and downregulated (blue). The genes highlighted in dark red (overexpressed) and dark blue (repressed) form part of the ETMR gene signature, which include *Lin28a/b, Prtg, Nr6a1*, and *Greb1*, known to regulate cell proliferation and are associated with ETMRs. Genes marked with an asterisk (*) are direct targets of the *let-7* miRNA family. Principal component analysis (PCA) shows clear separation between mutant and WT samples in both the neocortex (NCx) and septum (SP). **B)** Gene Ontology (GO) enrichment analysis of DEGs reveals that downregulated genes are primarily involved in neural differentiation and synaptic function, while upregulated genes are associated with processes such as immune response activation, cell migration, and apoptosis. **C)** Gene Set Enrichment Analysis (GSEA) confirms the activation of immune system pathways in *Rx-Dicer1* mutants and a concurrent reduction in neuronal differentiation pathways.

3.1.2. Mature rosettes (E17.5)

RNA-Seq analysis of protein-coding genes at E11 identified 6,441 DEGs between mutant and control embryos, combining NCx and SP samples using multifactor design (Love, Huber and Anders, 2014) (Padj < 0.05; Figure 23A). The top upregulated genes in *Dicer1* KO embryos included *Prss56, Igf1, Igf2bp1,* and *Pik3c2b*, all known to activate the *PI3K/AKT/mTOR* pathway (Zhang *et al.*, 2018; Margaria *et al.*, 2019; Mancarella, Morrione and Scotlandi, 2021; Li *et al.*, 2023), which promotes cell proliferation and malignancy in cancer (Glaviano *et al.*, 2023). Interestingly, another overexpressed gene within the ETMR signature, *Ccne1*, is dependent on *AKT* activity and stimulates the cell cycle via *Cdk2* activation (Hwang and Clurman, 2005; Au-Yeung *et al.*, 2017). Similarly, *Cdc25a*, also overexpressed at E17.5 and part of the ETMR signature, promotes replication stress, mitotic aberrations, and increases sensitivity to replication checkpoint inhibitors in tumor cell lines (Kok *et al.*, 2020). These upregulted genes in mature rosettes either activate or act as effectors of the *PI3K/AKT/mTOR* pathway, suggesting a critical role for this pathway in rosette formation and maturation. Although many ETMR signature genes and *let-7* targets, such as *Lin28a*, were differentially expressed, they were not among the most dysregulated or statistically significant DEGs (Figure 23A).

Conversely, the downregulation of genes in *Dicer1* KO embryos such as *Nkx2.2, Grid1, Faim2, Dpp10,* and *Gria1* points to impaired neuronal maturation and synapse formation. For example, *Nkx2.2* is essential for the development of specific neuronal subtypes (Kirjavainen *et al.*, 2022), while *Grid1* plays a key role in synapse formation, function, and plasticity (Ung *et al.*, 2024). *Gria1*, a subunit of the AMPA receptor, is crucial for excitatory synaptic transmission (Ismail *et al.*, 2022). These genes are integral to various aspects of neuronal development, including neurogenesis, differentiation, synaptic vesicle trafficking, and receptor function.

Consistent with this, GO analysis of *Rx-Dicer1* DEGs at E17.5 showed that upregulated genes were enriched in processes related to cell proliferation,

genome instability, DNA repair, and epithelial-to-mesenchymal transition (EMT), while downregulated genes were associated with neural differentiation (Figure 23B). GSEA further confirmed the activation of pathways involved in cell proliferation, such as *WNT* signaling, EMT, and stem cell proliferation, alongside a marked decrease in neuronal differentiation (Figure 23C).

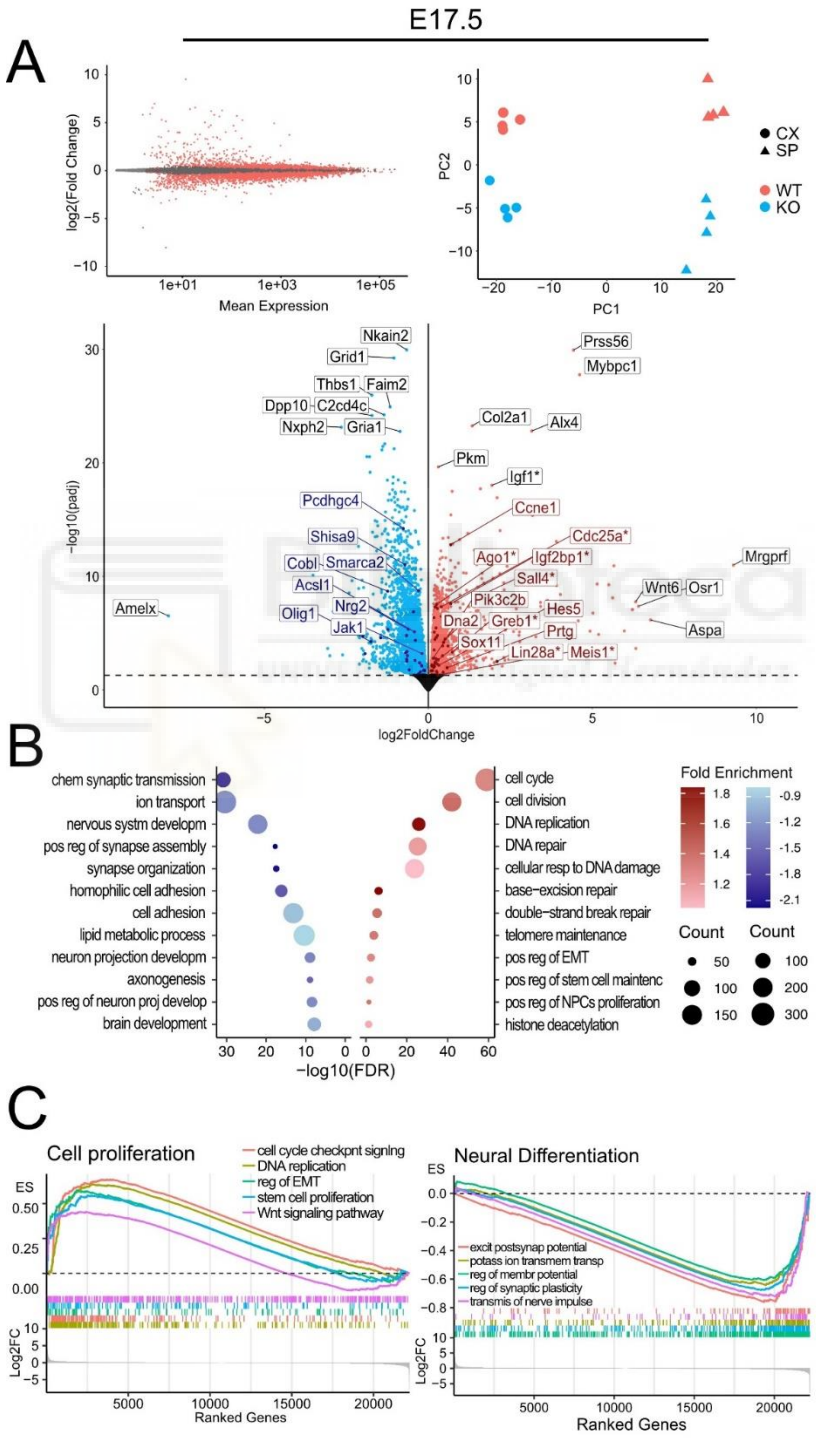


Figure 23. Transcriptomic and Functional Analysis of *Rx-Dicer1* Mutants at E17.5. **A)** RNA-seq analysis of *Rx-Dicer1* mutants at E17.5 compared to controls in both the cortex (CX) and septum (SP) regions identified several differentially expressed genes (DEGs). The volcano plot

highlights significant upregulation (red) and downregulation (blue) of key genes. The genes highlighted in dark red (overexpressed) and dark blue (repressed) form part of the ETMR gene signature. Genes such as *Prss56, Igf1, Igf2bp1, Pik3c2b,* and *Ccne1* all involved in the *PI3K/AKT/mTOR* pathway, were highly upregulated in the mutants. Principal component analysis (PCA) reveals clear separation between wildtype (WT) and knockout (KO) samples. **B)** Gene Ontology (GO) enrichment analysis of upregulated and downregulated genes at E17.5. Upregulated genes are enriched for cell cycle, DNA repair, and epithelial-to-mesenchymal transition (EMT) pathways, while downregulated genes are associated with neural differentiation and synaptic functions. **C)** Gene Set Enrichment Analysis (GSEA) confirms the activation of cell proliferation pathways, including *WNT* signaling, stem cell proliferation, and EMT, along with the suppression of neuronal differentiation pathways. *IL-6* signaling is also significantly upregulated, forming a positive feedback loop with the *AKT* pathway, further driving tumor progression.

3.2. Comparison of *Rx-Dicer1* mutant mouse embryos to ETMR tumors

3.2.1. Whole transcriptome

To assess the relevance of the *Rx-Dicer1* mutant model to human pediatric brain tumors, we performed a comparative transcriptomic analysis using bulk RNA-seq data from Jessa et al. (2019) (Jessa *et al.*, 2019). We also included the previously established ETMR model GBS (Neumann *et al.*, 2017) to determine which model more accurately mirrors human ETMRs. Scaled Log2FC values were computed across the transcriptome for human tumors versus WT fetal brain, and for mouse models versus their respective controls.

Hierarchical clustering based on these full transcriptomic signatures (Figure 24) shows that human ETMR samples cluster most closely with Other Embryonic Brain Tumor type (OTHER_EBT). Notably, the *Rx-Dicer1* model at both E11.5 and E17.5 stages exhibits strong similarity to ETMR, surpassing that of the GBS model, which shows a weaker association. These results suggest that *Rx-Dicer1* provides a more faithful transcriptomic representation of ETMR and may serve as a valuable model for further mechanistic and preclinical studies.

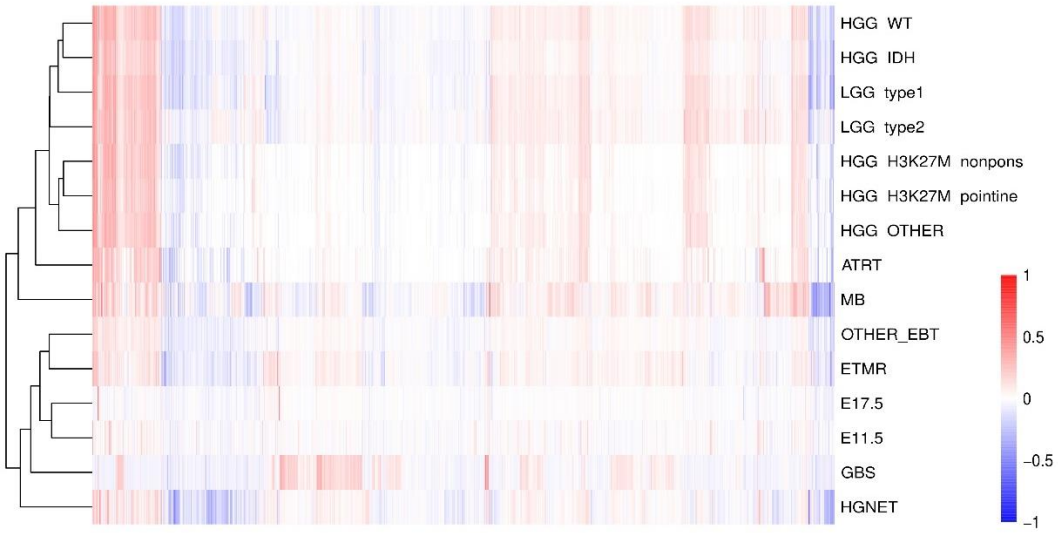


Figure 24. Comparative Full Transcriptomic Analysis of *Rx-Dicer1* Mutants and Human Brain Tumors. Heatmap that compares the full transcriptomes of *Rx-Dicer1* mutant mice (E11.5 and E17.5), the GBS mouse model, and various human brain tumors, including ETMR and high-grade gliomas (HGG), using Log2FC relative to wild-type or healthy controls. Only confident orthologous genes between human and mouse, as identified by Ensembl, were included. The heatmap columns are hierarchically clustered, with the dendrogram at the top illustrating similarities between the samples. The color scale represents the scaled Log2FC, with red indicating upregulation and blue indicating downregulation.

3.2.2. Differentially expressed genes (DEGs)

Having established that the *Rx-Dicer1* mutant mouse model most closely resembles human ETMR based on global transcriptomic profiling, we next performed GSEA to further assess functional similarities. For each tumor type, the top 500 DEGs (tumor vs. controls) were used as gene sets, while the ranked list of Log2FC values of global transcriptome from each mouse model (vs. respective controls) served as the input for enrichment analysis. The Normalized Enrichment Score (NES) represents a critical metric in GSEA, standardizing the Enrichment Score (ES) of gene sets to identify differences or similarities within the dataset. A NES of 0 indicates that the gene set is not statistically significant.

Figure 25A shows that in the top 500 upregulated DEGs gene sets, the *Rx-Dicer1* mouse model at E11.5 exhibits significant similarity to several types of brain tumors, suggesting a more general tumoral phenotype. In contrast, the E17.5 stage shows greater specificity for ETMR gene set, Atypical Teratoid/Rhabdoid Tumors (ATRT) gene set, and Medulloblastoma (MB) gene set. The GBS model, though significantly similar to ETMR, exhibits a lower NES in the ETMR gene set (GBS NES: 1.92, E17.5 NES: 2.13, E11.5 NES: 2.28) and instead demonstrates stronger enrichment for HGNET gene set.

In Figure 25B, we observe that in the top 500 downregulated DEGs gene set, the *Rx-Dicer1* model at E11.5 is highly specific to ETMR and Low-Grade Glioma (LGG) gene set, while the E17.5 stage is statistically similar to almost all types of brain tumors, but clearly more to ETMR gene set than E11.5 (GBS ETMR NES: 0, E17.5 ETMR NES: -2.42, E11.5 ETMR NES: -1.86). Once again, ATRT, ETMR, and MB gene sets are the most NES enriched sets in E17.5. The GBS model does not show significant NES for any of these gene sets and even correlates negatively with LGG type 2 gene set.

To simplify the analysis, we combined the upregulated and downregulated heatmaps and their respective NES values to create a comprehensive plot, shown in Figure 25C. This plot demonstrates that, overall, both the E11.5 and E17.5 *Rx-Dicer1* models are more similar to ETMR gene set than GBS when considering the top 500 upregulated and downregulated DEGs for each tumor type (GBS mean NES: 1.92, mean E17.5 NES: 4.55, mean E11.5 NES: 4.13). Notably, the E17.5 *Rx-Dicer1* mutant mouse is the most similar to ETMR gene set, and also shows significant similarity to ATRT and MB gene sets. The GBS

model is enriched in various tumor terms, but not specifically for ETMR gene set, primarily due to the lack of significance in the downregulated genes. In conclusion, these findings highlight the *Rx-Dicer1* mutant mouse, particularly at the E17.5 stage, as the most faithful model for human ETMR, with significant parallels to ATRT and MB. In contrast, the GBS model, while showing some similarities, lacks specificity for ETMR due to insufficient downregulated gene significance.

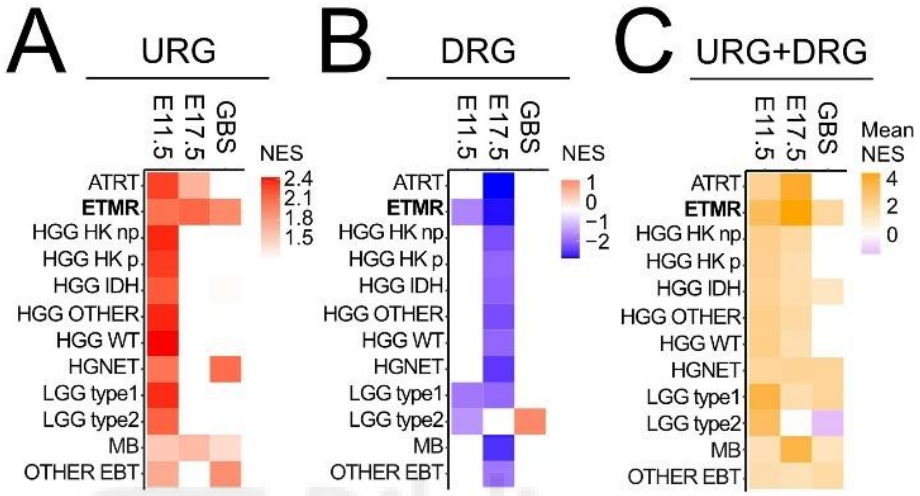


Figure 25. Gene Set Enrichment Analysis (GSEA) of *Rx-Dicer1* Mutant Mouse, GBS, and Human Brain Tumors. A) Heatmap showing Normalized Enrichment Scores (NES) for the top 500 upregulated DEGs (URG) from each tumor type. The *Rx-Dicer1* mouse model at E11.5 demonstrates a general tumoral phenotype across various brain tumor types, while at E17.5, it shows greater specificity for ETMR, ATRT, and MB. The GBS model exhibits a lower NES for ETMR but stronger enrichment for HGNET. **B)** Heatmap showing NES for the top 500 downregulated DEGs (DRG) from each tumor type. The E11.5 *Rx-Dicer1* model is highly specific to ETMR and LGG, while E17.5 shows similarity to almost all brain tumors, however having more similarity to ETMR than E11.5. GBS does not show significant NES for any gene sets and correlates negatively with LGG type 2. **C)** Combined heatmap of upregulated and downregulated NES values (URG + DRG). The E17.5 *Rx-Dicer1* mutant mouse is the most similar to ETMR, with significant similarities to ATRT and MB. In contrast, the GBS model shows some similarities but lacks specificity for ETMR due to the low significance of downregulated genes (GBS mean NES: 1.92, mean E17.5 NES: 4.55, mean E11.5 NES: 4.13).

3.2.3. Comparisons based on tumor signatures

To evaluate how closely our mouse models align with human ETMRs, we examined the expression of ETMR-upregulated signature genes in both *Rx-Dicer1* and GBS mice. Strikingly, the majority of ETMR upregulated signature genes are shared with the upregulated genes in the *Rx-Dicer1* model (marked in red and green), whereas only a small subset are upregulated and shared with the GBS model (blue) (Figure 26A). These findings suggest that the *Rx-Dicer1* model more closely reflects the gene expression patterns characteristic of human ETMRs than the GBS model. Moreover, the most significantly different DEGs, such as *DNMT3B, LIN28A, LIN28B, PRTG, GREB1,* and *IGF2BP1*, were specific

to *Rx-Dicer1* mutants and recognized markers in ETMR, playing key roles in the initiation and maintenance of the tumors (Cheng *et al.*, 2016; Sin-Chan *et al.*, 2019; Xiang *et al.*, 2021; Cotino-Nájera *et al.*, 2024). Among the top 26 upregulated ETMR signature genes, 17 were upregulated in the *Rx-Dicer1* mouse (approximately 65%), while only 6 were overexpressed in GBS (approximately 23%), representing a 2.83-fold increase in *Rx-Dicer1* compared to GBS. These results further supported the validity of *Rx-Dicer1* mutants as a better model of ETMR than GBS.

Regarding developmental stages, Figure 26B shows that the top 26 genes are slightly more similar to E11.5 (13/26) than to E17.5 (11/26). However, when considering the entire signature, E17.5 contains 27% of the signature genes, compared to 17% for E11.5, representing a 1.59-fold enrichment in E17.5. Figure 26C illustrates the similarity of each mouse model's DEGs to the ETMR signature on the left y-axis, as well as their similarity of each mouse model's DEGs to ETMR DEGs on the right y-axis. This figure indicates that E17.5, particularly the SP, is the most similar to ETMR in terms of both the signature and DEGs similarity.

This conclusion is further supported by the GSEA enrichment heatmap (Figure 26D), in which we used the tumor-specific signature genes (DEGs unique to each tumor) as gene sets. Notably, E17.5 displays the highest mean NES in the ETMR signature gene set (GBS mean NES: 0.00; E17.5 mean NES: 3.10; E11.5 mean NES: 0.11). These findings reinforce the notion that the *Rx-Dicer1* mouse model, particularly the E17.5 SP, provides a more accurate developmental and transcriptomic representation of human ETMR than the GBS model.

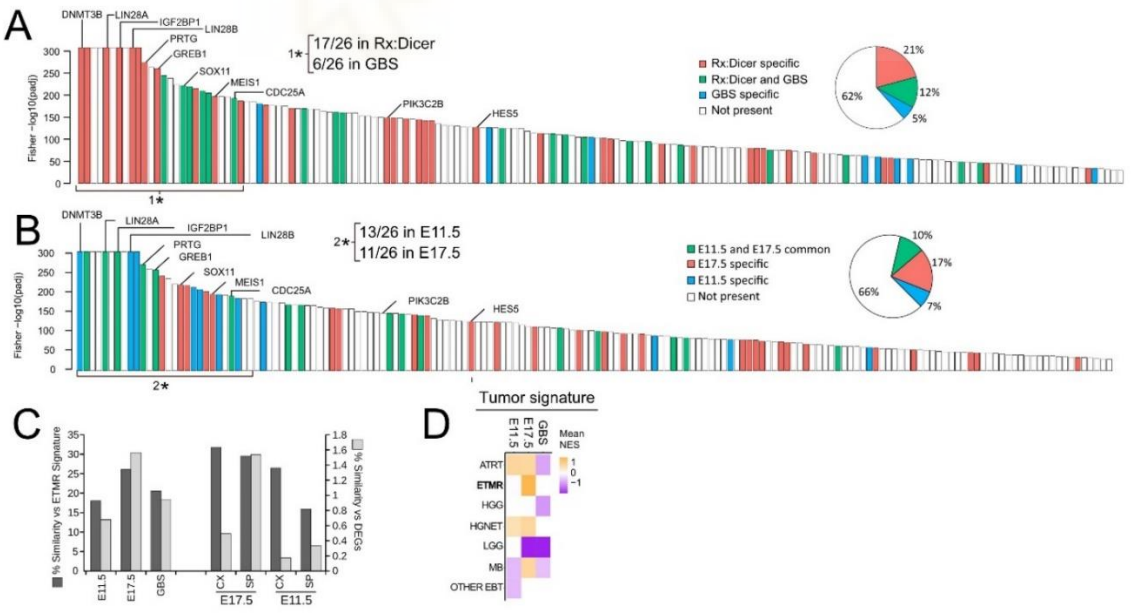


Figure 26. Comparative Analysis of the ETMR Gene Signature Across *Rx-Dicer1* and GBS Mouse Models. A) ETMR signature which comprises 164 genes consistently overexpressed in ETMR tumors compared to other brain tumors. The bar plot focus on the top 26 most significant genes in the ETMR signature, ordered by their Fisher's test p-value. Genes are color-coded: red indicates genes specific to the *Rx-Dicer1* model, green indicates genes shared between *Rx-*

Dicer1 and GBS, and blue represents genes specific to the GBS model. Most of the highlighted genes are either specific to the *Rx-Dicer1* model or shared between both models, with very few being exclusive to GBS. The pie chart shows that 65% (17/26) of the top 26 genes are present in the *Rx-Dicer1* model, compared to 23% (6/26) in GBS. **B)** Breakdown of the ETMR signature in the *Rx-Dicer1* model at two developmental stages (E11.5 and E17.5). Of the top 26 genes, 13 are present at E11.5, while 11 are found at E17.5. When considering the entire 164-gene signature, E17.5 exhibits a higher overlap, with 27% of the genes represented compared to 17% at E11.5, as illustrated in the pie chart. **C)** Bar plot showing the similarity of each mouse model's DEGs to the full ETMR signature (left y-axis) and to human ETMR DEGs (right y-axis). E17.5 shows the highest similarity to the ETMR gene signature and DEGs, particularly in the septum (SP) region, highlighting it as the most faithful model for ETMR. **D)** GSEA heatmap showing the normalized enrichment scores (NES) for the *Rx-Dicer1* and GBS models against different brain tumor signatures (GBS mean NES: 0, mean E17.5 NES: 3.10, mean E11.5 NES: 0.11).

Unraveling the mechanisms that initiate and expand ETMR tumors

4. Mechanisms leading to the initiation and expansion of ETMR tumors

4.1. Single-cell RNA-seq analysis of E12.5 NCx in Dicer1 mutant embryos

To investigate the mechanisms leading to the formation of ETMR, I performed scRNA-seq analysis on NCx tissue at E12.5, a stage where tumor rosettes are beginning to form. Although sequencing was also conducted on NCx at E17.5 and on SP tissue at both E12.5 and E17.5, these datasets were not analyzed within the scope of this thesis due to time constraints.

4.1.1. Cell types and marker genes

After preprocessing the data, we generated a two-dimensional UMAP of the scRNA-seq data from WT and *Dicer1* mutant (KO) samples. This dimensionality reduction enabled the visualization of distinct cellular populations in each condition. Next, the UMAP embedding was clustered into 36 groups to uncover subtle differences between the WT and KO samples (Figure 27). Notably, some clusters appear specific to either the KO or the WT condition, suggesting potentially significant condition-dependent variations in cellular composition.

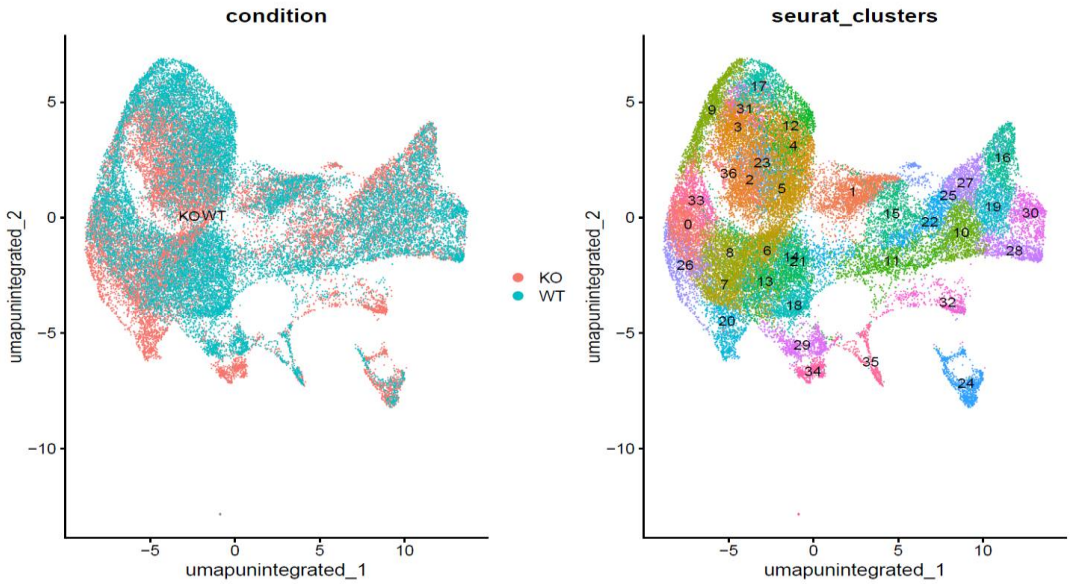


Figure 27. UMAP Visualization of Single-Cell RNA-Seq Data from NCx at E12.5 in WT and *Dicer1* Mutant Mice (KO). Single-cell RNA sequencing (scRNA-seq) performed on NCx tissue from E12.5 embryos, a stage when ETMR tumor rosettes begin to form. The left panel shows a UMAP plot, a two-dimensional representation of this scRNA-seq data, divided by wildtype (WT) and *Dicer1* mutant (KO) samples. Differences between the WT and KO conditions can already be observed, with some regions appearing to be condition specific. The right panel displays the UMAP clustered into 36 groups, which allows for the identification of subtle differences between WT and KO samples, highlighting the distinct cellular populations between the two conditions. These clusters provide insights into the early cellular changes that could contribute to the formation of ETMR.

To identify the different clusters as specific cell types, we utilized canonical gene markers for cortical development, as established in previous studies (Di Bella *et al.*, 2021; Ruan *et al.*, 2021; Dave *et al.*, 2023). Co-expression of *Wnt8b, Crabp2, Hmga2, Dlk1, Prtg, Ccnd1, Igf2bp1*, and *Lrrn1* identified NECs. For RGCs, we used markers including *Hes1*, *Hes5, Pax6, Sox2, Nes,* and *Aldoc*. IPCs were annotated using *Eomes*, *Neurog2, Gadd45g, Btg2*, and *Neurog1*. For early-born neurons, markers like *Neurod1, Tubb3, Nrp1, Gap43*, and *Dcx* were applied, while *Dlx2* and *Dlx1* were used for interneurons. Cycling markers used were *Top2a* and *Mki67*. *Lin28a* served as a tumor marker commonly associated with ETMRs, as its expression is not normally observed at this developmental stage.

Notably, many of the genes associated with specific cell types are clustered together, reinforcing their relevance as strong candidates. This clustering pattern provides further validation of our annotation approach and highlights the distinctive expression profiles that define these developmental stages (Figure 28A). Using these gene markers, we annotated all 36 clusters (Figure 28B). Then we reduced the number of clusters grouped into broader cell-type categories (Figure 29). To visualize the updated annotation, we generated a UMAP displaying all newly classified cell types (Figure 30A, Figure 30B).

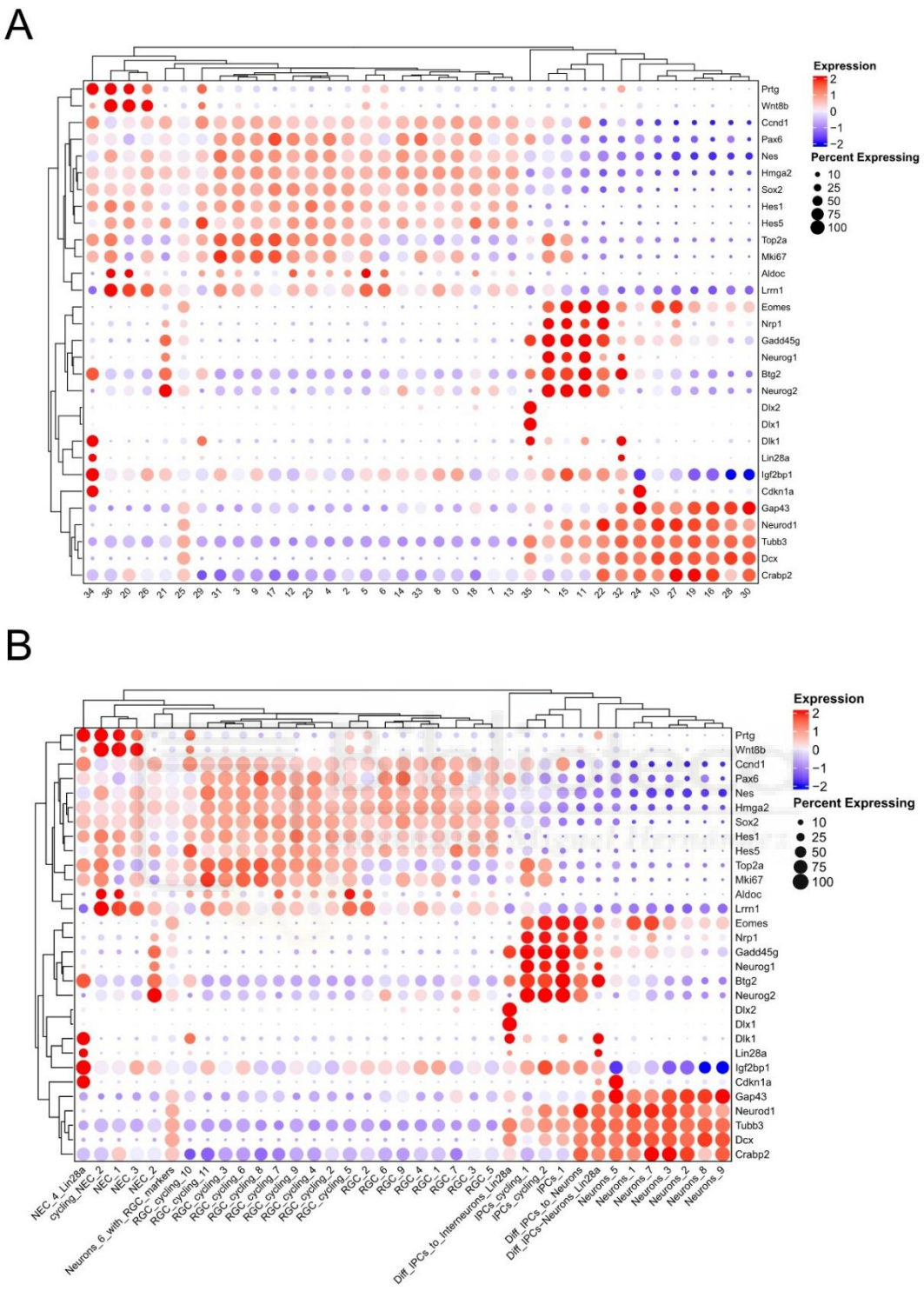


Figure 28. Dot Plot of Canonical Gene Markers for Cortical Development in Clusters. A) Dot plot showing the expression levels of canonical gene markers used to annotate various cell types during cortical development. Each row represents a specific gene marker, and each column represents a cluster from the UMAP. The color intensity indicates the average expression level of the gene in each cluster (ranging from blue for low expression to red for high expression), while the size of the dots represents the percentage of cells within the cluster expressing the gene. **B**) Dot plot displaying the expression of key gene markers across 36 annotated clusters.

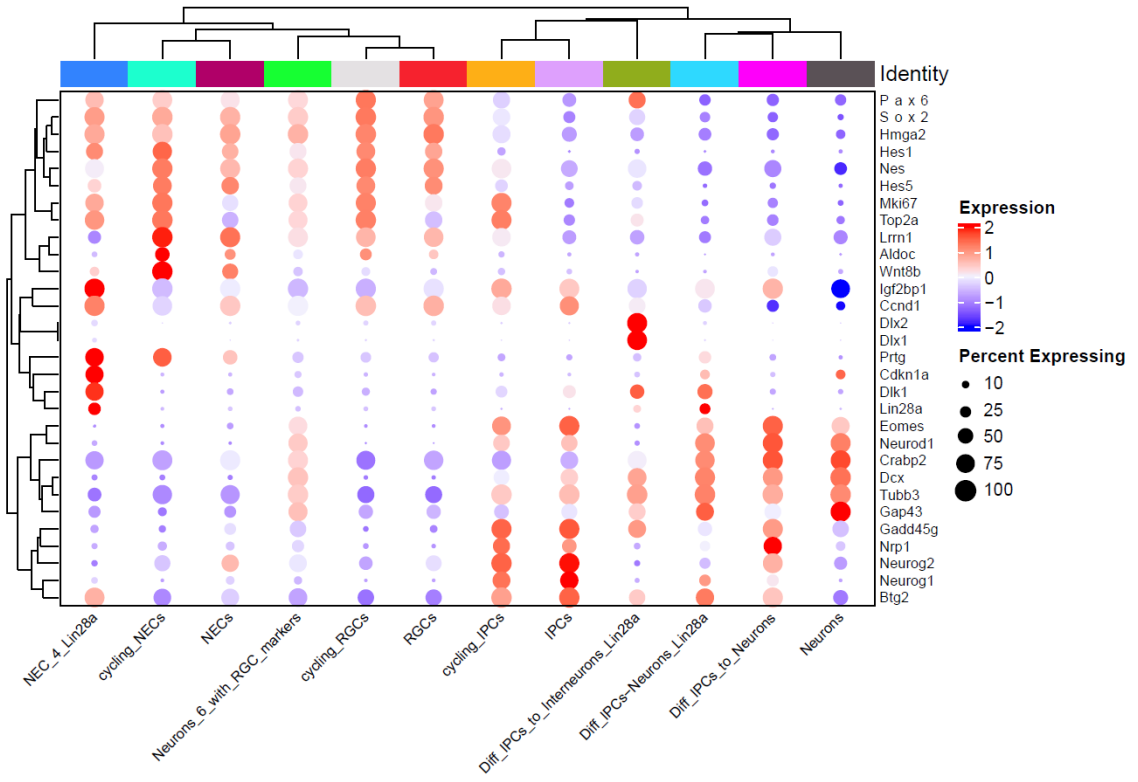


Figure 29. Simplified Cell Type Gene Expression Dot Plot. Dot plot showing the expression of key gene markers, grouped into broader cell-type categories. Dot size reflects the percentage of cells expressing each gene, with color intensity showing the average expression level.

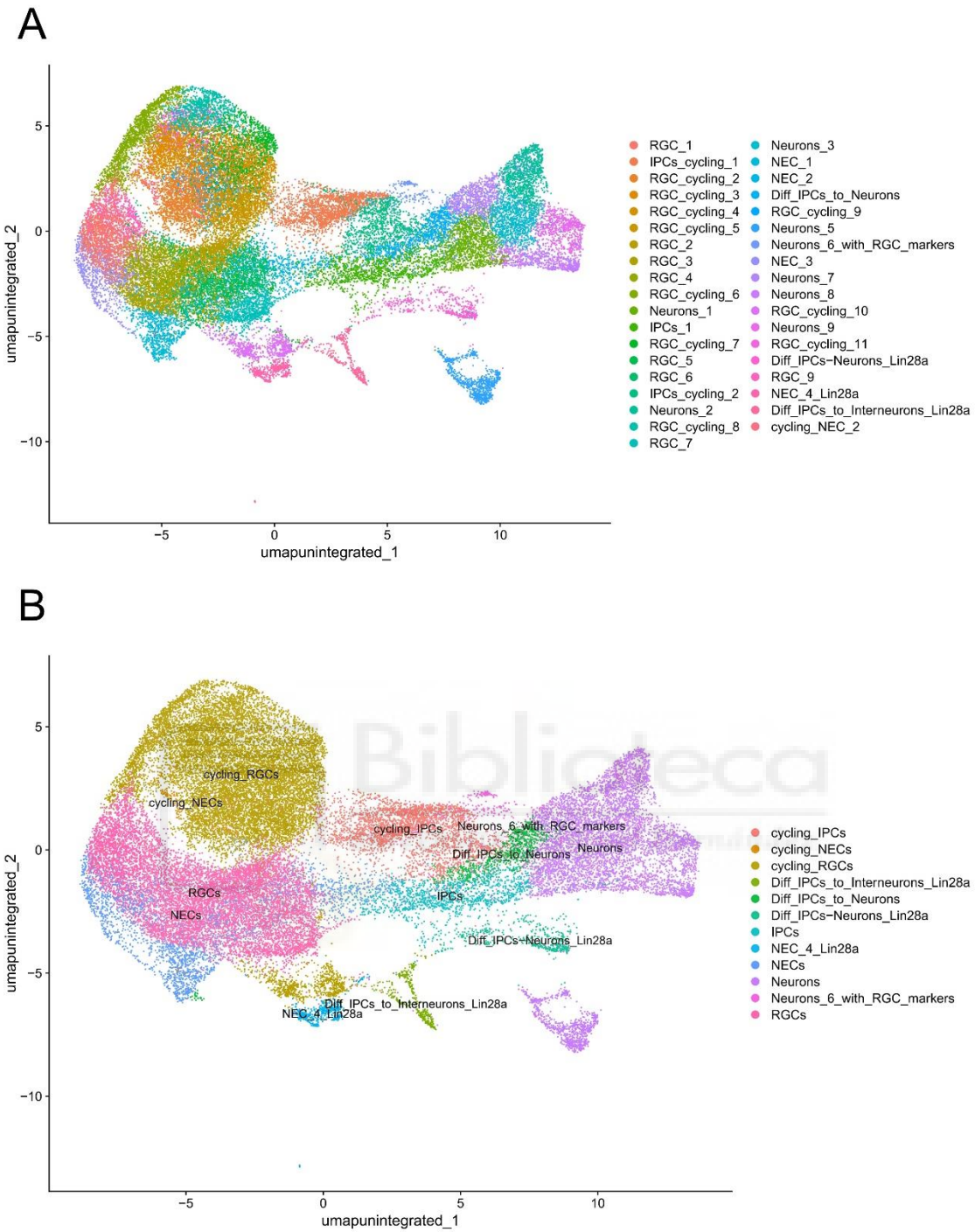


Figure 30. UMAP representation of all annotated cell types. A) The UMAP visualization shows the distribution of cell types across different clusters identified in the analysis. Each cluster is color-coded and labeled with its respective cell type, highlighting various progenitor populations, cycling cells, and differentiated neurons. **B**) This UMAP displays the higher-level groupings of cell types, simplifying the cellular landscape by aggregating clusters into broader categories such as RGCs, NECs, IPCs, neurons, and cycling progenitors.

4.1.2. Tumor-specific cell types

To identify which cell types are differentially abundant between the *Dicer1* KO and WT conditions, we employed the *propeller* method from the *speckle* R package

(Phipson *et al.*, 2022). This approach tests differences in cell type proportions in single-cell RNA-seq data by considering both the number of cells in each cluster and the biological replicates across conditions. Significance was assessed using FDR-adjusted p-values to correct for multiple testing.

Then, we performed a scatter plot of the log2FC in cell type proportions between KO and WT from *propeller*. Interestingly, the top three of the four KO-specific cell types (with positive log2FC) exhibit a NEC signature (Figure 31). Conversely, cell types more abundant in WT relative to KO (negative log2FC) are primarily RGCs (Figure 31). These findings suggest that in the KO mouse, RGCs lose their identity and revert to a NEC-like state, becoming trapped in this undifferentiated phase. The resulting accumulation of NEC-like cells likely drives the formation of highly proliferative rosettes characteristic of the *Dicer1* KO phenotype. In order to visualize this cell type proportions in UMAP, we displayed a overlay heatmap of these log2FC values (Figure 32).

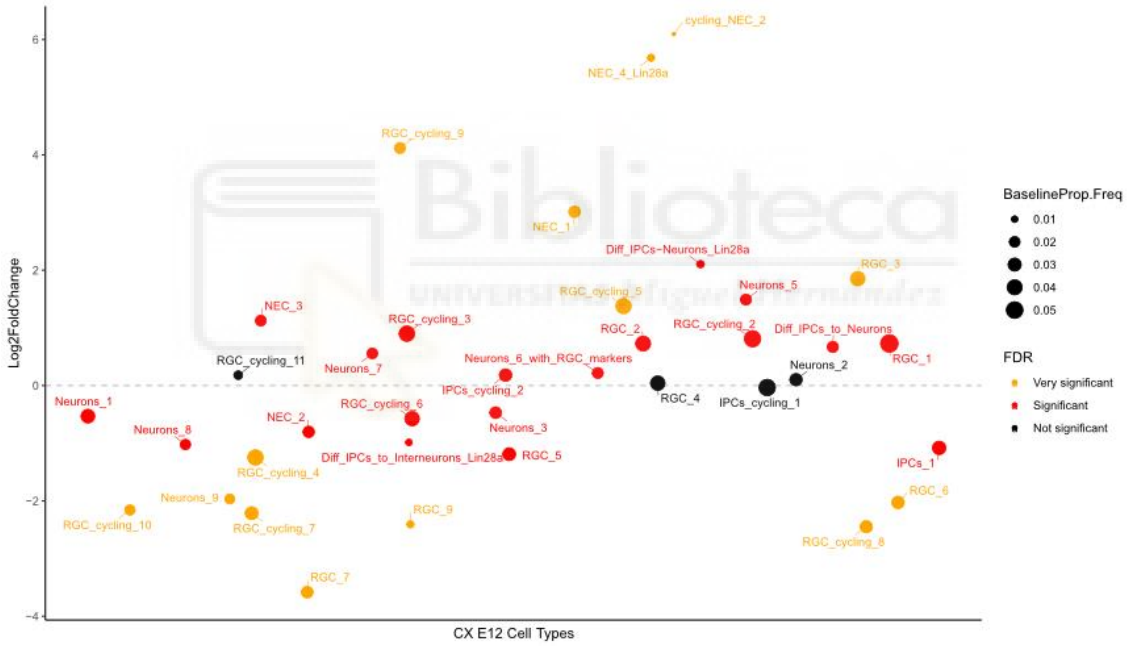


Figure 31. Log2 Fold Change in Cell Type Proportions between *Dicer1* KO and WT Conditions. This plot shows the log2 fold change (log2FC) in cell number proportions between *Dicer1* KO and WT conditions, with each point representing a distinct cell type. The y-axis indicates the log2FC, the x-axis shows the baseline proportion frequency, and the point size reflects this baseline proportion frequency. Points are colored by FDR-adjusted significance levels (Red: very significant, FDR < 1e-50; Orange: significant, FDR < 0.05; Black: not significant, FDR > 0.05). The dashed horizontal line at 0 separates cell types more abundant in KO (above the line) from those more abundant in WT (below the line).

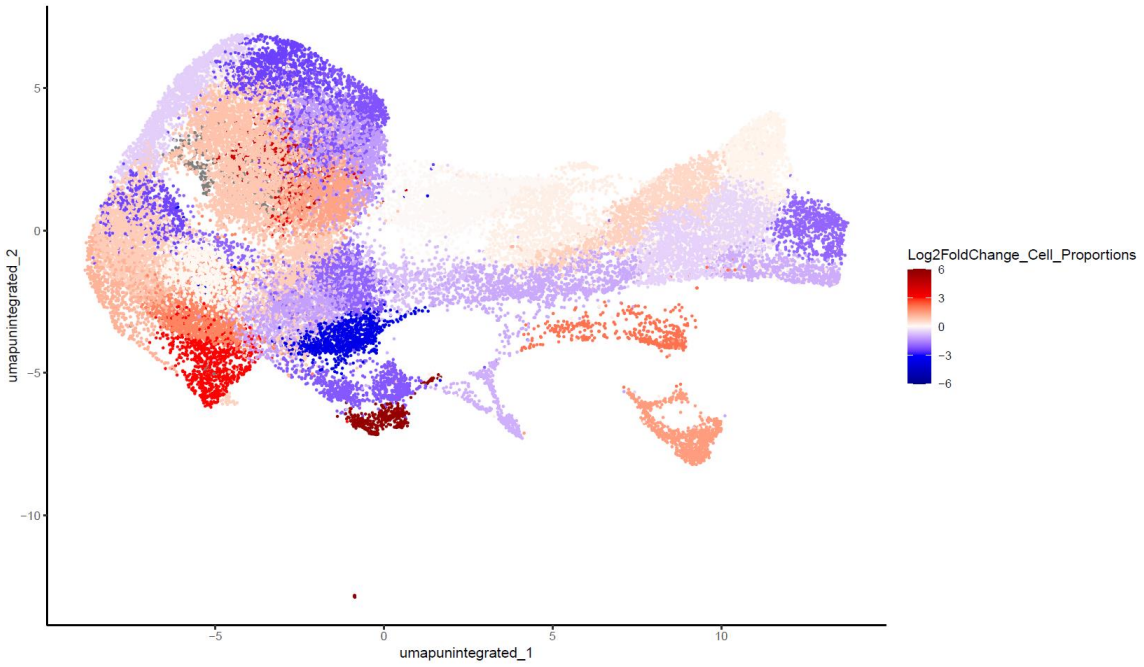


Figure 32: UMAP Representation of Cell Proportion Fold Changes. This UMAP visualizes the Log2FC in cell proportions between KO and WT conditions across the cell clusters. Red areas represent clusters enriched in KO conditions, mainly NEC-related, while blue areas represent clusters more abundant in WT conditions, primarily RGCs.

4.1.3. RNA velocity and transcriptional trajectories

We next aimed to explore the differences in cellular dynamics between WT and KO conditions, particularly focusing on how cell type differentiation trajectories are affected. To achieve this, we employed RNA velocity analysis using *Velocyto* and *scVelo* (La Manno *et al.*, 2018; Bergen *et al.*, 2020). This approach leverages the abundance of spliced and unspliced transcripts to infer the direction and speed of state transitions, thus providing a powerful way to track immature-to-mature RNA transitions and estimate single-cell RNA velocities. Notably, in WT (Figure 33A), RGCs exhibit a full cell cycle RNA trajectory, indicating normal cycling and maturation. In contrast, in KO mice (Figure 33B), the RNA velocity of RGCs suggests incomplete cell cycling, with RGCs either failing to complete the cell cycle or transitioning into NECs without returning to the RGC state. These findings imply that RGCs in KO mice lose their ability to properly differentiate, becoming either stalled in the cell cycle or reverting to an undifferentiated NEC state.

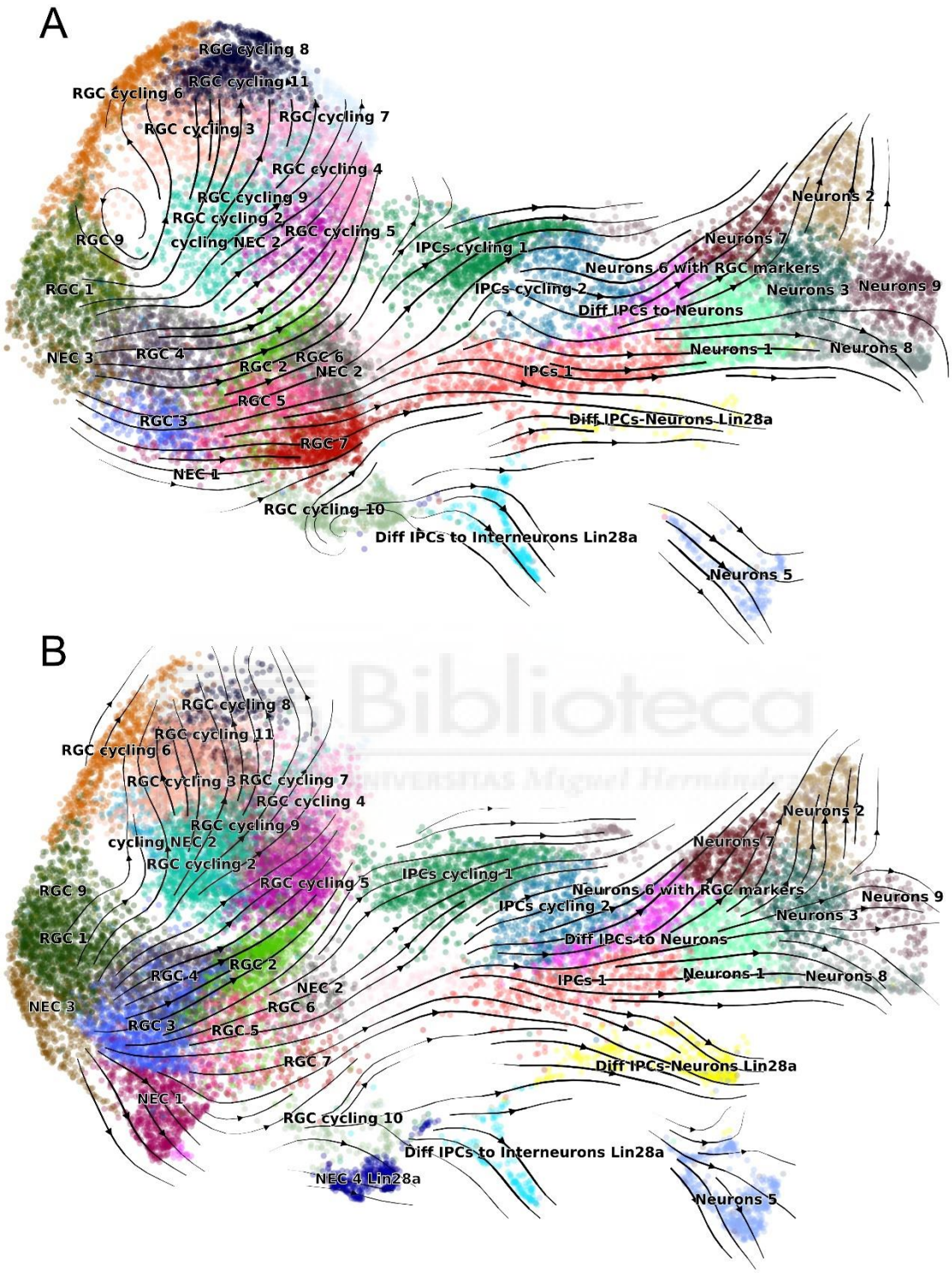
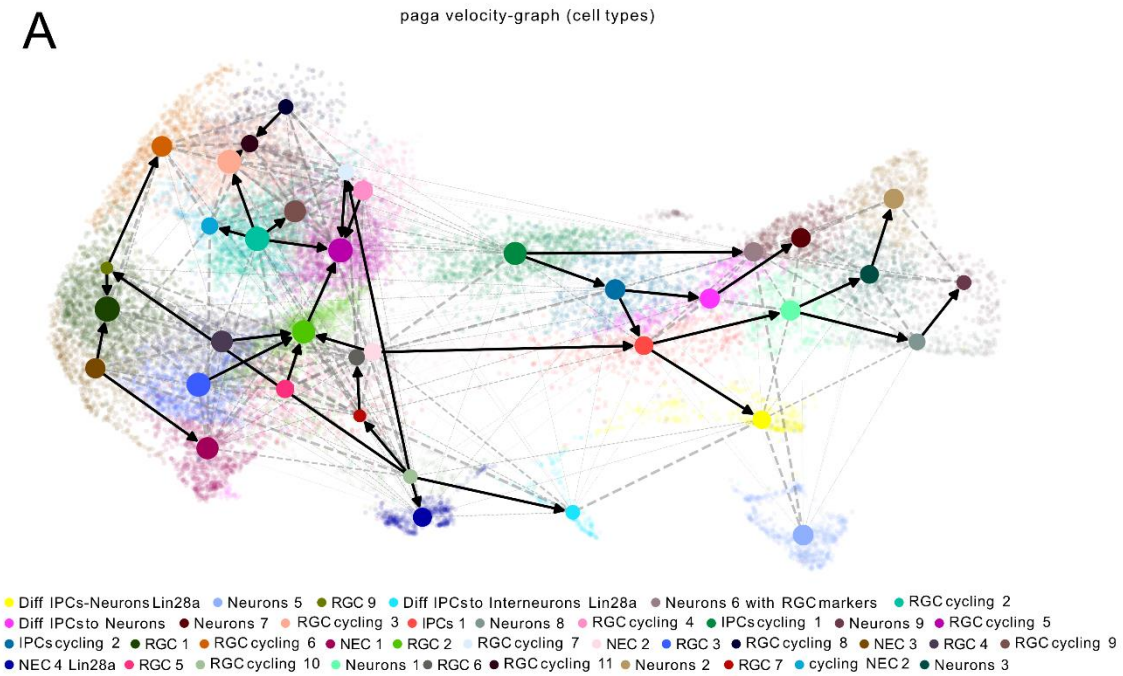


Figure 33. RNA velocity analysis of WT cells. This UMAP visualization shows RNA velocities of individual cells, tracking the dynamic transitions from immature to mature RNA in WT (**A**) and KO (**B**).

We next aimed to examine how global connectivity patterns of cellular states differ between WT and KO conditions, with a particular focus on identifying abnormal cell transitions. To achieve this, we employed Partition-based Graph Abstraction (*PAGA*) (Wolf *et al.*, 2019), which constructs a graph-based representation of single-cell data to highlight potential developmental trajectories and the connectivity between different cell states. By applying *PAGA*, we could better visualize and interpret how cells in the KO condition diverge from the typical differentiation routes observed in WT.

Using *PAGA*, we found that in WT (Figure 34A), *RGC_cycling_6* RNA points back to *RGC_9*, indicating normal cell cycle progression. However, in KO mice (Figure 34B), *RGC_9* RNA instead points toward *RGC_cycling_6*, suggesting an abnormal trajectory. Moreover, none of the trajectories from cycling cells, *NEC_1*, or *NEC_4* in KO mice point backward, indicating that these cells reach a point of no return and ultimately contribute to the formation of overproliferative rosettes. Additionally, in WT, *NEC_3* differentiates exclusively into RGCs, whereas in KO, *NEC_3* transitions into the large KO specific *NEC_1* cluster. This observation further supports the notion that RGCs in KO remain in an undifferentiated NEC state rather than committing to the RGC lineage.



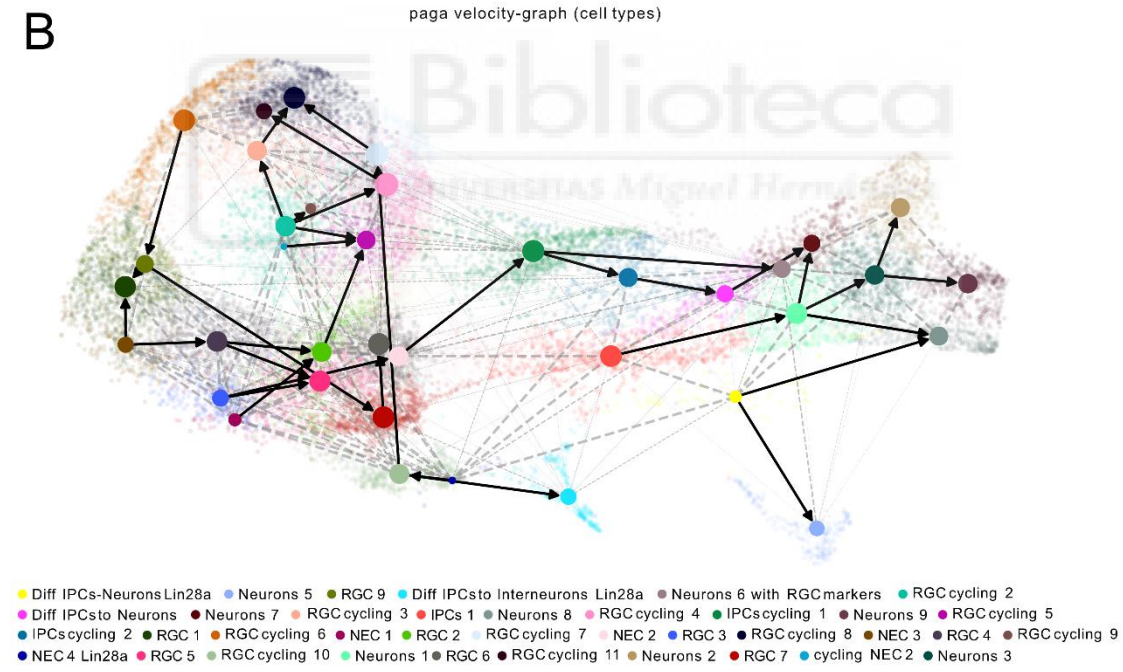


Figure 34. *PAGA* graph depicting the RNA trajectories. Partition-based Graph Abstraction (*PAGA*) constructs a graph-based representation of single-cell data to highlight potential developmental trajectories and the connectivity between different cell states. Nodes represent cell clusters, and edges indicate inferred connectivity based on RNA velocity. The direction of arrows reflects the predicted transitions between cell states. **A**) Represents the *PAGA* velocity graph for the WT condition, while **B**) depicts the corresponding graph for the KO condition.

4.1.4. ETMR marker genes in candidate tumor cells.

We conducted a preliminary study on the gene markers of *NEC_1*, *NEC_3*, and *NEC_4_Lin28a* cell types, as these are specific to the KO condition and exhibit a gene signature similar to NECs. Differential expression analysis in these cell types, compared to others, revealed several genes associated with the ETMR signature, including *Lin28a, Prtg*, and *Trim71*. These findings are illustrated in the gene expression plot (Figure 35), where *NEC_1, NEC_3*, and *NEC_4_Lin28a* show the highest expression levels of these genes.

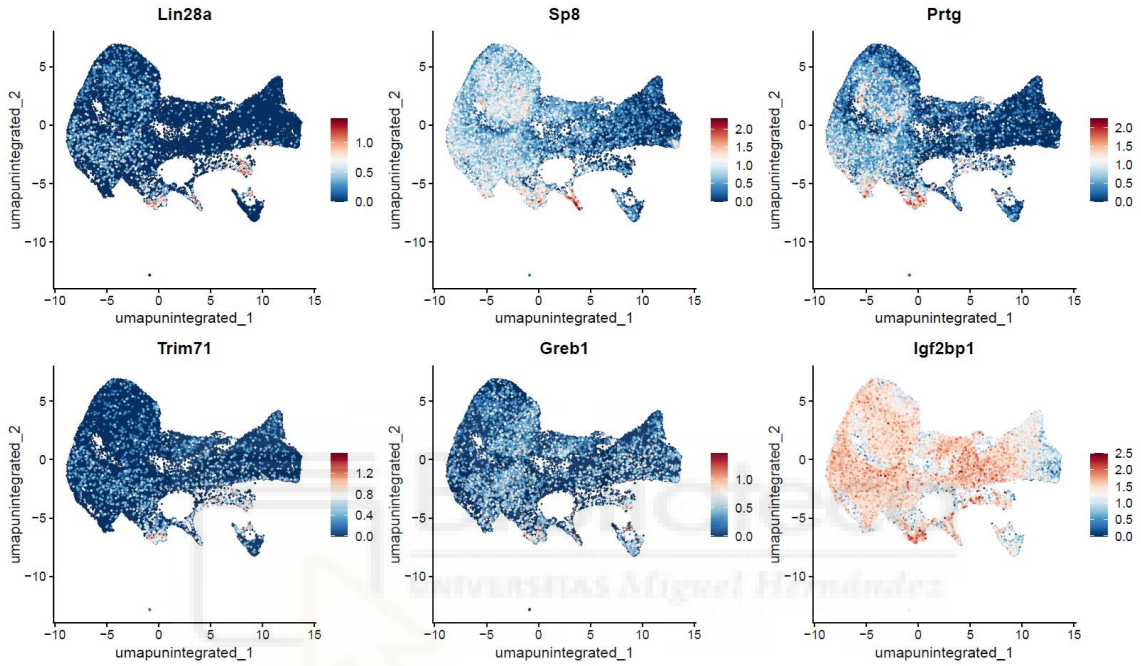


Figure 35. Expression of ETMR-Associated Genes in NEC Cell Types. This figure shows a feature plot with the expression of several key genes (*Lin28a, Sp8, Prtg, Trim71, Greb1,* and *Igf2bp1*) across different cell types in the UMAP plot. The genes are differentially expressed in KO-specific cell types (*NEC_1, NEC_3*, and *NEC_4_Lin28a*), which display a NEC-like gene signature. These genes, including *Lin28a* and *Prtg*, are part of the ETMR signature.

To confirm the overexpression of these genes in KO versus WT, violin plots were generated for two key ETMR markers, *Lin28a* and *Prtg* (Figure 36A, Figure 36B). The plots show *NEC_4_Lin28a* exhibiting the highest mean expression and the most distinct difference from WT both for *Lin28a* and *Prtg*. This suggests that *NEC_4_Lin28a* may be a strong candidate for contributing to the overproliferative cells forming rosettes in *Dicer1* mice, and it may also resemble the tumor-forming cells seen in human ETMR.

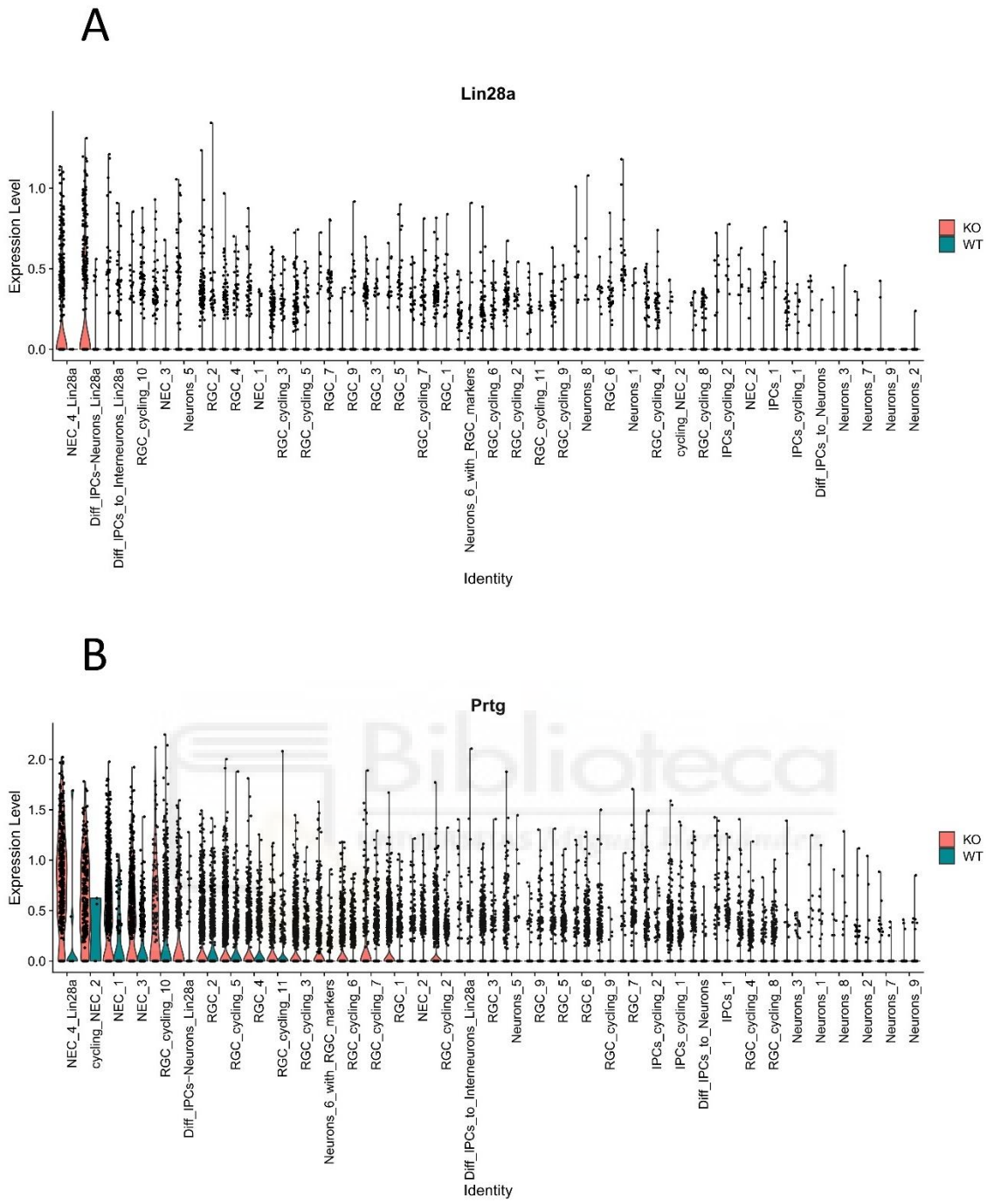


Figure 36. Violin Plots for *Lin28a* and *Prtg* Expression in KO vs. WT. A) Violin plot showing the expression levels of *Lin28a* across different cell types in KO and WT samples. The x-axis represents various cell types, ordered from left to right based on their mean expression levels, with the highest expression observed in *NEC_4_Lin28a*. The y-axis shows normalized expression levels. **B**) Violin plot depicting the expression levels of *Prtg* across different cell types in KO and WT samples. Similar to (A), the x-axis represents different cell types, with *NEC_4_Lin28a* showing the highest expression. The y-axis shows normalized expression values. In KO samples, *Lin28a* and *Prtg* are significantly overexpressed in *NEC_4_Lin28a*, further supporting its involvement in the overproliferation of cells, contributing to rosette formation in *Dicer1* mutants.

DISCUSSION

The main objective of this Doctoral Thesis was to uncover the genetic mechanisms underlying the evolution of cortical neurogenesis and its pathological dysregulation in the aggressive Embryonal Tumors with Multilayered Rosettes (ETMRs). To address this, a multidisciplinary approach was employed, integrating bulk RNA sequencing (RNA-seq), Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq), and single-cell RNA sequencing (scRNA-seq) across diverse model animal systems, including mice, chicks, snakes, and the *Rx-Dicer1* mutant mouse line. This comprehensive analysis provided novel insights into the following:

- **Mechanisms regulating neurogenic modes across species:** The balance between direct neurogenesis (DN) and indirect neurogenesis (IN) in apical radial glial cells (aRGCs) is modulated by species-specific transcriptomic and epigenomic features. Key transcriptomic regulators such as *Sall1*, *miR-3607*, *Robo/Dll1* signaling and *Cux2* were identified, providing insights into conserved and divergent pathways that drive cortical expansion and folding during evolution.
- **Discovery and characterization of membrane transcription factors:** Hundreds of potential MTFs were identified in human and mouse proteomes using *DeepTFactor*, revealing an abundance of candidates that localize to the nucleus and interact with DNA. Notably, ROBO1, a key regulator of neurogenic modes, emerged as a strong MTF candidate, underscoring the potential role of MTFs in modulating neurogenesis.
- **Validation of the *Rx-Dicer1* mutant mouse as ETMR model:** Transcriptomic comparisons with human ETMRs, the mutant mouse GBS (only available ETMR mouse model) and public bulk RNA-seq datasets of brain tumors, demonstrated that *Rx-Dicer1* mutants recapitulate most accurately the molecular and histological features of ETMRs.
- **Developmental trajectories of cortical progenitors in *Rx-Dicer1* mutants:** scRNA-seq analyses reconstructed the developmental pathways of cortical progenitors, identifying tumor-specific neuroepithelial-like populations in the *Rx-Dicer1* mutant mouse. These populations exhibit features that mirror those of human ETMRs.

1. Mechanisms regulating neurogenic modes across species

Previous research in our lab by Cárdenas et al. (2018) provided critical insights into the evolutionary mechanisms driving cortical neurogenesis. The study established that DN, characterized by limited neuron production, predominates in the paleocortex of birds, reptiles, and mammals. In contrast, in the evolutionarily newer mammalian NCx predominates IN, via Intermediate Progenitor Cells (IPCs) and basal Radial Glia Cells (bRGCs). This study also highlighted that modulation of conserved signaling pathways determines the mode of

neurogenesis: high *Slit*/*Robo* and low *Dll1* signaling via *Jag1* and *Jag2* drives DN, while attenuating *Robo* signaling and enhancing *Dll1* promotes IN. This shift in neurogenic modes was shown to contribute significantly to expansion and increased complexity of the cerebral cortex in mammals, particularly in the NCx (Cárdenas *et al.*, 2018).

Building on this framework, in this Thesis we have investigated the transcriptomic and epigenomic mechanisms underpinning neurogenic modes across amniotes. By analyzing key species, including mice, chicks, and snakes, this work aims to uncover the conserved and divergent pathways that regulate the balance between DN and IN modes, ultimately shaping cortical complexity during evolution. To achieve this, different brain regions and developmental stages, with varying proportions of DN and IN, were analyzed in each species, enabling the identification of genes specifically linked to these two neurogenic modes.

The project first focuses on mouse models to investigate gene expression changes underlying the transition between neurogenic modes. Transcriptomic comparisons between IN-enriched conditions (NCx E12) and DN-enriched conditions (OB E12, NCx E14) revealed thousands of DEGs. IN-enriched conditions were characterized by upregulation of genes like *Ccnd1/2* (Tsunekawa *et al.*, 2012; Bonnefont *et al.*, 2019), *Dmrta1* and *Neurog2* (Noack *et al.*, 2022; Shimojo, Masaki and Kageyama, 2024), and *Hmga2* (Kuwayama *et al.*, 2023), which promote progenitor proliferation and IPC formation. In contrast, DN-enriched conditions showed increased expression of genes associated with neuronal differentiation and synaptic development, such as *Gria3* and *Grin2b* (Karaca *et al.*, 2015; Platzer *et al.*, 2017) and Dscaml1 (Mayo *et al.*, 2023). GO analyses further highlighted this functional divergence, with IN enriched in cell cycle processes and DN-enriched enriched in neuronal development pathways, reflecting the distinct roles of aRGCs in proliferation versus differentiation.

In chick embryos, the distinction between proliferation and differentiation in IN-enriched and DN-enriched conditions is less clear compared to mouse models. While *DUSP16, RSPO2* and *RSPO3* promote neural differentiation and are more expressed in DN-enriched condition (Zega *et al.*, 2017; Gyllborg *et al.*, 2018; Zhang *et al.*, 2020), *EPHA3*, *CXCR7* and *LMO3* contribute to proliferation and are more expressed in IN-enriched condition (Kelava *et al.*, 2012; Day *et al.*, 2013; Chen *et al.*, 2015; Isogai *et al.*, 2015). Additionally, *WNT7B*, *WNT9A* and *NFIA/NFIB* are more expressed in DN-enriched condition, and critical for symmetric proliferation (Papachristou *et al.*, 2014; Bunt *et al.*, 2017; Zhu *et al.*, 2024).

In addition, high-confidence orthologous genes more expressed in IN-enriched conditions in both species were predominantly linked to IPC production (as expected), such as *NR2E1/NR2F1* (Faedo *et al.*, 2008; Drill, 2009; Sobhan and Funa, 2017), *GADD45G* (Kawaue *et al.*, 2014), and *CTNNA1* (Lien *et al.*, 2006; Stocker and Chenn, 2009; Bedogni and Hevner, 2021). On the other hand, *SFRP1*, which is more expressed in IN-enriched conditions, reduces neuron production by limiting the number of BPs (Miao *et al.*, 2018). Genes more expressed in DN-enriched conditions, like *JAG2* and *ZNRF2*, promote DN mode (Pinto *et*

al., 2008; Cárdenas *et al.*, 2018). Further research is needed to determine whether these genes conserve the same roles across species, which would validate and clarify the mixed pattern observed between chick and mouse neurogenic mechanisms. This need arises because almost all of the existing functional experiments have been performed in mammals, rather than in chicks or snakes, underscoring the importance of additional experimental validation in non-mammalian models.

The clear involvement of *JAG2* in promoting direct neurogenesis presents an intriguing avenue for further investigation (Cárdenas *et al.*, 2018). Future research should explore how *JAG2* contributes to the evolutionary adaptations of neurogenic modes, shedding light on its broader role in cortical development across amniotes. Deciphering the precise mechanisms through which *JAG2* promotes DN mode would be particularly valuable for understanding its functional impact on aRGC behavior and its evolutionary implications in shaping cortical development.

Epigenomic analyses were also performed, which revealed insights into chromatin accessibility and its role in regulating neurogenic modes. In the mouse, ATAC-seq identified regions of differential accessibility (DARs) linked to both IN-enriched and DN-enriched samples. In IN-enriched conditions, DARs were associated with genes like *Dmrta1*, *Jun*, and *Fzd8*, which are also DEGs. These genes promote aRGC proliferation and IPC production, underscoring their role in expanding the progenitor pool (Shaulian and Karin, 2001; Siegenthaler, Tremper-Wells and Miller, 2008; Boyd *et al.*, 2015). On the other hand, *MIR100HG*, which had DARs enriched in DN-enriched condition and it is differentially expressed, encodes miRNAs that deplete IPCs, highlighting its role in suppressing the IN mode (Lu *et al.*, 2017; Han *et al.*, 2023).

In chick embryos, ATAC-seq also uncovered DARs associated with distinct neurogenic modes. In IN-enriched regions, genes such as *IGDCC3* or *FGFR3* were linked to aRGC proliferation and cortical progenitor expansion. Interestingly, *IGDCC3* is also a proliferation marker in ETMR-associated aRGCs, reinforcing its potential significance (Inglis-Broadgate *et al.*, 2005; Goggolidou *et al.*, 2013; Zega *et al.*, 2017). However, DUSP16, which is associated with neuronal differentiation, also exhibited DARs that were more accessible in IN-enriched regions. In DN-enriched regions, genes such as *NFIB* and *HMGA2*, which typically promote IPC proliferation in mice, were more accessible and expressed, revealing a lack of correlation with the expected IN phenotype in chick. These findings, coupled with the presence of DARs in genes related to neuronal maturation, underscore the complexity of regulatory landscapes in chick neurogenesis (Pruunsild and Timmusk, 2005; Kishi *et al.*, 2012; Kuwayama *et al.*, 2023).

In principle, in chick embryos we would expect that genes involved in proliferation would be enriched in samples with higher proportions of IN, while genes involved in differentiation would be enriched in samples dominated by DN. However, we observed a mixed pattern, where proliferation-associated genes appear in DN-enriched regions, and differentiation-associated genes appear in IN-enriched regions. Three main factors may explain this blurred boundary: (1) the lateral and

medial dorsal pallium in E4 chick may not be as distinctly separated as the IN- and DN-enriched regions in mouse, obscuring clearer partitioning of these gene expression patterns; or (2) chick neurogenesis could be governed by different genetic mechanisms that balance DN and IN in a more overlapping way than what is observed in mouse. (3) Moreover, these inter-species discrepancies may reflect intrinsic differences in progenitor heterogeneity and potency, suggesting that neurogenic lineages in chick, mouse, and other amniotes constitute a spectrum of mixed division modes rather than a simple binary system, and that each species may deploy unique combinations of progenitor subtypes to achieve its characteristic cortical architecture.

Additionally, genes traditionally viewed as “proliferation-promoting” can in some contexts be time dependent. *Foxp1* exemplifies how “proliferation-promoting” genes can be highly time dependent in their function. When *Foxp1* is deleted early, aRGCs lose their self-renewing capacity and shift toward producing more IPCs, effectively reducing the population of aRGCs. Conversely, elevating *Foxp1* levels prolongs the period during which aRGCs undergo symmetric divisions, thereby reducing the generation of IPCs. Notably, maintaining high Foxp1 expression beyond the usual developmental window can lead to the emergence of bRGCs (Pearson *et al.*, 2020). Thus, *Foxp1* illustrates how the timing and duration of gene expression critically modulate whether a cell continues to expand via self-renewal (favoring certain forms of “proliferation”) or transitions toward BP production. This time dependence underscores the complexity of defining genes strictly as “proliferation-promoting,” as their roles may vary depending on developmental context and species-specific regulatory network. This temporal component, together with species-specific progenitor heterogeneity, further underscores that the regulation of neurogenic modes is more nuanced than a strict IN-versus-DN dichotomy.

Several pivotal regulatory elements emerge as central to these complex processes, offering insights into both conserved mechanisms and species-specific adaptations that shape cortical development. Below, we discuss how these key regulatory genes (*Mir3607*, *Sall1*, *Cux2*, and *Robo/Dll1*) contribute to the modulation of neurogenic modes, providing a framework for understanding the interplay between genetic and epigenomic factors across species.

We discovered that *Mir3607* plays a pivotal role in regulating aRGC expansion and cortical size. It is expressed embryonically in gyrencephalic mammals like humans and ferrets but absent in lissencephalic rodents. We found that experimental expression of *Mir3607* in embryonic mouse cortex amplified aRGC and IPC proliferation via *Wnt*/β-catenin signaling by inhibiting APC (Chinnappa *et al.*, 2022), a negative β-catenin regulator (Farkas *et al.*, 2008; Moore *et al.*, 2020). Its loss in ferrets reduced RGC proliferation, while its overexpression in human cerebral organoids promoted VZ expansion. We also discovered that differences in TF binding sites and structural variations in pre-*Mir3607* likely explain its divergent expression between species (Chinnappa *et al.*, 2022).

We found that *Sall1* is more expressed in IN-enriched conditions in all species analyzed, and it is predicted to bind to several genomic regions that are differentially-accessible between modes of neurogenesis. *Sall1* can function as both an activator and a repressor, binding to DARs associated with key IN-enriched genes like *Fzd8*, *Dmrta1*, and *Mycn*, all implicated in decisions between symmetric and asymmetric divisions (Kikkawa *et al.*, 2013; Boyd *et al.*, 2015; Otte *et al.*, 2021). Additionally, it has been shown that *Sall1* influences *WNT* signaling through activation of phosphorylated *JUN* (p-JUN) (Yuan *et al.*, 2023). Interestingly, *Sall1* is also predicted to bind to DARs in DN-enriched genes that are linked to neuronal genes such as *Gap43* and *Grin2b*. While their exact role in neurogenesis remains unclear, our hypothesis is that some neuronal genes expressed in RGCs under DN-enriched conditions may actively promote direct neurogenesis. While *Sall1* likely regulates modes of neurogenesis across species, determining its dual role as an activator or repressor requires further investigation using ChIP-seq and RNA-seq experiments.

We identified that *Cux2*, a typically neuronal transcription-factor expressed in a salt-and-pepper pattern among RGCs, is crucial for DN. In ferrets, *Cux2* was highly expressed in RGCs at the LS, aligning with DN-dominated regions, and its overexpression reduced immature RGC populations while increasing DN. These findings reinforce the hypothesis that certain neuronal genes expressed in RGCs actively promote DN, suggesting a conserved mechanism by which these genes drive the transition of progenitors to a committed neuronal fate (Singh *et al.*, 2024).

We further discovered that the synergistic interaction between *Robo1/2* and *Dll1* strongly promotes DN. Overexpression of *Robo1/2* combined with CRISPR-mediated *Dll1* knockdown reprogrammed RGCs, including enhancement of mitochondrial oxidative phosphorylation (OxPhos), and cell cycle arrest via upregulation of *Cdkn1a* (*p21*) (Karimian, Ahmadi and Yousefi, 2016). Interestingly, the transcriptomic shift included upregulation of genes linked to OxPhos and downregulation of progenitor maintenance genes, favoring DN mode (Iwata and Vanderhaeghen, 2021). However, we observed that the downregulation of neuronal genes in this context contrasts with previous findings where neuronal gene expression in RGCs promotes DN. This paradox warrants further exploration to clarify how neuronal genes modulate progenitor behavior in response to *Robo/Dll1* signaling.

In light of these discoveries, *Sall1* emerges as a pivotal TF, or somewhat master regulator, weaving together the threads of neurogenesis. Its ability to influence *Robo1* expression, modulate *WNT* signaling, and drive IPC production in mice positions *Sall1* at the heart of this intricate biological tapestry. Its dual role as both an activator and a repressor hints at a delicate balance, a molecular choreography that defines the fates of progenitor cells. To fully understand how *Sall1* orchestrates these processes across species is a necessary next step.

2. Discovery and characterization of membrane transcription factors

We have observed that several transmembrane receptor proteins, such as ROBO1 and DLL1, play essential roles in regulating cortical neurogenesis across evolutionary lineages. However, the mechanisms by which these membrane proteins influence gene expression remain largely unknown. Recent studies have introduced the intriguing concept of membrane-bound transcription factors (MTFs), a class of proteins capable of linking membrane signaling to transcriptional regulation (Seo, Kim and Park, 2008; Liu *et al.*, 2018; Yang *et al.*, 2020). Using the deep learning tool *DeepTFactor* (Kim *et al.*, 2021), we systematically predicted the likelihood of transmembrane proteins functioning as MTFs.

To expand our understanding, we applied the analysis pipeline to the entire human and mouse transmembrane proteomes, as detailed below. Using *DeepTFactor*, we established a step-by-step analysis pipeline that led to identifying, in a completely unsupervised manner, a set of 144 MTF candidates in humans and 144 in mice, with 77 overlapping between them. This indicates that only 0.31% of transmembrane proteins qualify as predicted MTFs (with a TF score above 0.5). This list includes the small set of 13 mammalian MTFs previously validated by UniProt [ATF6, ATF6B, CREB3, CREB3L1, CREB3L2, CREB3L3, CREB3L4, JPH2, MYRF, PLSCR1, SREBF1, SREBF2, and XBP1], which confers high confidence to the analysis pipeline and the validity of these results. Most importantly, the new list significantly expands the collection of putative MTFs in the mammalian proteome.

TFs are typically characterized as small soluble proteins with a key DNA-binding domain (Mitsis *et al.*, 2020). In agreement with this, we found that the presence of transmembrane domains as part of the protein frequently (~71% half TMD, ~95% full TMD) negatively impacts the MTF prediction score. Using experimentally validated MTFs, we conducted in-silico tests to discover which domains are recognized by *DeepTFactor* as important for identifying a protein as an MTF. Remarkably, the DNA-binding domain was key in only 71.42% of cases, while other domains were recognized as critical in the remaining cases. Because TFs interact with DNA to regulate gene expression, their transport to, and localization at, the cell nucleus are essential conditions of TFs (Lambert *et al.*, 2018; Mitsis *et al.*, 2020). By using the Human Protein Atlas (HPA) repository (Thul *et al.*, 2017), we determined the subcellular localization of the novel MTFs and confirmed that ~84% of the novel MTFs have nuclear localization in human cell lines. In summary, these results provide a robust prediction of MTF candidate transmembrane proteins.

To gauge tissue-specific relevance, we cross-referenced the 144 human candidates with the Human Protein Atlas ‘brain-enriched’ categories (Thul *et al.*, 2017) and found 29 matches (≈ 20 % of the set) while an equivalent 26/144 (≈ 18 %) were classified as brain-enriched in the mouse Brain RNA-seq Atlas (Harris *et al.*, 2019). Thus, roughly one in five predicted MTFs appears preferentially expressed in the nervous system.

Closer inspection suggests three functional “convergence hubs”: (i) ER-stress sensors (ATF6/CREB3 family, XBP1) that activate the unfolded-protein response; (ii) lipid-homeostasis regulators (SREBF1/2, MYRF, CREB3L2) that up-shift

membrane-biosynthetic genes; and (iii) contact/adhesion ICD-type MTFs (ROBO1-ICD, NOTCH-ICD, PCDH-ICDs) whose shared targets are enriched for cytoskeletal and axon-guidance genes. These patterns suggest that MTFs, more than soluble TFs, act as “hardware sensors,” coupling the physical state of a membrane compartment (secretory load, lipid scarcity, cell-cell contact) directly to the transcriptional programmes that expand or remodel that compartment. This is an important requirement during cortical neurogenesis, when radial-glial end-feet, elongating neurites and nascent oligodendrocytes all experience extreme membrane demands.

From this perspective, we speculate that brain-biased MTFs could (a) gate the switch from symmetric to asymmetric divisions by linking ER/mitotic load to lineage decisions, (b) synchronise lipid supply with dendrite and spine growth, and (c) translate contact-mediated cues into nuclear signals that refine axon targeting. Systematic mapping of cleavage triggers and DNA-binding partners for each candidate will reveal whether these subclasses converge on common regulons or partition labour across complementary facets of neurodevelopment.

Finally, we identified ROBO1 as a strong MTF candidate, raising the possibility that it not only regulates neurogenesis through canonical signaling but also directly influences gene expression as an MTF. Human ROBO1 is indeed proteolytically cleaved by metalloproteinases and γ-secretase, producing a C-terminal fragment capable of translocating to the nucleus in cancer cells (Seki *et al.*, 2010; Bianchi *et al.*, 2019). Moreover, the tumor suppressor APC has been shown to regulate ROBO1 internalization and trafficking by dissociating from its cytoplasmic domain upon SLIT2 binding, which facilitates clathrin-mediated endocytosis, a step required for subsequent nuclear localization (Huang *et al.*, 2024). Notably, APC is a well-known negative regulator of *Wnt/β-catenin* signaling in the developing cortex, restricting RGC proliferation and cortical expansion (Chinnappa *et al.*, 2022)

We know from previous results from the lab that simultaneous loss of *Slit1* and *Slit2* produces a phenotype similar to *Robo1/2* mutants, characterized by a marked increase in *Tbr2*-positive progenitors, a feature indicative of the promotion of IN mode. These observations suggest the presence of an interconnected gene regulatory network whereby *Slit1/2* and *Robo1* (possibly via nuclear functions) partner with *APC* to control progenitor dynamics and direct neurogenesis. By contrast, DLL1 consistently scored below the 0.5 threshold in our *DeepTFactor* predictions, indicating that it is not predicted to function as an MTF.

While *DeepTFactor* is a powerful tool to predict MTFs, several important limitations must be considered. First, the prediction accuracy heavily relies on the availability and quality of the training dataset, which may exclude some important but yet undiscovered TFs or MTFs. Second, the relatively low ability of the tool to recognize DNA-binding domains, with an accuracy of ca. 71%, may limit its ability to predict certain TFs. However, when predicting canonical TFs, the deep learning algorithm also considers other domains beyond the DNA-binding domain. Despite this, the tool has an impressive accuracy of ~0.98/1 (Kim *et al.*, 2021). Moreover, it is important to consider that some MTFs can regulate gene expression in

non-canonical ways, or in protein complexes where they do not bind DNA directly, and this may limit their recognition as TFs. Another important consideration is that nearly 50% of transmembrane proteins that have undergone manual curation in UniProt lack topological domain annotations, indicating that the list of potential MTFs identified in this study is likely to be substantially greater (The UniProt Consortium, 2021). Finally, significant variability was observed in the subcellular localization of predicted MTFs, where not all were detected in the cell nucleus. Subcellular localization may depend heavily on the cellular context, and MTFs, in particular, may remain in a non-active, latent state, not shuttling into the nucleus or acting as TFs until activated.

In conclusion, our *DeepTFactor*-based pipeline greatly narrows down the pool of putative MTFs from the transmembrane proteome, making targeted experimental validation more tractable. The identification of *Robo1* as a strong MTF candidate, contrasted with *Dll1* low score, highlights the power of this approach. As we refine our understanding of membrane-integrated signaling and transcription, these newly uncovered candidate MTFs, including Robo family members, represent promising avenues for investigating how membrane-bound signals directly modulate gene expression in complex developmental processes like cortical neurogenesis.

3. Validation of the *Rx-Dicer1* mutant mouse as a model of ETMR

Embryonal Tumor with Multilayered Rosettes (ETMR) is a highly aggressive pediatric brain tumor characterized by the excessive proliferation of tumor cells, resulting in the formation of distinctive rosettes (Jessa *et al.*, 2019). In prior research from our lab, the *Rx-Dicer1* mutant mouse was identified as a promising candidate for modeling ETMR due to its phenotypic and molecular similarities to the human disease (Fernández *et al.*, 2020). To really validate this model, we conducted a comprehensive evaluation using transcriptomic analyses to determine whether the *Rx-Dicer1* mutant faithfully replicates key features of ETMR.

Our findings demonstrate that the *Rx-Dicer1* mutant mouse effectively recapitulates hallmark characteristics of human ETMRs. Transcriptomic analysis at E11.5, before the onset of rosette formation, identified 4,642 DEGs in the NCx and septum. Among these, genes such as *Lin28a/b, Prtg, Nr6a1,* and *Greb1*, which are known to drive proliferation and maintain stemness, were significantly upregulated—mirroring key traits of ETMRs. Furthermore, the strong overlap between these DEGs and ETMR-associated genes, including *Lin28a* and *Dnmt3b*, underscores the critical role of these drivers in establishing the molecular phenotype observed in *Rx-Dicer1* mutants (Cheng *et al.*, 2016; Sin-Chan *et al.*, 2019; Xiang *et al.*, 2021; Cotino-Nájera *et al.*, 2024).

Reduced levels of mature *let-7* miRNA in *Rx-Dicer1* mutants lead to the upregulation of *let-7* target genes, including *Lin28a/b, Prtg, Nr6a1*, and *Greb1*, further enhancing proliferative and undifferentiated states. This disruption is compounded by a feedback loop, where elevated *Lin28a/b* levels suppress *Dicer1* expression, hindering miRNA maturation. These molecular interactions underline

a positive feed-back mechanism that perpetuates the ETMR-like state in the mutant model (Leone *et al.*, 2008).

At E17.5, the transcriptomic landscape of mature rosettes demonstrates persistent dysregulation of ETMR-associated genes, though less pronounced than at earlier stages. *Lin28a* and other *let-7* targets remain altered, but genes such as *Prss56, Igf1, Igf2bp1, and Pik3c2b* are significantly upregulated, driving the activation of the *PI3K/AKT/mTOR* pathway. This pathway is critical for promoting proliferation and malignancy, further defining the aggressive phenotype of the *Rx-Dicer1* mutants (Zhang *et al.*, 2018; Margaria *et al.*, 2019; Mancarella, Morrione and Scotlandi, 2021; Li *et al.*, 2023).

Upregulation of *Ccne1* and *Cdc25a* in mature rosettes underscores enhanced cell cycle progression and replication stress in *Rx-Dicer1* mutants. These changes indicate an increased proliferative capacity, consistent with the malignant characteristics of ETMRs (Li *et al.*, 2012). Additionally, downregulated genes such as *Nkx2.2, Grid1*, and *Gria1* suggest impaired neuronal maturation and disrupted synaptic function, highlighting the loss of neurogenic potential in *Rx-Dicer1* tumors (Ismail *et al.*, 2022; Kirjavainen *et al.*, 2022; Ung *et al.*, 2024).

Gene ontology analysis reveals that early stages are dominated by reduced neural differentiation, increased proliferation pathways, and activation of immune system. Similarly, mature rosettes show enrichment of proliferation, *IL-6* signaling pathway, and epithelial-to-mesenchymal transition (EMT) pathways, with diminished neuronal differentiation.

This transition appears to occur in two distinct phases driven by the absence of mature *let-7* miRNA. During the initial stage, the lack of mature *let-7* leads to the overexpression of its target genes, including *Lin28a/b, Prtg, Nr6a1*, and *Greb1*, many of which are canonical markers of ETMR. These genes establish a proliferative and undifferentiated state characteristic of early ETMR-like pathology in the *Rx-Dicer1* mutants. By E17.5, the transcriptomic landscape evolves to reflect a combination of the ongoing lack of mature *let-7* and the cumulative effects of the sustained overexpression of its targets. This creates a new transcriptomic profile marked by persistent dysregulation of ETMR-associated genes alongside the activation of secondary pathways, such as *PI3K/AKT/mTOR*, *WNT* or *IL-6* (Wegiel *et al.*, 2008; Chen *et al.*, 2016; Kim, Kim and Park, 2019). The result is a more complex molecular phenotype that integrates the initial effects of *let-7* target overexpression with the downstream consequences of prolonged signaling disruption.

Interestingly, *IL-6* enhances *AKT* activity, forming a positive feedback loop with *MSI1-AKT*. This interaction amplifies proliferative signaling through the *PI3K/AKT/mTOR* pathway. Similar *IL-6*/*AKT* crosstalk has been documented in various tumors, where it plays a role in sustaining cancer cell proliferation and survival. Importantly, inhibiting *IL-6* has been shown to resensitize cancer cells that are resistant to *PI3K/AKT/mTOR* inhibitors, highlighting the therapeutic potential of targeting this pathway (Wegiel *et al.*, 2008; Chen *et al.*, 2016).

After transcriptomically characterizing the *Rx-Dicer1* mutant mouse and establishing its ETMR-like features, we extended our analysis to assess its specificity for ETMR by integrating publicly available RNA-seq data. Using datasets from Jessa et al. (2019) and Neumann et al. (2017), we compared the full transcriptome of *Rx-Dicer1* mutants to human brain tumor bulk RNA-seq data, including ETMR and other brain tumors, as well as to the previously established GBS ETMR mouse model. Hierarchical clustering and heatmap visualization revealed that *Rx-Dicer1* mutants at both E11.5 and E17.5 closely clustered with human ETMRs, showing strong transcriptomic similarity. In contrast, the GBS model and HGNET samples displayed moderate similarity, underscoring the *Rx-Dicer1* mutant's specificity and robustness as an ETMR model.

As we validated that the *Rx-Dicer1* mutant mouse aligns most closely with human ETMRs when considering the full transcriptome, we also confirmed its specificity using a DEG-based comparison. GSEA was performed on the top 500 DEGs for each tumor model compared to controls, further supporting the *Rx-Dicer1* model's relevance. When integrating both upregulated and downregulated DEGs, *Rx-Dicer1* mutants at E11.5 and E17.5 exhibited a higher overall similarity to ETMR than the GBS model. At E17.5, *Rx-Dicer1* mutants showed the strongest alignment with ETMR while also overlapping with ATRT and MB. This overlap is particularly noteworthy given that *Dicer1* mutations have also been implicated in ATRT pathogenesis, suggesting a shared molecular framework between these tumor types. Although GBS was enriched in several tumor-related signatures, it lacked the robust ETMR specificity observed in *Rx-Dicer1* mutants due to non-significant downregulation patterns.

Building on our DEG-based comparison, we further evaluated tumor-specific signatures to refine the alignment of *Rx-Dicer1* mutants with ETMR. Tumor-specific signatures were generated by identifying DEGs unique to each tumor type from Jessa et al. (2019). Analysis of the ETMR-specific signature, comprising genes consistently overexpressed in ETMR compared to other brain tumors, revealed that most of these DEGs are also overexpressed in the *Rx-Dicer1* model. Key ETMR markers such as *DNMT3B, LIN28A, LIN28B, PRTG, GREB1*, and *IGF2BP1*, essential for tumor initiation and maintenance, were exclusively found in *Rx-Dicer1* compared to GBS (Cheng *et al.*, 2016; Sin-Chan *et al.*, 2019; Xiang *et al.*, 2021; Cotino-Nájera *et al.*, 2024). These findings underscore the closer alignment of *Rx-Dicer1* with human ETMRs compared to GBS, which exhibited minimal overlap with the ETMR-specific signature.

Additionally, the developmental stage comparison highlighted the stronger representation of ETMR signature genes in *Rx-Dicer1* mutants at E17.5, particularly in the SP region. This stage demonstrated the highest overlap in both ETMR-specific genes and DEGs, with GSEA enrichment confirming E17.5’s strongest ETMR-like profile. These results emphasize the *Rx-Dicer1* mutant’s fidelity as an ETMR model at later developmental stages, further distinguishing it from GBS.

In conclusion, the *Rx-Dicer1* mutant mouse emerges as a highly faithful model of human ETMR, accurately recapitulating key molecular and transcriptomic features of the disease. Its closer alignment with human ETMRs compared to GBS

solidifies its utility for advancing our understanding of ETMR pathogenesis and therapeutic development.

Although the Rx-Dicer1 mouse captures many molecular hallmarks of human ETMR, several caveats must be acknowledged. First, DICER1 loss alone does not generate the C19MC micro-RNA cluster amplification that defines ~90 % of human ETMRs, so key oncogenic drivers such as miR-517c/miR-520g are absent. Second, pan-retinal deletion of Dicer1 perturbs the entire miRNA repertoire; consequently, the resulting tumor phenotype may reflect broad miRNA collapse rather than the selective let-7 axis typical of ETMR. Third, early embryonic lethality in complete Dicer1 KOs necessitates tissue-restricted or hypomorphic alleles, which may not fully mirror the timing or mosaicism of somatic mutations in patients. Fourth, the murine forebrain lacks certain human-specific progenitor subtypes (e.g., outer radial glia) that could modulate ETMR biology, potentially limiting the model's predictive power for therapeutic responses. Taken together, the Rx-Dicer1 model excels at dissecting let-7/LIN28-driven proliferation but should be complemented with C19MC-focused or human organoid systems to capture the full genetic spectrum of ETMR.

4. Developmental trajectories of cortical progenitors in *Rx-Dicer1* mutants

Once the *Dicer1* mutant mouse was validated as a faithful model of ETMR, we performed scRNA-seq to uncover cell composition differences and developmental trajectories of cortical progenitors in the embryonic NCx of knockout (KO) versus wild-type (WT) embryos at E12.5. Our analyses revealed distinct cellular populations between WT and KO samples. Condition-dependent shifts in cell composition were evident, with clusters specific to either genotype. Detailed subclustering further divided the cell populations into 36 groups, highlighting subtle differences in cell types and states. Canonical developmental marker genes were used to annotate these clusters, identifying higher-level categories such as RGCs, neuroepithelial cells (NECs), IPCs, neurons, and cycling progenitors (Di Bella *et al.*, 2021; Ruan *et al.*, 2021; Dave *et al.*, 2023). These results reveal early transcriptional changes in *Dicer1* mutants, particularly in KO-specific clusters

Comparing cell-type proportions between KO and WT samples revealed striking differences in cell composition (Phipson *et al.*, 2022). Notably, KO samples showed a significant enrichment of NEC-like cells, with three of the four most KO-specific cell types exhibiting a NEC signature. This enrichment suggests that in the absence of functional *Dicer1*, RGCs revert to a more primitive NEC-like identity. On the other hand, WT samples were enriched in RGCs, indicating that RGCs in KO samples fail to maintain their radial glial phenotype and become "trapped" in an undifferentiated NEC state. This reversion to a primitive NEC-like state in KO samples may underpin the development of the highly proliferative rosettes characteristic of ETMR.

To further explore the dynamics of cell-state transitions, RNA velocity analysis was performed (La Manno *et al.*, 2018; Bergen *et al.*, 2021). In WT samples, RNA velocity vectors showed that RGCs follow a complete cell cycle trajectory, reflecting normal proliferation and differentiation processes. In contrast, KO samples

displayed incomplete cycling, with velocity vectors indicating that RGCs transition into NEC-like states without returning to the RGC state. This suggests a disruption in normal differentiation, as KO RGCs become stalled or revert to an undifferentiated NEC phenotype.

Partition-based Graph Abstraction (*PAGA*) trajectory analysis further illuminated these cell-state progressions (Wolf *et al.*, 2019). In WT samples, cycling RGCs linked back to non-cycling RGCs, demonstrating standard differentiation loops. However, in KO samples, trajectories shifted abnormally. RGCs transitioned toward cycling progenitors but failed to revert to RGC states. Additionally, NECs and cycling progenitors in KO samples exhibited no trajectory back to RGCs, indicating a “point of no return” that drives hyperproliferative rosette formation. This disrupted trajectory underscores the pathological mechanisms triggered by *Dicer1* loss, establishing a foundation for the development of ETMR-like features in the *Rx-Dicer1* mutant mouse model.

Finally, the analysis of KO-specific NEC clusters revealed significant insights into the molecular underpinnings of rosette formation in *Dicer1* mutants (Hao *et al.*, 2024). Clusters such as *NEC_1, NEC_3*, and *NEC_4_Lin28a* were unique to the KO condition and exhibited a gene signature highly similar to NECs. Differential expression analysis showed that these clusters displayed elevated levels of ETMR marker genes, including *Lin28a*, *Prtg*, and *Trim71*. Among these, *NEC_4_Lin28a* stood out with markedly higher expression of *Lin28a* and other critical tumor markers, as highlighted in Figure 35A and 35B. Violin plots further confirmed the consistent overexpression of *Lin28a* and *Prtg* in *NEC_4_Lin28a* compared to other cell types, reinforcing its prominence in the KO condition. These findings suggest that *NEC_4_Lin28a* could play a pivotal role in driving overproliferation and rosette formation, aligning closely with human ETMR pathology. This cluster may represent a murine equivalent of ETMR tumor cells, offering a valuable model for studying the mechanisms underlying this aggressive pediatric brain tumor.

While RNA-velocity and PAGA give a first glimpse of progenitor dynamics in the Rx-Dicer1 cortex, several caveats temper how literally their arrows and graphs should be read. (i) Both methods derive pseudotime from a single snapshot (E12.5); any true difference in developmental timing between WT and KO can skew vector direction. (ii) RNA-velocity assumes uniform transcription/splicing kinetics across cells, yet Dicer1 loss perturbs miRNA biogenesis and secondarily alters the stability of thousands of transcripts, potentially violating that assumption and biasing the “unspliced-to-spliced” ratio. (iii) Neither velocity nor PAGA incorporates tissue-spatial information, so branches may reflect transcriptional proximity rather than bona-fide lineage neighbourhoods in vivo. (iv) Resolution is limited by sequencing depth: rapidly cycling, low-complexity progenitors (e.g., NECs) generate noisy profiles that can cluster artefactually. (v) Crucially, computational trajectories do not prove causality; the inferred paths must be validated with experimental lineage tracing (fate-mapping, bar-coding) and functional assays.

CONCLUSIONS

1. Genes controlling progenitor proliferation versus neuronal differentiation vary across amniotes, with correspondence by differentially accessible chromatin regions.

2. Individual TFs can simultaneously drive or repress proliferation, differentiation, and cortical folding processes, as observed for *Sall1* and *Cux2*.

3. The *Robo/Dll1* and *Mir3607* pathways have profound influences on the molecular biology and fate of radial glia progenitor cells, and these regulatory mechanisms are conserved across amniotes.

4. *DeepTFactor* identifies known MTFs and uncovers new candidates, significantly increasing their repertoire. *Robo1* stands out as a new MTF, suggesting a main role for these molecules in regulating neurogenic modes.

5. *Rx-Dicer1* mutant mouse embryos display molecular phenotypes that mimic ETMR, including reduced expression of *let-7* miRNA, overexpression of *Lin28a/b*, and activation of signaling pathways typical of ETMR.

6. *Rx-Dicer1* mutants are transcriptomically more similar to ETMR than the GBS mouse, validating *Rx-Dicer1* as a superior model for ETMR research.

7. Persistent dysregulation of *PI3K/AKT/mTOR* and *IL-6* signaling by late embryogenesis (E17.5) suggests potential new targets for therapeutic intervention.

8. Single-cell analyses reveal *Rx-Dicer1* mutants trap radial glial cells in an undifferentiated NEC-like state, preventing normal cortical progenitor progression and driving abnormal rosette formation.

CONCLUSIONES

1. Los genes que controlan la proliferación de progenitores frente a la diferenciación neuronal varían entre los amniotas, con correspondencia de las regiones cromatínicas diferencialmente accesibles.

2. Factores de transcripción individuales pueden simultáneamente promover o reprimir la proliferación, la diferenciación y los procesos de plegamiento cortical, tal como se observa con *Sall1* y *Cux2*.

3. Las vías *Robo/Dll1* y *Mir3607* ejercen profundos efectos sobre la biología molecular y el destino de las células progenitoras de la glía radial, y estos mecanismos reguladores están conservados entre los amniotas.

4. *DeepTFactor* identifica MTF conocidos y descubre nuevos candidatos, ampliando significativamente su repertorio. *Robo1* destaca como un nuevo MTF, lo que sugiere un papel importante en la regulación de los modos neurogénicos.

5. Los ratones mutantes *Rx-Dicer1* imitan estrechamente ETMR al reducir el miARN *let-7*, sobreexpresar *Lin28a/b* y activar vías clave de proliferación y pluripotencia, dando lugar a fenotipos hiperproliferativos similares a tumores.

6. Los mutantes *Rx-Dicer1* muestran una mayor similitud transcriptómica con el ETMR que el ratón GBS, lo que valida *Rx-Dicer1* como un modelo superior para la investigación del ETMR.

7. La disrupción persistente de las vías *PI3K/AKT/mTOR* e *IL-6* en etapas tardías del desarrollo embrionario (E17.5) sugiere posibles nuevas dianas terapéuticas.

8. Los análisis de células individuales revelan que los mutantes *Rx-Dicer1* mantienen a las células gliales radiales en un estado similar al de células neuroepiteliales indiferenciadas, impidiendo la progresión normal de los progenitores corticales y favoreciendo la formación anómala de rosetas.

REFERENCES

Acemel, R.D. *et al.* (2016) 'A single three-dimensional chromatin compartment in amphioxus indicates a stepwise evolution of vertebrate Hox bimodal regulation', *Nature Genetics*, 48(3), pp. 336–341. Available at: https://doi.org/10.1038/ng.3497.

Aleksic, T. *et al.* (2018) 'Nuclear IGF1R Interacts with Regulatory Regions of Chromatin to Promote RNA Polymerase II Recruitment and Gene Expression Associated with Advanced Tumor Stage', *Cancer Research*, 78(13), pp. 3497–3509. Available at: https://doi.org/10.1158/0008-5472.CAN-17-3498.

Amemiya, H.M., Kundaje, A. and Boyle, A.P. (2019) 'The ENCODE Blacklist: Identification of Problematic Regions of the Genome', *Scientific Reports*, 9(1), p. 9354. Available at: https://doi.org/10.1038/s41598-019-45839-z.

Anders, S., Pyl, P.T. and Huber, W. (2015) 'HTSeq—a Python framework to work with high-throughput sequencing data', *Bioinformatics*, 31(2), pp. 166–169. Available at: https://doi.org/10.1093/bioinformatics/btu638.

Andrews, M.G. and Pearson, C.A. (2023) 'Toward an understanding of glucose metabolism in radial glial biology and brain development', *Life Science Alliance*, 7(1), p. e202302193. Available at: https://doi.org/10.26508/lsa.202302193.

Andrews, S. (2010) *FastQC: a quality control tool for high throughput sequence data*. Available at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (Accessed: 9 December 2024).

Angevine Jr., J.B. *et al.* (1970) 'Embryonic vertebrate central nervous system: Revised terminology', *The Anatomical Record*, 166(2), pp. 257–261. Available at: https://doi.org/10.1002/ar.1091660214.

Antonacci, F. *et al.* (2014) 'Palindromic GOLGA8 core duplicons promote chromosome 15q13.3 microdeletion and evolutionary instability', *Nature genetics*, 46(12), pp. 1293–1302. Available at: https://doi.org/10.1038/ng.3120.

Arrázola, M.S. *et al.* (2019) 'Mitochondria in Developmental and Adult Neurogenesis', *Neurotoxicity Research*, 36(2), pp. 257–267. Available at: https://doi.org/10.1007/s12640-018-9942-y.

Attardo, A. *et al.* (2008) 'Live Imaging at the Onset of Cortical Neurogenesis Reveals Differential Appearance of the Neuronal Phenotype in Apical versus Basal Progenitor Progeny', *PLOS ONE*, 3(6), p. e2388. Available at: https://doi.org/10.1371/journal.pone.0002388.

Au-Yeung, G. *et al.* (2017) 'Selective targeting of Cyclin E1 amplified high grade serous ovarian cancer by cyclin-dependent kinase 2 and AKT inhibition', *Clinical cancer research : an official journal of the American Association for Cancer Research*, 23(7), pp. 1862–1874. Available at: https://doi.org/10.1158/1078-0432.CCR-16-0620.

Bailey, J.A. *et al.* (2002) 'Recent Segmental Duplications in the Human Genome', *Science*, 297(5583), pp. 1003–1007. Available at: https://doi.org/10.1126/science.1072047.

Bailey, T.L. *et al.* (2015) 'The MEME Suite', *Nucleic Acids Research*, 43(W1), pp. W39–W49. Available at: https://doi.org/10.1093/nar/gkv416.

Bailey, T.L. and Elkan, C. (1994) 'Fitting a mixture model by expectation maximization to discover motifs in biopolymers', *Proceedings. International Conference on Intelligent Systems for Molecular Biology*, 2, pp. 28–36.

Bandopadhayay, P. and Chi, S.N. (2022) 'The challenges in treating embryonal tumors with multilayered rosettes (ETMR) and other infant brain tumors', *Neuro-oncology*, 24(1), pp. 138–140. Available at: https://doi.org/10.1093/neuonc/noab206.

Barembaum, M. and Bronner-Fraser, M. (2004) 'A novel spalt gene expressed in branchial arches affects the ability of cranial neural crest cells to populate sensory ganglia', *Neuron Glia Biology*, 1(1), pp. 57–63. Available at: https://doi.org/10.1017/s1740925x04000080.

Barski, A. *et al.* (2007) 'High-resolution profiling of histone methylations in the human genome', *Cell*, 129(4), pp. 823–837. Available at: https://doi.org/10.1016/j.cell.2007.05.009.

Basson, M. and Horvitz, H.R. (1996) 'The Caenorhabditis elegans gene sem-4 controls neuronal and mesodermal cell development and encodes a zinc finger protein', *Genes & Development*, 10(15), pp. 1953–1965. Available at: https://doi.org/10.1101/gad.10.15.1953.

Bedogni, F. and Hevner, R.F. (2021) 'Cell-Type-Specific Gene Expression in Developing Mouse Neocortex: Intermediate Progenitors Implicated in Axon Development', *Frontiers in Molecular Neuroscience*, 14, p. 686034. Available at: https://doi.org/10.3389/fnmol.2021.686034.

Bergen, V. *et al.* (2020) 'Generalizing RNA velocity to transient cell states through dynamical modeling', *Nature Biotechnology*, 38(12), pp. 1408–1414. Available at: https://doi.org/10.1038/s41587-020-0591-3.

Bergen, V. *et al.* (2021) 'RNA velocity—current challenges and future perspectives', *Molecular Systems Biology*, 17(8), p. e10282. Available at: https://doi.org/10.15252/msb.202110282.

Betancourt, J., Katzman, S. and Chen, B. (2014) 'Nuclear factor one B regulates neural stem cell differentiation and axonal projection of corticofugal neurons', *Journal of Comparative Neurology*, 522(1), pp. 6–35. Available at: https://doi.org/10.1002/cne.23373.

Betizeau, M. *et al.* (2013) 'Precursor diversity and complexity of lineage relationships in the outer subventricular zone of the primate', *Neuron*, 80(2), pp. 442–457. Available at: https://doi.org/10.1016/j.neuron.2013.09.032.

Bianchi, G. *et al.* (2019) 'The Transmembrane Receptor Roundabout 1 (ROBO1) Is Necessary for Multiple Myeloma Proliferation and Homing to the Bone Marrow Niche', *Blood*, 134, p. 507. Available at: https://doi.org/10.1182/blood-2019-128778.

Biegel, J.A. *et al.* (2002) 'Alterations of the hSNF5/INI1 gene in central nervous system atypical teratoid/rhabdoid tumors and renal and extrarenal rhabdoid tumors', *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, 8(11), pp. 3461–3467.

Biegel, J.A., Busse, T.M. and Weissman, B.E. (2014) 'SWI/SNF Chromatin Remodeling Complexes and Cancer', *American journal of medical genetics. Part C, Seminars in medical genetics*, 0(3), pp. 350–366. Available at: https://doi.org/10.1002/ajmg.c.31410.

Bishop, K.M., Goudreau, G. and O'Leary, D.D. (2000) 'Regulation of area identity in the mammalian neocortex by Emx2 and Pax6', *Science (New York, N.Y.)*, 288(5464), pp. 344–349. Available at: https://doi.org/10.1126/science.288.5464.344.

Bonev, B., Stanley, P. and Papalopulu, N. (2012) 'MicroRNA-9 Modulates Hes1 ultradian oscillations by forming a double-negative feedback loop', *Cell Reports*, 2(1), pp. 10–18. Available at: https://doi.org/10.1016/j.celrep.2012.05.017.

Bonnefont, J. *et al.* (2019) 'Cortical Neurogenesis Requires Bcl6-Mediated Transcriptional Repression of Multiple Self-Renewal-Promoting Extrinsic Pathways', *Neuron*, 103(6), pp. 1096-1108.e4. Available at: https://doi.org/10.1016/j.neuron.2019.06.027.

Borrell, V. *et al.* (2012) 'Slit/Robo signaling modulates the proliferation of central nervous system progenitors', *Neuron*, 76(2), pp. 338–352. Available at: https://doi.org/10.1016/j.neuron.2012.08.003.

Borrell, V. (2019) 'Recent advances in understanding neocortical development', *F1000Research*, 8, p. 1791. Available at: https://doi.org/10.12688/f1000research.20332.1.

Borrell, V. and Götz, M. (2014) 'Role of radial glial cells in cerebral cortex folding', *Current Opinion in Neurobiology*, 27, pp. 39–46. Available at: https://doi.org/10.1016/j.conb.2014.02.007.

Boyd, J.L. *et al.* (2015) 'Human-chimpanzee differences in a FZD8 enhancer alter cell-cycle dynamics in the developing neocortex', *Current biology: CB*, 25(6), pp. 772–779. Available at: https://doi.org/10.1016/j.cub.2015.01.041.

Brand, C.M. *et al.* (2024) 'Sequence-Based Machine Learning Reveals 3D Genome Differences between Bonobos and Chimpanzees', *Genome Biology and Evolution*, 16(11), p. evae210. Available at: https://doi.org/10.1093/gbe/evae210.

Briscoe, S.D. and Ragsdale, C.W. (2018) 'Homology, neocortex, and the evolution of developmental mechanisms', *Science (New York, N.Y.)*, 362(6411), pp. 190–193. Available at: https://doi.org/10.1126/science.aau3711.

Buenrostro, J.D. *et al.* (2013) 'Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position', *Nature Methods*, 10(12), pp. 1213–1218. Available at: https://doi.org/10.1038/nmeth.2688.

Bujalka, H. *et al.* (2013) 'MYRF Is a Membrane-Associated Transcription Factor That Autoproteolytically Cleaves to Directly Activate Myelin Genes', *PLoS Biology*, 11(8), p. e1001625. Available at: https://doi.org/10.1371/journal.pbio.1001625.

Bunt, J. *et al.* (2017) 'Combined allelic dosage of Nfia and Nfib regulates cortical development', *Brain and Neuroscience Advances*, 1, p. 2398212817739433. Available at: https://doi.org/10.1177/2398212817739433.

Camp, J.G. *et al.* (2015) 'Human cerebral organoids recapitulate gene expression programs of fetal neocortex development', *Proceedings of the National Academy of Sciences*, 112(51), pp. 15672–15677. Available at: https://doi.org/10.1073/pnas.1520760112.

Cantera, R. *et al.* (2002) 'Mutations in spalt cause a severe but reversible neurodegenerative phenotype in the embryonic central nervous system of Drosophila melanogaster', *Development (Cambridge, England)*, 129(24), pp. 5577–5586. Available at: https://doi.org/10.1242/dev.00158.

Capra, J.A. *et al.* (2013) 'Many human accelerated regions are developmental enhancers', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 368(1632), p. 20130025. Available at: https://doi.org/10.1098/rstb.2013.0025.

Cárdenas, A. *et al.* (2018) 'Evolution of Cortical Neurogenesis in Amniotes Controlled by Robo Signaling Levels', *Cell*, 174(3), pp. 590-606.e21. Available at: https://doi.org/10.1016/j.cell.2018.06.007.

Cárdenas, A. and Borrell, V. (2020) 'Molecular and cellular evolution of corticogenesis in amniotes', *Cellular and Molecular Life Sciences*, 77(8), pp. 1435–1460. Available at: https://doi.org/10.1007/s00018-019-03315-x.

Carroll, S.B. (2005) 'Evolution at Two Levels: On Genes and Form', *PLOS Biology*, 3(7), p. e245. Available at: https://doi.org/10.1371/journal.pbio.0030245.

Carroll, S.B. (2008) 'Evo-Devo and an Expanding Evolutionary Synthesis: A Genetic Theory of Morphological Evolution', *Cell*, 134(1), pp. 25–36. Available at: https://doi.org/10.1016/j.cell.2008.06.030.

Caviness, V.S., Takahashi, T. and Nowakowski, R.S. (1995) 'Numbers, time and neocortical neuronogenesis: a general developmental and evolutionary model', *Trends in Neurosciences*, 18(9), pp. 379–383. Available at: https://doi.org/10.1016/0166-2236(95)93933-o.

de Celis, J.F., Barrio, R. and Kafatos, F.C. (1999) 'Regulation of the spalt/spalt-related gene complex and its function during sensory organ development in the Drosophila thorax', *Development (Cambridge, England)*, 126(12), pp. 2653–2662. Available at: https://doi.org/10.1242/dev.126.12.2653.

Chang, W. *et al.* (2024) *shiny: Web Application Framework for R*. Available at: https://shiny.posit.co/.

Charvet, C.J., Owerkowicz, T. and Striedter, G.F. (2009) 'Phylogeny of the telencephalic subventricular zone in sauropsids: evidence for the sequential evolution of pallial and subpallial subventricular zones', *Brain, Behavior and Evolution*, 73(4), pp. 285–294. Available at: https://doi.org/10.1159/000230673.

Chen, H.-Y. *et al.* (2016) 'Musashi-1 regulates AKT-derived IL-6 autocrinal/paracrinal malignancy and chemoresistance in glioblastoma', *Oncotarget*, 7(27), pp. 42485–42501. Available at: https://doi.org/10.18632/oncotarget.9890.

Chen, Q. *et al.* (2015) 'CXCR7 Mediates Neural Progenitor Cells Migration to CXCL12 Independent of CXCR4', *Stem cells (Dayton, Ohio)*, 33(8), pp. 2574–2585. Available at: https://doi.org/10.1002/stem.2022.

Chen, S., Krinsky, B.H. and Long, M. (2013) 'New genes as drivers of phenotypic evolution', *Nature Reviews Genetics*, 14(9), pp. 645–660. Available at: https://doi.org/10.1038/nrg3521.

Chen, Y. and Wang, X. (2020) 'miRDB: an online database for prediction of functional microRNA targets', *Nucleic Acids Research*, 48(D1), pp. D127–D131. Available at: https://doi.org/10.1093/nar/gkz757.

Cheng, G. *et al.* (2016) 'Positive expression of NR6A1/CT150 as a predictor of biochemical recurrence-free survival in prostate cancer patients', *Oncotarget*, 8(38), pp. 64427–64439. Available at: https://doi.org/10.18632/oncotarget.11749.

Chenn, A. and Walsh, C.A. (2002) 'Regulation of cerebral cortical size by control of cell cycle exit in neural precursors', *Science (New York, N.Y.)*, 297(5580), pp. 365–369. Available at: https://doi.org/10.1126/science.1074192.

Cheung, A.F.P. *et al.* (2007) 'Comparative aspects of cortical neurogenesis in vertebrates', *Journal of Anatomy*, 211(2), pp. 164–176. Available at: https://doi.org/10.1111/j.1469-7580.2007.00769.x.

Cheung, A.F.P. *et al.* (2010) 'The subventricular zone is the developmental milestone of a 6-layered neocortex: comparisons in metatherian and eutherian mammals', *Cerebral Cortex (New York, N.Y.: 1991)*, 20(5), pp. 1071–1081. Available at: https://doi.org/10.1093/cercor/bhp168.

Chinnappa, K. *et al.* (2022) 'Secondary loss of miR-3607 reduced cortical progenitor amplification during rodent evolution', *Science Advances*, 8(2), p. eabj4010. Available at: https://doi.org/10.1126/sciadv.abj4010.

Choi, J. *et al.* (2018) 'Elucidating the transactivation domain of the pleiotropic transcription factor Myrf', *Scientific Reports*, 8(1), p. 13075. Available at: https://doi.org/10.1038/s41598-018-31477-4.

Chou, F.-S., Li, R. and Wang, P.-S. (2018) 'Molecular components and polarity of radial glial cells during cerebral cortex development', *Cellular and Molecular Life Sciences*, 75(6), pp. 1027–1041. Available at: https://doi.org/10.1007/s00018-017-2680-0.

Clinton, B.K. *et al.* (2014) 'Radial glia in the proliferative ventricular zone of the embryonic and adult turtle, Trachemys scripta elegans', *Neurogenesis*, 1(1), p. e970905. Available at: https://doi.org/10.4161/23262125.2014.970905.

Clough, E. *et al.* (2024) 'NCBI GEO: archive for gene expression and epigenomics data sets: 23-year update', *Nucleic Acids Research*, 52(D1), pp. D138–D144. Available at: https://doi.org/10.1093/nar/gkad965.

Cotino-Nájera, S. *et al.* (2024) 'The role of Lin28A and Lin28B in cancer beyond Let-7', *FEBS Letters*, 598(24), pp. 2963–2979. Available at: https://doi.org/10.1002/1873-3468.15004.

Cubelos, B. *et al.* (2010) 'Cux1 and Cux2 regulate dendritic branching, spine morphology and synapses of the upper layer neurons of the cortex', *Neuron*, 66(4), pp. 523–535. Available at: https://doi.org/10.1016/j.neuron.2010.04.038.

Cubillos-Ruiz, J.R. *et al.* (2015) 'ER Stress Sensor XBP1 Controls Anti-tumor Immunity by Disrupting Dendritic Cell Homeostasis', *Cell*, 161(7), pp. 1527–1538. Available at: https://doi.org/10.1016/j.cell.2015.05.025.

Dal Col, J. *et al.* (2022) 'Phospholipid scramblase 1: a protein with multiple functions via multiple molecular interactors', *Cell Communication and Signaling*, 20(1), p. 78. Available at: https://doi.org/10.1186/s12964-022-00895-3.

Danielli, S.G. *et al.* (2024) 'Single cell transcriptomic profiling identifies tumor-acquired and therapy-resistant cell states in pediatric rhabdomyosarcoma', *Nature Communications*, 15(1), p. 6307. Available at: https://doi.org/10.1038/s41467-024-50527-2.

D'Arcy, B.R. *et al.* (2023) 'Non-muscle myosins control radial glial basal endfeet to mediate interneuron organization', *PLOS Biology*, 21(2), p. e3001926. Available at: https://doi.org/10.1371/journal.pbio.3001926.

D'Arcy, B.R. and Silver, D.L. (2020) 'Local gene regulation in radial glia: Lessons from across the nervous system', *Traffic (Copenhagen, Denmark)*, 21(12), pp. 737–748. Available at: https://doi.org/10.1111/tra.12769.

Dave, B.M. *et al.* (2023) 'Directed differentiation of human hindbrain neuroepithelial stem cells recapitulates cerebellar granule neurogenesis', *Development*, 150(13), p. dev201534. Available at: https://doi.org/10.1242/dev.201534.

Day, B.W. *et al.* (2013) 'EphA3 Maintains Tumorigenicity and Is a Therapeutic Target in Glioblastoma Multiforme', *Cancer Cell*, 23(2), pp. 238–248. Available at: https://doi.org/10.1016/j.ccr.2013.01.007.

De Juan Romero, C. and Borrell, V. (2015) 'Coevolution of radial glial cells and the cerebral cortex', *Glia*, 63(8), pp. 1303–1319. Available at: https://doi.org/10.1002/glia.22827.

De Lima, L. *et al.* (2020) 'Central nervous system high-grade neuroepithelial tumor with BCOR alteration (CNS HGNET-BCOR)—case-based reviews', *Child's Nervous System*, 36(8), pp. 1589–1599. Available at: https://doi.org/10.1007/s00381-020-04692-6.

Dehay, C. and Kennedy, H. (2007) 'Cell-cycle control and cortical development', *Nature Reviews Neuroscience*, 8(6), pp. 438–450. Available at: https://doi.org/10.1038/nrn2097.

Delgado, R.N. and Lim, D.A. (2015) 'Embryonic *Nkx2.1*-expressing neural precursor cells contribute to the regional heterogeneity of adult V–SVZ neural stem cells', *Developmental Biology*, 407(2), pp. 265–274. Available at: https://doi.org/10.1016/j.ydbio.2015.09.008.

Del-Valle-Anton, L. *et al.* (2024) 'Multiple parallel cell lineages in the developing mammalian cerebral cortex', *Science Advances*, 10(13), p. eadn9998. Available at: https://doi.org/10.1126/sciadv.adn9998.

Di Bella, D.J. *et al.* (2021) 'Molecular logic of cellular diversification in the mouse cerebral cortex', *Nature*, 595(7868), pp. 554–559. Available at: https://doi.org/10.1038/s41586-021-03670-5.

Diaz, J.L. *et al.* (2020) 'An evolutionarily acquired microRNA shapes development of mammalian cortical projections', *Proceedings of the National Academy of Sciences of the United States of America*, 117(46), pp. 29113–29122. Available at: https://doi.org/10.1073/pnas.2006700117.

Dj, T. *et al.* (2002) 'Coordination of ATF6-mediated transcription and ATF6 degradation by a domain that is shared with the viral transcription factor, VP16', *The Journal of biological chemistry*, 277(23). Available at: https://doi.org/10.1074/jbc.M201749200.

Docampo-Seara, A. *et al.* (2018) 'Study of pallial neurogenesis in shark embryos and the evolutionary origin of the subventricular zone', *Brain Structure & Function*, 223(8), pp. 3593–3612. Available at: https://doi.org/10.1007/s00429-018-1705-2.

Drill, E.A. (2009) *Deciphering the role of tlx in dorsal neural progenitors and its contribution to brain structure and behavior*. University of Pittsburgh. Available at: https://d-scholarship.pitt.edu/8513/ (Accessed: 30 December 2024).

Dugas-Ford, J. and Ragsdale, C.W. (2015) 'Levels of Homology and the Problem of Neocortex', *Annual Review of Neuroscience*, 38(1), pp. 351–368. Available at: https://doi.org/10.1146/annurev-neuro-071714-033911.

Englund, C. *et al.* (2005) 'Pax6, Tbr2, and Tbr1 are expressed sequentially by radial glia, intermediate progenitor cells, and postmitotic neurons in developing neocortex', *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 25(1), pp. 247–251. Available at: https://doi.org/10.1523/JNEUROSCI.2899-04.2005.

Espinós, A. *et al.* (2022) 'Evolution of genetic mechanisms regulating cortical neurogenesis', *Developmental Neurobiology*, 82(5), pp. 428–453. Available at: https://doi.org/10.1002/dneu.22891.

Faedo, A. *et al.* (2008) 'COUP-TFI Coordinates Cortical Patterning, Neurogenesis, and Laminar Fate and Modulates MAPK/ERK, AKT, and ß-Catenin Signaling', *Cerebral Cortex*, 18(9), pp. 2117–2131. Available at: https://doi.org/10.1093/cercor/bhm238.

Fan, C. *et al.* (2021) 'Functional mechanisms of MYRF DNA-binding domain mutations implicated in birth defects', *Journal of Biological Chemistry*, 296. Available at: https://doi.org/10.1016/j.jbc.2021.100612.

Farkas, L.M. *et al.* (2008) 'Insulinoma-Associated 1 Has a Panneurogenic Role and Promotes the Generation and Expansion of Basal Progenitors in the Developing Mouse Neocortex', *Neuron*, 60(1), pp. 40–55. Available at: https://doi.org/10.1016/j.neuron.2008.09.020.

Federico, A. *et al.* (2022) 'ATRT–SHH comprises three molecular subgroups with characteristic clinical and histopathological features and prognostic significance', *Acta Neuropathologica*, 143(6), pp. 697–711. Available at: https://doi.org/10.1007/s00401-022-02424-5.

Fernandez, A.S. *et al.* (1998) 'Expression of the Emx-1 and Dlx-1 homeobox genes define three molecularly distinct domains in the telencephalon of mouse, chick, turtle and frog embryos: implications for the evolution of telencephalic subdivisions in amniotes', *Development*, 125(11), pp. 2099–2111. Available at: https://doi.org/10.1242/dev.125.11.2099.

Fernández, V. *et al.* (2020) 'Repression of Irs2 by let-7 miRNAs is essential for homeostasis of the telencephalic neuroepithelium', *The EMBO Journal*, 39(21), p. e105479. Available at: https://doi.org/10.15252/embj.2020105479.

Fiddes, I.T. *et al.* (2018) 'Human-Specific NOTCH2NL Genes Affect Notch Signaling and Cortical Neurogenesis', *Cell*, 173(6), pp. 1356-1369.e22. Available at: https://doi.org/10.1016/j.cell.2018.03.051.

Fietz, S.A. *et al.* (2010) 'OSVZ progenitors of human and ferret neocortex are epithelial-like and expand by integrin signaling', *Nature Neuroscience*, 13(6), pp. 690–699. Available at: https://doi.org/10.1038/nn.2553.

Filbin, M.G. *et al.* (2018) 'Developmental and oncogenic programs in H3K27M gliomas dissected by single-cell RNA-seq', *Science*, 360(6386), pp. 331–335. Available at: https://doi.org/10.1126/science.aao4750.

Fisher, R.A. (1932) *Statistical methods for research workers, 4th ed*. Oxford, England: Oliver & Boyd (Statistical methods for research workers, 4th ed), pp. xi, 307.

Florio, M. *et al.* (2015) 'Human-specific gene ARHGAP11B promotes basal progenitor amplification and neocortex expansion', *Science*, 347(6229), pp. 1465–1470. Available at: https://doi.org/10.1126/science.aaa1975.

Florio, M. *et al.* (2016) 'A single splice site mutation in human-specific ARHGAP11B causes basal progenitor amplification', *Science Advances*, 2(12), p. e1601941. Available at: https://doi.org/10.1126/sciadv.1601941.

Florio, M. *et al.* (2018) 'Evolution and cell-type specificity of human-specific genes preferentially expressed in progenitors of fetal neocortex', *eLife*. Edited by J.G. Gleeson, 7, p. e32332. Available at: https://doi.org/10.7554/eLife.32332.

Fortna, A. *et al.* (2004) 'Lineage-Specific Gene Duplication and Loss in Human and Great Ape Evolution', *PLOS Biology*, 2(7), p. e207. Available at: https://doi.org/10.1371/journal.pbio.0020207.

Franco, S.J. *et al.* (2012) 'Fate-Restricted Neural Progenitors in the Mammalian Cerebral Cortex', *Science*, 337(6095), pp. 746–749. Available at: https://doi.org/10.1126/science.1223616.

Fu, Y. *et al.* (2014) 'Gene expression regulation mediated through reversible m^6A RNA methylation', *Nature Reviews. Genetics*, 15(5), pp. 293–306. Available at: https://doi.org/10.1038/nrg3724.

Furukawa, T., Kozak, C.A. and Cepko, C.L. (1997) 'rax, a novel paired-type homeobox gene, shows expression in the anterior neural fold and developing retina', *Proceedings of the National Academy of Sciences*, 94(7), pp. 3088–3093. Available at: https://doi.org/10.1073/pnas.94.7.3088.

Galderisi, U., Jori, F.P. and Giordano, A. (2003) 'Cell cycle regulation and neural differentiation', *Oncogene*, 22(33), pp. 5208–5219. Available at: https://doi.org/10.1038/sj.onc.1206558.

García-Moreno, F. *et al.* (2012) 'Compartmentalization of Cerebral Cortical Germinal Zones in a Lissencephalic Primate and Gyrencephalic Rodent', *Cerebral Cortex*, 22(2), pp. 482–492. Available at: https://doi.org/10.1093/cercor/bhr312.

García-Moreno, F. *et al.* (2018) 'Absence of Tangentially Migrating Glutamatergic Neurons in the Developing Avian Brain', *Cell Reports*, 22(1), pp. 96–109. Available at: https://doi.org/10.1016/j.celrep.2017.12.032.

Geller, E. *et al.* (2024) 'Massively parallel disruption of enhancers active in human neural stem cells', *Cell Reports*, 43(2), p. 113693. Available at: https://doi.org/10.1016/j.celrep.2024.113693.

Gessi, M. *et al.* (2009) 'Embryonal Tumors With Abundant Neuropil and True Rosettes', *The American journal of surgical pathology*, 33(2), pp. 211–217. Available at: https://doi.org/10.1097/PAS.0b013e318186235b.

Gil-Sanz, C. *et al.* (2015) 'Matters Arising: Lineage tracing using Cux2-Cre and Cux2-CreERT2 mice', *Neuron*, 86(4), pp. 1091–1099. Available at: https://doi.org/10.1016/j.neuron.2015.04.019.

Girskis, K.M. *et al.* (2021) 'Rewiring of human neurodevelopmental gene regulatory programs by human accelerated regions', *Neuron*, 109(20), pp. 3239-3251.e7. Available at: https://doi.org/10.1016/j.neuron.2021.08.005.

Glaviano, A. *et al.* (2023) 'PI3K/AKT/mTOR signaling transduction pathway and targeted therapies in cancer', *Molecular Cancer*, 22(1), p. 138. Available at: https://doi.org/10.1186/s12943-023-01827-6.

Goggolidou, P. *et al.* (2013) 'A chronological expression profile of gene activity during embryonic mouse brain development', *Mammalian Genome*, 24(11), pp. 459–472. Available at: https://doi.org/10.1007/s00335-013-9486-7.

Goodfellow, M. *et al.* (2014) 'microRNA input into a neural ultradian oscillator controls emergence and timing of alternative cell states', *Nature Communications*, 5(1), p. 3399. Available at: https://doi.org/10.1038/ncomms4399.

Götz, M. and Huttner, W.B. (2005) 'The cell biology of neurogenesis', *Nature Reviews. Molecular Cell Biology*, 6(10), pp. 777–788. Available at: https://doi.org/10.1038/nrm1739.

Götz, M., Stoykova, A. and Gruss, P. (1998) 'Pax6 controls radial glia differentiation in the cerebral cortex', *Neuron*, 21(5), pp. 1031–1044. Available at: https://doi.org/10.1016/s0896-6273(00)80621-2.

Grant, C.E., Bailey, T.L. and Noble, W.S. (2011) 'FIMO: scanning for occurrences of a given motif', *Bioinformatics*, 27(7), pp. 1017–1018. Available at: https://doi.org/10.1093/bioinformatics/btr064.

Greb-Markiewicz, B. *et al.* (2019) 'The subcellular localization of bHLH transcription factor TCF4 is mediated by multiple nuclear localization and nuclear export signals', *Scientific Reports*, 9(1), p. 15629. Available at: https://doi.org/10.1038/s41598-019-52239-w.

Gu, S. *et al.* (2012) 'The Loop Position of shRNAs and Pre-miRNAs Is Critical for the Accuracy of Dicer Processing In Vivo', *Cell*, 151(4), pp. 900–911. Available at: https://doi.org/10.1016/j.cell.2012.09.042.

Gualano, F.M. *et al.* (2023) 'Embryonal tumor with multilayered rosettes: Post-treatment maturation and implications for future therapy', *Cancer Reports*, 6(5), p. e1812. Available at: https://doi.org/10.1002/cnr2.1812.

Gupta, A., Read, D.E. and Gupta, S. (2015) 'Assays for induction of the unfolded protein response and selective activation of the three major pathways', *Methods in Molecular Biology (Clifton, N.J.)*, 1292, pp. 19–38. Available at: https://doi.org/10.1007/978-1-4939-2522-3_2.

Gupta, S. *et al.* (2007) 'Quantifying similarity between motifs', *Genome Biology*, 8(2), p. R24. Available at: https://doi.org/10.1186/gb-2007-8-2-r24.

Gyllborg, D. *et al.* (2018) 'The Matricellular Protein R-Spondin 2 Promotes Midbrain Dopaminergic Neurogenesis and Differentiation', *Stem Cell Reports*, 11(3), pp. 651–664. Available at: https://doi.org/10.1016/j.stemcr.2018.07.014.

Han, J.S. *et al.* (2023) 'Notch directs telencephalic development and controls neocortical neuron fate determination by regulating microRNA levels', *Development (Cambridge, England)*, 150(11), p. dev201408. Available at: https://doi.org/10.1242/dev.201408.

Hansen, D.V. *et al.* (2010) 'Neurogenic radial glia in the outer subventricular zone of human neocortex', *Nature*, 464(7288), pp. 554–561. Available at: https://doi.org/10.1038/nature08845.

Hao, Y. *et al.* (2024) 'Dictionary learning for integrative, multimodal and scalable single-cell analysis', *Nature Biotechnology*, 42(2), pp. 293–304. Available at: https://doi.org/10.1038/s41587-023-01767-y.

Harris, J.A. *et al.* (2019) 'Hierarchical organization of cortical and thalamic connectivity', *Nature*, 575(7781), pp. 195–202. Available at: https://doi.org/10.1038/s41586-019-1716-z.

Harrison, P.W. *et al.* (2024) 'Ensembl 2024', *Nucleic Acids Research*, 52(D1), pp. D891–D899. Available at: https://doi.org/10.1093/nar/gkad1049.

Harrison, S.J. *et al.* (2012) 'Sall1 regulates cortical neurogenesis and laminar fate specification in mice: implications for neural abnormalities in Townes-Brocks syndrome', *Disease Models & Mechanisms*, 5(3), pp. 351–365. Available at: https://doi.org/10.1242/dmm.002873.

Harrison, S.J., Nishinakamura, R. and Monaghan, A.P. (2008) 'Sall1 regulates mitral cell development and olfactory nerve extension in the developing olfactory bulb', *Cerebral Cortex (New York, N.Y.: 1991)*, 18(7), pp. 1604–1617. Available at: https://doi.org/10.1093/cercor/bhm191.

Harrison, S.J., Parrish, M. and Monaghan, A.P. (2008) 'Sall3 is required for the terminal maturation of olfactory glomerular interneurons', *The Journal of Comparative Neurology*, 507(5), pp. 1780–1794. Available at: https://doi.org/10.1002/cne.21650.

Haubensak, W. *et al.* (2004) 'Neurons arise in the basal neuroepithelium of the early mammalian telencephalon: a major site of neurogenesis', *Proceedings of the National Academy of Sciences of the United States of America*, 101(9), pp. 3196–3201. Available at: https://doi.org/10.1073/pnas.0308600100.

Haubst, N. *et al.* (2004) 'Molecular dissection of Pax6 function: the specific roles of the paired domain and homeodomain in brain development', *Development (Cambridge, England)*, 131(24), pp. 6131–6140. Available at: https://doi.org/10.1242/dev.01524.

He, X. *et al.* (2020) 'RTP4 inhibits IFN-I response and enhances experimental cerebral malaria and neuropathology', *Proceedings of the National Academy of Sciences of the United States of America*, 117(32), pp. 19465–19474. Available at: https://doi.org/10.1073/pnas.2006492117.

Heide, M. and Huttner, W.B. (2021) 'Human-Specific Genes, Cortical Progenitor Cells, and Microcephaly', *Cells*, 10(5), p. 1209. Available at: https://doi.org/10.3390/cells10051209.

Heins, N. *et al.* (2002) 'Glial cells generate neurons: the role of the transcription factor Pax6', *Nature Neuroscience*, 5(4), pp. 308–315. Available at: https://doi.org/10.1038/nn828.

Heintzman, N.D. *et al.* (2007) 'Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome', *Nature Genetics*, 39(3), pp. 311–318. Available at: https://doi.org/10.1038/ng1966.

Heumos, L. *et al.* (2023) 'Best practices for single-cell analysis across modalities', *Nature Reviews Genetics*, 24(8), pp. 550–572. Available at: https://doi.org/10.1038/s41576-023-00586-w.

Hevner, R.F. *et al.* (2006) 'Transcription factors in glutamatergic neurogenesis: conserved programs in neocortex, cerebellum, and adult hippocampus', *Neuroscience Research*, 55(3), pp. 223–233. Available at: https://doi.org/10.1016/j.neures.2006.03.004.

Hevner, R.F. (2023) 'Intermediate Progenitors in Neocortical Development and Evolution', in *Neocortical Neurogenesis in Development and Evolution*. John Wiley & Sons, Ltd, pp. 41–61. Available at: https://doi.org/10.1002/9781119860914.ch3.

Hinds, J.W. and Ruffett, T.L. (1971) 'Cell proliferation in the neural tube: An electron microscopic and Golgi analysis in the mouse cerebral vesicle', *Zeitschrift für Zellforschung und Mikroskopische Anatomie*, 115(2), pp. 226–264. Available at: https://doi.org/10.1007/BF00391127.

Hitz, B.C. *et al.* (2023) 'The ENCODE Uniform Analysis Pipelines', *bioRxiv: The Preprint Server for Biology*, p. 2023.04.04.535623. Available at: https://doi.org/10.1101/2023.04.04.535623.

Ho, B. *et al.* (2020) 'Molecular subgrouping of atypical teratoid/rhabdoid tumors-a reinvestigation and current consensus', *Neuro-Oncology*, 22(5), pp. 613–624. Available at: https://doi.org/10.1093/neuonc/noz235.

Horwitz, M. *et al.* (2016) 'Embryonal tumors with multilayered rosettes in children: the SFCE experience', *Child's Nervous System: ChNS: Official Journal of the International Society for Pediatric Neurosurgery*, 32(2), pp. 299–305. Available at: https://doi.org/10.1007/s00381-015-2920-2.

Hsu, A.P. *et al.* (2019) 'Dominant activating RAC2 mutation with lymphopenia, immunodeficiency, and cytoskeletal defects', *Blood*, 133(18), pp. 1977–1988. Available at: https://doi.org/10.1182/blood-2018-11-886028.

Huang, H. *et al.* (2021) 'MYRF: A Mysterious Membrane-Bound Transcription Factor Involved in Myelin Development and Human Diseases', *Neuroscience Bulletin*, 37(6), pp. 881–884. Available at: https://doi.org/10.1007/s12264-021-00678-9.

Huang, P. *et al.* (2020) 'Nuclear translocation of PLSCR1 activates STAT1 signaling in basal-like breast cancer', *Theranostics*, 10(10), pp. 4644–4658. Available at: https://doi.org/10.7150/thno.43150.

Huang, Y.-W. *et al.* (2024) 'Adenomatous polyposis coli (APC) regulates internalization and signaling of the chemorepellent receptor, Roundabout (ROBO) 1'. bioRxiv, p. 2024.01.12.574717. Available at: https://doi.org/10.1101/2024.01.12.574717.

Huilgol, D. *et al.* (2023) 'Direct and indirect neurogenesis generate a mosaic of distinct glutamatergic projection neuron types in cerebral cortex', *Neuron*, 111(16), pp. 2557-2569.e4. Available at: https://doi.org/10.1016/j.neuron.2023.05.021.

Hwang, H.C. and Clurman, B.E. (2005) 'Cyclin E in normal and neoplastic cell cycles', *Oncogene*, 24(17), pp. 2776–2786. Available at: https://doi.org/10.1038/sj.onc.1208613.

Imayoshi, I. *et al.* (2013) 'Oscillatory control of factors determining multipotency and fate in mouse neural progenitors', *Science (New York, N.Y.)*, 342(6163), pp. 1203–1208. Available at: https://doi.org/10.1126/science.1242366.

Imayoshi, I., Ishidate, F. and Kageyama, R. (2015) 'Real-time imaging of bHLH transcription factors reveals their dynamic control in the multipotency and fate choice of neural stem cells', *Frontiers in Cellular Neuroscience*, 9, p. 288. Available at: https://doi.org/10.3389/fncel.2015.00288.

Imitola, J. *et al.* (2023) 'Stat1 is an inducible transcriptional repressor of neural stem cells self-renewal program during neuroinflammation', *Frontiers in Cellular Neuroscience*, 17, p. 1156802. Available at: https://doi.org/10.3389/fncel.2023.1156802.

Inglis-Broadgate, S.L. *et al.* (2005) 'FGFR3 regulates brain size by controlling progenitor cell proliferation and apoptosis during embryonic development', *Developmental Biology*, 279(1), pp. 73–85. Available at: https://doi.org/10.1016/j.ydbio.2004.11.035.

Ismail, V. *et al.* (2022) 'Identification and functional evaluation of GRIA1 missense and truncation variants in individuals with ID: An emerging neurodevelopmental syndrome', *The American Journal of Human Genetics*, 109(7), pp. 1217–1241. Available at: https://doi.org/10.1016/j.ajhg.2022.05.009.

Isogai, E. *et al.* (2015) 'Oncogenic Lmo3 cooperates with Hen2 to induce hydrocephalus in mice', *Experimental Animals*, 64(4), pp. 407–414. Available at: https://doi.org/10.1538/expanim.15-0026.

Iwata, R. and Vanderhaeghen, P. (2021) 'Regulatory roles of mitochondria and metabolism in neurogenesis', *Current Opinion in Neurobiology*, 69, pp. 231–240. Available at: https://doi.org/10.1016/j.conb.2021.05.003.

Izumi, H. and Kaneko, Y. (2012) 'Evidence of asymmetric cell division and centrosome inheritance in human neuroblastoma cells', *Proceedings of the National Academy of Sciences of the United States of America*, 109(44), pp. 18048–18053. Available at: https://doi.org/10.1073/pnas.1205525109.

Jarvis, E.D. *et al.* (2005) 'Avian brains and a new understanding of vertebrate brain evolution', *Nature Reviews. Neuroscience*, 6(2), pp. 151–159. Available at: https://doi.org/10.1038/nrn1606.

Jaworski, A., Long, H. and Tessier-Lavigne, M. (2010) 'Collaborative and Specialized Functions of Robo1 and Robo2 in Spinal Commissural Axon Guidance', *The Journal of Neuroscience*, 30(28), pp. 9445–9453. Available at: https://doi.org/10.1523/JNEUROSCI.6290-09.2010.

Jeong, H. *et al.* (2021) 'Evolution of DNA methylation in the human brain', *Nature Communications*, 12(1), p. 2021. Available at: https://doi.org/10.1038/s41467-021-21917-7.

Jessa, S. *et al.* (2019) 'Stalled developmental programs at the root of pediatric brain tumors', *Nature genetics*, 51(12), pp. 1702–1713. Available at: https://doi.org/10.1038/s41588-019-0531-7.

Jiang, F. *et al.* (2023) 'Simultaneous profiling of spatial gene expression and chromatin accessibility during mouse brain development', *Nature Methods*, 20(7), pp. 1048–1057. Available at: https://doi.org/10.1038/s41592-023-01884-1.

Jo, M. *et al.* (2024) 'Fluorescent tagging of endogenous IRS2 with an auxin-dependent degron to assess dynamic intracellular localization and function', *Journal of Biological Chemistry*, 300(11). Available at: https://doi.org/10.1016/j.jbc.2024.107796.

Johann, P.D. *et al.* (2016) 'Atypical Teratoid/Rhabdoid Tumors Are Comprised of Three Epigenetic Subgroups with Distinct Enhancer Landscapes', *Cancer Cell*, 29(3), pp. 379–393. Available at: https://doi.org/10.1016/j.ccell.2016.02.001.

Johnson, D.S. *et al.* (2007) 'Genome-wide mapping of in vivo protein-DNA interactions', *Science (New York, N.Y.)*, 316(5830), pp. 1497–1502. Available at: https://doi.org/10.1126/science.1141319.

Kageyama, R. *et al.* (2008) 'Dynamic Notch signaling in neural progenitor cells and a revised view of lateral inhibition', *Nature Neuroscience*, 11(11), pp. 1247–1251. Available at: https://doi.org/10.1038/nn.2208.

Kaise, T. and Kageyama, R. (2021) 'Hes1 oscillation frequency correlates with activation of neural stem cells', *Gene Expression Patterns*, 40, p. 119170. Available at: https://doi.org/10.1016/j.gep.2021.119170.

Kamihara, J. *et al.* (2020) '*DICER1*-associated central nervous system sarcoma in children: comprehensive clinicopathologic and genetic analysis of a newly described rare tumor', *Modern Pathology*, 33(10), pp. 1910–1921. Available at: https://doi.org/10.1038/s41379-020-0516-1.

Kan, W. *et al.* (2024) 'Downregulating DNA methyltransferase 3B by suppressing the PI3K/Akt signaling pathway enhances the chemosensitivity of glioblastoma to temozolomide', *Molecular Neurobiology*, 61(9), pp. 7066–7074. Available at: https://doi.org/10.1007/s12035-024-04041-7.

Kaneko, K. *et al.* (2016) 'Neogenin, Defined as a GD3-associated Molecule by Enzyme-mediated Activation of Radical Sources, Confers Malignant Properties via Intracytoplasmic Domain in Melanoma Cells *', *Journal of Biological Chemistry*, 291(32), pp. 16630–16643. Available at: https://doi.org/10.1074/jbc.M115.708834.

Kang, H.J. *et al.* (2011) 'Spatio-temporal transcriptome of the human brain', *Nature*, 478(7370), pp. 483–489. Available at: https://doi.org/10.1038/nature10523.

Karaca, E. *et al.* (2015) 'Genes that Affect Brain Structure and Function Identified by Rare Variant Analyses of Mendelian Neurologic Disease', *Neuron*, 88(3), pp. 499–513. Available at: https://doi.org/10.1016/j.neuron.2015.09.048.

Karimian, A., Ahmadi, Y. and Yousefi, B. (2016) 'Multiple functions of p21 in cell cycle, apoptosis and transcriptional regulation after DNA damage', *DNA Repair*, 42, pp. 63–71. Available at: https://doi.org/10.1016/j.dnarep.2016.04.008.

Katane, M. *et al.* (2018) 'Structural and enzymatic properties of mammalian d-glutamate cyclase', *Archives of Biochemistry and Biophysics*, 654, pp. 10–18. Available at: https://doi.org/10.1016/j.abb.2018.07.005.

Kawase, T. *et al.* (2009) 'PH Domain-Only Protein PHLDA3 Is a p53-Regulated Repressor of Akt', *Cell*, 136(3), pp. 535–550. Available at: https://doi.org/10.1016/j.cell.2008.12.002.

Kawaue, T. *et al.* (2014) 'Neurogenin2-d4Venus and Gadd45g-d4Venus transgenic mice: Visualizing mitotic and migratory behaviors of cells committed to the neuronal lineage in the developing mammalian brain', *Development, Growth & Differentiation*, 56(4), pp. 293–304. Available at: https://doi.org/10.1111/dgd.12131.

Kelava, I. *et al.* (2012) 'Abundant Occurrence of Basal Radial Glia in the Subventricular Zone of Embryonic Neocortex of a Lissencephalic Primate, the Common Marmoset Callithrix jacchus', *Cerebral Cortex (New York, NY)*, 22(2), pp. 469–481. Available at: https://doi.org/10.1093/cercor/bhr301.

Keough, K.C. *et al.* (2023) 'Three-dimensional genome rewiring in loci with human accelerated regions', *Science (New York, N.Y.)*, 380(6643), p. eabm1696. Available at: https://doi.org/10.1126/science.abm1696.

Kerimoglu, C. *et al.* (2021) 'H3 acetylation selectively promotes basal progenitor proliferation and neocortex expansion', *Science Advances*, 7(38), p. eabc6792. Available at: https://doi.org/10.1126/sciadv.abc6792.

Khacho, M., Harris, R. and Slack, R.S. (2019) 'Mitochondria as central regulators of neural stem cell fate and cognitive function', *Nature Reviews Neuroscience*, 20(1), pp. 34–48. Available at: https://doi.org/10.1038/s41583-018-0091-3.

Kikkawa, T. *et al.* (2013) 'Dmrta1 regulates proneural gene expression downstream of Pax6 in the mammalian telencephalon', *Genes to Cells*, 18(8), pp. 636–649. Available at: https://doi.org/10.1111/gtc.12061.

Kikkawa, T. and Osumi, N. (2021) 'Multiple Functions of the Dmrt Genes in the Development of the Central Nervous System', *Frontiers in Neuroscience*, 15, p. 789583. Available at: https://doi.org/10.3389/fnins.2021.789583.

Kim, D., Langmead, B. and Salzberg, S.L. (2015) 'HISAT: a fast spliced aligner with low memory requirements', *Nature Methods*, 12(4), pp. 357–360. Available at: https://doi.org/10.1038/nmeth.3317.

Kim, G.B. *et al.* (2021) 'DeepTFactor: A deep learning-based tool for the prediction of transcription factors', *Proceedings of the National Academy of Sciences*, 118(2), p. e2021171118. Available at: https://doi.org/10.1073/pnas.2021171118.

Kim, J.H., Kim, W.S. and Park, C. (2019) 'Interleukin-6 mediates resistance to PI3K-pathway–targeted therapy in lymphoma', *BMC Cancer*, 19(1), p. 936. Available at: https://doi.org/10.1186/s12885-019-6057-7.

Kim, J.-Y. *et al.* (2017) 'The value of phosphohistone H3 as a proliferation marker for evaluating invasive breast cancers: A comparative study with Ki67', *Oncotarget*, 8(39), pp. 65064–65076. Available at: https://doi.org/10.18632/oncotarget.17775.

Kirjavainen, A. *et al.* (2022) 'Gata2, Nkx2-2 and Skor2 form a transcription factor network regulating development of a midbrain GABAergic neuron subtype with characteristics of REM-sleep regulatory neurons', *Development*, 149(14), p. dev200937. Available at: https://doi.org/10.1242/dev.200937.

Kirsch, M. *et al.* (2009) 'Frequent loss of the CDKN2C (p18INK4c) gene product in pituitary adenomas', *Genes, Chromosomes & Cancer*, 48(2), pp. 143–154. Available at: https://doi.org/10.1002/gcc.20621.

Kishi, Y. *et al.* (2012) 'HMGA regulates the global chromatin state and neurogenic potential in neocortical precursor cells', *Nature Neuroscience*, 15(8), pp. 1127–1133. Available at: https://doi.org/10.1038/nn.3165.

Kleinman, C.L. *et al.* (2014) 'Fusion of TTYH1 with the C19MC microRNA cluster drives expression of a brain-specific DNMT3B isoform in the embryonal brain tumor ETMR', *Nature Genetics*, 46(1), pp. 39–44. Available at: https://doi.org/10.1038/ng.2849.

Koelsche, C. *et al.* (2018) 'Primary intracranial spindle cell sarcoma with rhabdomyosarcoma-like features share a highly distinct methylation profile and DICER1 mutations', *Acta Neuropathologica*, 136(2), pp. 327–337. Available at: https://doi.org/10.1007/s00401-018-1871-6.

Kok, Y.P. *et al.* (2020) 'Overexpression of Cyclin E1 or Cdc25A leads to replication stress, mitotic aberrancies, and increased sensitivity to replication checkpoint inhibitors', *Oncogenesis*, 9(10), pp. 1–15. Available at: https://doi.org/10.1038/s41389-020-00270-2.

Kokhlikyan, N. *et al.* (2020) 'Captum: A unified and generic model interpretability library for PyTorch'.

Kolde, R. (2018) *pheatmap: Pretty Heatmaps*. Available at: https://github.com/raivokolde/pheatmap.

Korshunov, A. *et al.* (2012) 'LIN28A immunoreactivity is a potent diagnostic marker of embryonal tumor with multilayered rosettes (ETMR)', *Acta Neuropathologica*, 124(6), pp. 875–881. Available at: https://doi.org/10.1007/s00401-012-1068-3.

Korshunov, A. *et al.* (2014) 'Embryonal tumor with abundant neuropil and true rosettes (ETANTR), ependymoblastoma, and medulloepithelioma share molecular similarity and comprise a single clinicopathological entity', *Acta Neuropathologica*, 128(2), pp. 279–289. Available at: https://doi.org/10.1007/s00401-013-1228-0.

Kovaleva, I.E. *et al.* (2020) 'Mitochondrial localization of SESN2', *PLOS ONE*, 15(4), p. e0226862. Available at: https://doi.org/10.1371/journal.pone.0226862.

Kowalczyk, T. *et al.* (2009) 'Intermediate Neuronal Progenitors (Basal Progenitors) Produce Pyramidal–Projection Neurons for All Layers of Cerebral Cortex', *Cerebral Cortex*, 19(10), pp. 2439–2450. Available at: https://doi.org/10.1093/cercor/bhn260.

Koyama, Y. (2014) 'Signaling molecules regulating phenotypic conversions of astrocytes and glial scar formation in damaged nerve tissues', *Neurochemistry International*, 78, pp. 35–42. Available at: https://doi.org/10.1016/j.neuint.2014.08.005.

Kriegstein, A., Noctor, S. and Martínez-Cerdeño, V. (2006) 'Patterns of neural stem and progenitor cell division may underlie evolutionary cortical expansion', *Nature Reviews. Neuroscience*, 7(11), pp. 883–890. Available at: https://doi.org/10.1038/nrn2008.

Kuwayama, N. *et al.* (2023) 'HMGA2 directly mediates chromatin condensation in association with neuronal fate regulation', *Nature Communications*, 14, p. 6420. Available at: https://doi.org/10.1038/s41467-023-42094-9.

La Manno, G. *et al.* (2018) 'RNA velocity of single cells', *Nature*, 560(7719), pp. 494–498. Available at: https://doi.org/10.1038/s41586-018-0414-6.

Laguesse, S. *et al.* (2015) 'A Dynamic Unfolded Protein Response Contributes to the Control of Cortical Neurogenesis', *Developmental Cell*, 35(5), pp. 553–567. Available at: https://doi.org/10.1016/j.devcel.2015.11.005.

Lambert, S.A. *et al.* (2018) 'The Human Transcription Factors', *Cell*, 172(4), pp. 650–665. Available at: https://doi.org/10.1016/j.cell.2018.01.029.

Lambo, S. *et al.* (2019) 'The Molecular Landscape of ETMR at Diagnosis and Relapse', *Nature*, 576(7786), pp. 274–280. Available at: https://doi.org/10.1038/s41586-019-1815-x.

Lambo, S. *et al.* (2020) 'ETMR: a tumor entity in its infancy', *Acta Neuropathologica*, 140(3), pp. 249–266. Available at: https://doi.org/10.1007/s00401-020-02182-2.

LaMonica, B.E. *et al.* (2013) 'Mitotic spindle orientation predicts outer radial glial cell generation in human neocortex', *Nature Communications*, 4, p. 1665. Available at: https://doi.org/10.1038/ncomms2647.

Lee, T.I. *et al.* (2006) 'Control of Developmental Regulators by Polycomb in Human Embryonic Stem Cells', *Cell*, 125(2), pp. 301–313. Available at: https://doi.org/10.1016/j.cell.2006.02.043.

Leng, P. and Zhao, J. (2020) 'Transcription factors as molecular switches to regulate drought adaptation in maize', *Theoretical and Applied Genetics*, 133(5), pp. 1455–1465. Available at: https://doi.org/10.1007/s00122-019-03494-y.

Leone, P.E. *et al.* (2008) 'Deletions of CDKN2C in Multiple Myeloma: Biological and Clinical Implications', *Clinical Cancer Research*, 14(19), pp. 6033–6041. Available at: https://doi.org/10.1158/1078-0432.CCR-08-0347.

LeRoith, D., Scheinman, E.J. and Bitton-Worms, K. (2011) 'The Role of Insulin and Insulin-like Growth Factors in the Increased Risk of Cancer in Diabetes', *Rambam Maimonides Medical Journal*, 2(2), p. e0043. Available at: https://doi.org/10.5041/RMMJ.10043.

Lewis, N.A. *et al.* (2022) 'Histone H3.3 K27M chromatin functions implicate a network of neurodevelopmental factors including ASCL1 and NEUROD1 in DIPG', *Epigenetics & Chromatin*, 15, p. 18. Available at: https://doi.org/10.1186/s13072-022-00447-6.

Li, D. *et al.* (2023) 'Serine protease PRSS56, a novel cancer-testis antigen activated by DNA hypomethylation, promotes colorectal and gastric cancer progression via PI3K/AKT axis', *Cell & Bioscience*, 13(1), p. 124. Available at: https://doi.org/10.1186/s13578-023-01060-0.

Li, N. *et al.* (2012) 'Lin-28 Homologue A (LIN28A) Promotes Cell Cycle Progression via Regulation of Cyclin-dependent Kinase 2 (CDK2), Cyclin D1 (CCND1), and Cell Division Cycle 25 Homolog A (CDC25A) Expression in Cancer*', *Journal of Biological Chemistry*, 287(21), pp. 17386–17397. Available at: https://doi.org/10.1074/jbc.M111.321158.

Liao, Y. *et al.* (2021) 'OLIG2 maintenance is not essential for diffuse intrinsic pontine glioma cell line growth but regulates tumor phenotypes', *Neuro-Oncology*, 23(7), pp. 1183–1196. Available at: https://doi.org/10.1093/neuonc/noab016.

Lichtenthaler, S.F., Lemberg, M.K. and Fluhrer, R. (2018) 'Proteolytic ectodomain shedding of membrane proteins in mammals—hardware, concepts, and recent developments', *The EMBO Journal*, 37(15), p. e99456. Available at: https://doi.org/10.15252/embj.201899456.

Lien, W.-H. *et al.* (2006) 'αE-Catenin Controls Cerebral Cortical Size by Regulating the Hedgehog Signaling Pathway', *Science*, 311(5767), pp. 1609–1612. Available at: https://doi.org/10.1126/science.1121449.

Lin, Z. *et al.* (2022) 'Lin28 Regulates Cancer Cell Stemness for Tumour Progression', *Cancers*, 14(19), p. 4640. Available at: https://doi.org/10.3390/cancers14194640.

Lindberg, N. *et al.* (2014) 'Oncogenic Signaling Is Dominant to Cell of Origin and Dictates Astrocytic or Oligodendroglial Tumor Development from Oligodendrocyte Precursor Cells', *The Journal of Neuroscience*, 34(44), pp. 14644–14651. Available at: https://doi.org/10.1523/JNEUROSCI.2977-14.2014.

Lister, R. *et al.* (2013) 'Global epigenomic reconfiguration during mammalian brain development', *Science (New York, N.Y.)*, 341(6146), p. 1237905. Available at: https://doi.org/10.1126/science.1237905.

Liu, C. *et al.* (2011) 'Mosaic Analysis with Double Markers (MADM) Reveals Tumor Cell-of-Origin in Glioma', *Cell*, 146(2), pp. 209–221. Available at: https://doi.org/10.1016/j.cell.2011.06.014.

Liu, J. *et al.* (2025) 'A human-specific enhancer fine-tunes radial glia potency and corticogenesis', *Nature*, pp. 1–12. Available at: https://doi.org/10.1038/s41586-025-09002-1.

Liu, W. and Wang, X. (2019) 'Prediction of functional microRNA targets by integrative modeling of microRNA binding and target expression data', *Genome Biology*, 20(1), p. 18. Available at: https://doi.org/10.1186/s13059-019-1629-z.

Liu, Y. *et al.* (2018) 'The nuclear transportation routes of membrane-bound transcription factors', *Cell Communication and Signaling*, 16(1), p. 12. Available at: https://doi.org/10.1186/s12964-018-0224-3.

Loaeza-Loaeza, J. *et al.* (2022) 'DNMT3B overexpression downregulates genes with CpG islands, common motifs, and transcription factor binding sites that interact with DNMT3B', *Scientific Reports*, 12(1), p. 20839. Available at: https://doi.org/10.1038/s41598-022-24186-6.

Lomperta, K. *et al.* (2020) 'Insulin receptor substrate 1 may play divergent roles in human colorectal cancer development and progression', *World Journal of Gastroenterology*, 26(28), pp. 4140–4150. Available at: https://doi.org/10.3748/wjg.v26.i28.4140.

Louis, D.N. *et al.* (2014) 'International Society of Neuropathology-Haarlem Consensus Guidelines for Nervous System Tumor Classification and Grading', *Brain Pathology*, 24(5), pp. 429–435. Available at: https://doi.org/10.1111/bpa.12171.

Louis, D.N. *et al.* (2016) 'The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary', *Acta Neuropathologica*, 131(6), pp. 803–820. Available at: https://doi.org/10.1007/s00401-016-1545-1.

Louis, D.N. *et al.* (2021) 'The 2021 WHO Classification of Tumors of the Central Nervous System: a summary', *Neuro-Oncology*, 23(8), pp. 1231–1251. Available at: https://doi.org/10.1093/neuonc/noab106.

Love, M.I., Huber, W. and Anders, S. (2014) 'Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2', *Genome Biology*, 15(12), p. 550. Available at: https://doi.org/10.1186/s13059-014-0550-8.

Lu, Y. *et al.* (2017) 'lncRNA MIR100HG-derived miR-100 and miR-125b mediate cetuximab resistance via Wnt/β-catenin signaling', *Nature Medicine*, 23(11), pp. 1331–1341. Available at: https://doi.org/10.1038/nm.4424.

Lun, A.T.L. and Smyth, G.K. (2016) 'csaw: a Bioconductor package for differential binding analysis of ChIP-seq data using sliding windows', *Nucleic Acids Research*, 44(5), pp. e45–e45. Available at: https://doi.org/10.1093/nar/gkv1191.

Luo, X. *et al.* (2021) '3D Genome of macaque fetal brain reveals evolutionary innovations during primate corticogenesis', *Cell*, 184(3), pp. 723-740.e21. Available at: https://doi.org/10.1016/j.cell.2021.01.001.

Luo, Y. *et al.* (2020) 'New developments on the Encyclopedia of DNA Elements (ENCODE) data portal', *Nucleic Acids Research*, 48(D1), pp. D882–D889. Available at: https://doi.org/10.1093/nar/gkz1062.

Luzzati, F. (2015) 'A hypothesis for the evolution of the upper layers of the neocortex through co-option of the olfactory cortex developmental program', *Frontiers in Neuroscience*, 9, p. 162. Available at: https://doi.org/10.3389/fnins.2015.00162.

Mackay, A. *et al.* (2017) 'Integrated Molecular Meta-Analysis of 1,000 Pediatric High-Grade and Diffuse Intrinsic Pontine Glioma', *Cancer Cell*, 32(4), pp. 520-537.e5. Available at: https://doi.org/10.1016/j.ccell.2017.08.017.

Macosko, E.Z. *et al.* (2015) 'Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets', *Cell*, 161(5), pp. 1202–1214. Available at: https://doi.org/10.1016/j.cell.2015.05.002.

Maeda, Y. *et al.* (2023) 'Differential cell-cycle control by oscillatory versus sustained Hes1 expression via p21', *Cell Reports*, 42(5), p. 112520. Available at: https://doi.org/10.1016/j.celrep.2023.112520.

Malatesta, P. *et al.* (2003) 'Neuronal or glial progeny: regional differences in radial glia fate', *Neuron*, 37(5), pp. 751–764. Available at: https://doi.org/10.1016/s0896-6273(03)00116-8.

Malatesta, P., Hartfuss, E. and Götz, M. (2000) 'Isolation of radial glial cells by fluorescent-activated cell sorting reveals a neuronal lineage', *Development (Cambridge, England)*, 127(24), pp. 5253–5263. Available at: https://doi.org/10.1242/dev.127.24.5253.

Mancarella, C., Morrione, A. and Scotlandi, K. (2021) 'Novel Regulators of the IGF System in Cancer', *Biomolecules*, 11(2), p. 273. Available at: https://doi.org/10.3390/biom11020273.

Maness, P.F. and Schachner, M. (2007) 'Neural recognition molecules of the immunoglobulin superfamily: signaling transducers of axon guidance and neuronal migration', *Nature Neuroscience*, 10(1), pp. 19–26. Available at: https://doi.org/10.1038/nn1827.

Mani, S. *et al.* (2001) 'Failure to express GAP-43 during neurogenesis affects cell cycle regulation and differentiation of neural precursors and stimulates apoptosis of neurons', *Molecular and Cellular Neurosciences*, 17(1), pp. 54–66. Available at: https://doi.org/10.1006/mcne.2000.0931.

Manning, C.S. *et al.* (2019) 'Quantitative single-cell live imaging links HES5 dynamics with cell-state and fate in murine neurogenesis', *Nature Communications*, 10(1), p. 2835. Available at: https://doi.org/10.1038/s41467-019-10734-8.

Manuel, M.N. *et al.* (2015) 'Regulation of cerebral cortical neurogenesis by the Pax6 transcription factor', *Frontiers in Cellular Neuroscience*, 9, p. 70. Available at: https://doi.org/10.3389/fncel.2015.00070.

Margaria, J.P. *et al.* (2019) 'Class II PI3Ks at the Intersection between Signal Transduction and Membrane Trafficking', *Biomolecules*, 9(3), p. 104. Available at: https://doi.org/10.3390/biom9030104.

Martínez-Cerdeño, V. *et al.* (2016) 'Evolutionary origin of Tbr2-expressing precursor cells and the subventricular zone in the developing cortex', *The Journal of Comparative Neurology*, 524(3), pp. 433–447. Available at: https://doi.org/10.1002/cne.23879.

Martínez-Cerdeño, V. *et al.* (2018) 'Update on forebrain evolution: From neurogenesis to thermogenesis', *Seminars in Cell & Developmental Biology*, 76, pp. 15–22. Available at: https://doi.org/10.1016/j.semcdb.2017.09.034.

Martínez-Cerdeño, V., Noctor, S.C. and Kriegstein, A.R. (2006) 'The role of intermediate progenitor cells in the evolutionary expansion of the cerebral cortex', *Cerebral Cortex (New York, N.Y.: 1991)*, 16 Suppl 1, pp. i152-161. Available at: https://doi.org/10.1093/cercor/bhk017.

Martínez-Martínez, M.Á. *et al.* (2016) 'A restricted period for formation of outer subventricular zone defined by Cdh1 and Trnp1 levels', *Nature Communications*, 7(1), p. 11812. Available at: https://doi.org/10.1038/ncomms11812.

Martins, M. *et al.* (2021) 'A eutherian-specific microRNA controls the translation of Satb2 in a model of cortical differentiation', *Stem Cell Reports*, 16(6), pp. 1496–1509. Available at: https://doi.org/10.1016/j.stemcr.2021.04.020.

Matsumoto, N. *et al.* (2020) 'A discrete subtype of neural progenitor crucial for cortical folding in the gyrencephalic mammalian brain', *eLife*. Edited by F. Guillemot et al., 9, p. e54873. Available at: https://doi.org/10.7554/eLife.54873.

Mayo, S. *et al.* (2023) 'N-Type Ca Channel in Epileptic Syndromes and Epilepsy: A Systematic Review of Its Genetic Variants', *International Journal of Molecular Sciences*, 24(7), p. 6100. Available at: https://doi.org/10.3390/ijms24076100.

Mckenzie, M. and Ryan, M.T. (2010) 'Assembly factors of human mitochondrial complex I and their defects in disease', *IUBMB Life*, 62(7), pp. 497–502. Available at: https://doi.org/10.1002/iub.335.

Meyer, M.M. (2017) 'The role of mRNA structure in bacterial translational regulation', *Wiley interdisciplinary reviews. RNA*, 8(1). Available at: https://doi.org/10.1002/wrna.1370.

Mi, D. *et al.* (2013) 'Pax6 exerts regional control of cortical progenitor proliferation via direct repression of Cdk6 and hypophosphorylation of pRb', *Neuron*, 78(2), pp. 269–284. Available at: https://doi.org/10.1016/j.neuron.2013.02.012.

Miao, N. *et al.* (2018) 'Opposite Roles of Wnt7a and Sfrp1 in Modulating Proper Development of Neural Progenitors in the Mouse Cerebral Cortex', *Frontiers in Molecular Neuroscience*, 11, p. 247. Available at: https://doi.org/10.3389/fnmol.2018.00247.

Mihalas, A.B. *et al.* (2016) 'Intermediate Progenitor Cohorts Differentially Generate Cortical Layers and Require Tbr2 for Timely Acquisition of Neuronal Subtype Identity', *Cell Reports*, 16(1), pp. 92–105. Available at: https://doi.org/10.1016/j.celrep.2016.05.072.

Mihalas, A.B. and Hevner, R.F. (2018) 'Clonal analysis reveals laminar fate multipotency and daughter cell apoptosis of mouse cortical intermediate progenitors', *Development (Cambridge, England)*, 145(17). Available at: https://doi.org/10.1242/dev.164335.

Milan, M. *et al.* (2020) 'Pancreatic Cancer Cells Require the Transcription Factor MYRF to Maintain ER Homeostasis', *Developmental Cell*, 55(4), pp. 398-412.e7. Available at: https://doi.org/10.1016/j.devcel.2020.09.011.

Miškić, T. *et al.* (2021) 'Adult Upper Cortical Layer Specific Transcription Factor CUX2 Is Expressed in Transient Subplate and Marginal Zone Neurons of the Developing Human Brain', *Cells*, 10(2), p. 415. Available at: https://doi.org/10.3390/cells10020415.

Mitsis, T. *et al.* (2020) 'Transcription factors and evolution: An integral part of gene expression (Review)', *World Academy of Sciences Journal*, 2(1), pp. 3–8. Available at: https://doi.org/10.3892/wasj.2020.32.

Mitsogiannis, M.D. *et al.* (2021) 'Subtle Roles of Down Syndrome Cell Adhesion Molecules in Embryonic Forebrain Development and Neuronal Migration', *Frontiers in Cell and Developmental Biology*, 8. Available at: https://doi.org/10.3389/fcell.2020.624181.

Miyata, T. *et al.* (2001) 'Asymmetric inheritance of radial glial fibers by cortical neurons', *Neuron*, 31(5), pp. 727–741. Available at: https://doi.org/10.1016/s0896-6273(01)00420-2.

Miyata, T. *et al.* (2004) 'Asymmetric production of surface-dividing and non-surface-dividing cortical progenitor cells', *Development (Cambridge, England)*, 131(13), pp. 3133–3145. Available at: https://doi.org/10.1242/dev.01173.

Miyata, T. *et al.* (2014) 'Interkinetic nuclear migration generates and opposes ventricular-zone crowding: insight into tissue mechanics', *Frontiers in Cellular Neuroscience*, 8, p. 473. Available at: https://doi.org/10.3389/fncel.2014.00473.

Miyoshi, K., Miyoshi, T. and Siomi, H. (2010) 'Many ways to generate microRNA-like small RNAs: non-canonical pathways for microRNA production', *Molecular genetics and genomics: MGG*, 284(2), pp. 95–103. Available at: https://doi.org/10.1007/s00438-010-0556-1.

Moini, J., Gutierrez, A. and Avgeropoulos, N. (2023) 'Chapter 9 - Function and dysfunction of the cerebral lobes', in J. Moini, A. Gutierrez, and N. Avgeropoulos (eds) *Clinical Neuroepidemiology of Acute and Chronic Disorders*. Academic Press, pp. 133–152. Available at: https://doi.org/10.1016/B978-0-323-95901-8.00012-2.

Molenaar, J.J. *et al.* (2012) 'LIN28B induces neuroblastoma and enhances MYCN levels via let-7 suppression', *Nature Genetics*, 44(11), pp. 1199–1206. Available at: https://doi.org/10.1038/ng.2436.

Molnár, Z. (2011) 'Evolution of cerebral cortical development', *Brain, Behavior and Evolution*, 78(1), pp. 94–107. Available at: https://doi.org/10.1159/000327325.

Montiel, J.F. *et al.* (2016) 'From sauropsids to mammals and back: New approaches to comparative cortical development', *Journal of Comparative Neurology*, 524(3), pp. 630–645. Available at: https://doi.org/10.1002/cne.23871.

Moore, J.E. *et al.* (2020) 'Expanded encyclopaedias of DNA elements in the human and mouse genomes', *Nature*, 583(7818), pp. 699–710. Available at: https://doi.org/10.1038/s41586-020-2493-4.

Moore, L.D., Le, T. and Fan, G. (2013) 'DNA methylation and its basic function', *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, 38(1), pp. 23–38. Available at: https://doi.org/10.1038/npp.2012.112.

Mootha, V.K. *et al.* (2003) 'PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes', *Nature Genetics*, 34(3), pp. 267–273. Available at: https://doi.org/10.1038/ng1180.

Mukhtar, T. *et al.* (2020) 'Tead transcription factors differentially regulate cortical development', *Scientific Reports*, 10, p. 4625. Available at: https://doi.org/10.1038/s41598-020-61490-5.

Mukhtar, T. *et al.* (2022) 'Temporal and sequential transcriptional dynamics define lineage shifts in corticogenesis', *The EMBO Journal*, 41(24), p. e111132. Available at: https://doi.org/10.15252/embj.2022111132.

Mukhtar, T. and Taylor, V. (2023) 'Dynamic transcriptional programs define distinct mammalian cortical lineages', *Neural Regeneration Research*, 19(2), pp. 387–389. Available at: https://doi.org/10.4103/1673-5374.377589.

Muskovic, W. and Powell, J.E. (2021) 'DropletQC: improved identification of empty droplets and damaged cells in single-cell RNA-seq data', *Genome Biology*, 22(1), p. 329. Available at: https://doi.org/10.1186/s13059-021-02547-0.

Naftelberg, S. *et al.* (2015) 'Regulation of alternative splicing through coupling with transcription and chromatin structure', *Annual Review of Biochemistry*, 84, pp. 165–198. Available at: https://doi.org/10.1146/annurev-biochem-060614-034242.

Nagaraja, S. *et al.* (2019) 'Histone Variant and Cell Context Determine H3K27M Reprogramming of the Enhancer Landscape and Oncogenic State', *Molecular Cell*, 76(6), pp. 965-980.e12. Available at: https://doi.org/10.1016/j.molcel.2019.08.030.

Namba, T. *et al.* (2021) 'Metabolic Regulation of Neocortical Expansion in Development and Evolution', *Neuron*, 109(3), pp. 408–419. Available at: https://doi.org/10.1016/j.neuron.2020.11.014.

Nano, P.R. *et al.* (2023) 'A Meta-Atlas of the Developing Human Cortex Identifies Modules Driving Cell Subtype Specification', *bioRxiv*, p. 2023.09.12.557406. Available at: https://doi.org/10.1101/2023.09.12.557406.

Nesvick, C.L. *et al.* (2020) 'Atypical Teratoid Rhabdoid Tumor: Molecular Insights and Translation to Novel Therapeutics', *Journal of neuro-oncology*, 150(1), pp. 47–56. Available at: https://doi.org/10.1007/s11060-020-03639-w.

Neumann, J.E. *et al.* (2017) 'A mouse model for embryonal tumors with multilayered rosettes uncovers the therapeutic potential of Sonic-hedgehog inhibitors', *Nature Medicine*, 23(10), pp. 1191–1202. Available at: https://doi.org/10.1038/nm.4402.

Nguyen, L.H. *et al.* (2014) 'Lin28b is sufficient to drive liver cancer and necessary for its maintenance in murine models', *Cancer Cell*, 26(2), pp. 248–261. Available at: https://doi.org/10.1016/j.ccr.2014.06.018.

Nie, K. *et al.* (2020) 'COX6B2 drives metabolic reprogramming toward oxidative phosphorylation to promote metastasis in pancreatic ductal cancer cells', *Oncogenesis*, 9(5), pp. 1–13. Available at: https://doi.org/10.1038/s41389-020-0231-2.

Nishio, Y. *et al.* (2023) 'Gain-of-function MYCN causes a megalencephaly-polydactyly syndrome manifesting mirror phenotypes of Feingold syndrome', *HGG advances*, 4(4), p. 100238. Available at: https://doi.org/10.1016/j.xhgg.2023.100238.

Noack, F. *et al.* (2022) 'Multimodal profiling of the transcriptional regulatory landscape of the developing mouse cortex identifies Neurog2 as a key epigenome remodeler', *Nature Neuroscience*, 25(2), pp. 154–167. Available at: https://doi.org/10.1038/s41593-021-01002-4.

Noack, F. and Calegari, F. (2018) 'Epitranscriptomics: A New Regulatory Mechanism of Brain Development and Function', *Frontiers in Neuroscience*, 12. Available at: https://doi.org/10.3389/fnins.2018.00085.

Noctor, S.C. *et al.* (2001) 'Neurons derived from radial glial cells establish radial units in neocortex', *Nature*, 409(6821), pp. 714–720. Available at: https://doi.org/10.1038/35055553.

Noctor, S.C. *et al.* (2004) 'Cortical neurons arise in symmetric and asymmetric division zones and migrate through specific phases', *Nature Neuroscience*, 7(2), pp. 136–144. Available at: https://doi.org/10.1038/nn1172.

Noctor, S.C., Martínez-Cerdeño, V. and Kriegstein, A.R. (2008) 'Distinct behaviors of neural stem and progenitor cells underlie cortical neurogenesis', *The Journal of Comparative Neurology*, 508(1), pp. 28–44. Available at: https://doi.org/10.1002/cne.21669.

Nomura, T. *et al.* (2013) 'Reptiles: A New Model for Brain Evo-Devo Research', *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, 320(2), pp. 57–73. Available at: https://doi.org/10.1002/jez.b.22484.

Nomura, T. *et al.* (2016) 'The evolution of basal progenitors in the developing non-mammalian brain', *Development (Cambridge, England)*, 143(1), pp. 66–74. Available at: https://doi.org/10.1242/dev.127100.

Nomura, T. *et al.* (2018) 'Species-Specific Mechanisms of Neuron Subtype Specification Reveal Evolutionary Plasticity of Amniote Brain Development', *Cell Reports*, 22(12), pp. 3142–3151. Available at: https://doi.org/10.1016/j.celrep.2018.02.086.

Nomura, T., Gotoh, H. and Ono, K. (2013) 'Changes in the regulation of cortical neurogenesis contribute to encephalization during amniote brain evolution', *Nature Communications*, 4(1), p. 2206. Available at: https://doi.org/10.1038/ncomms3206.

Nonaka-Kinoshita, M. *et al.* (2013) 'Regulation of cerebral cortex size and folding by expansion of basal progenitors', *The EMBO Journal*, 32(13), pp. 1817–1828. Available at: https://doi.org/10.1038/emboj.2013.96.

Northcott, P.A. *et al.* (2017) 'The whole-genome landscape of medulloblastoma subtypes', *Nature*, 547(7663), pp. 311–317. Available at: https://doi.org/10.1038/nature22973.

Ochi, S. *et al.* (2020) 'Oscillatory expression of Hes1 regulates cell proliferation and neuronal differentiation in the embryonic brain', *Development (Cambridge, England)*, 147(4), p. dev182204. Available at: https://doi.org/10.1242/dev.182204.

Ochi, S. *et al.* (2022) 'Thirty Years' History since the Discovery of Pax6: From Central Nervous System Development to Neurodevelopmental Disorders', *International Journal of Molecular Sciences*, 23(11), p. 6115. Available at: https://doi.org/10.3390/ijms23116115.

Okano, M. *et al.* (1999) 'DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development', *Cell*, 99(3), pp. 247–257. Available at: https://doi.org/10.1016/s0092-8674(00)81656-6.

Ostrom, Q.T. *et al.* (2015) 'Alex's Lemonade Stand Foundation Infant and Childhood Primary Brain and Central Nervous System Tumors Diagnosed in the United States in 2007–2011', *Neuro-Oncology*, 16(Suppl 10), pp. x1–x36. Available at: https://doi.org/10.1093/neuonc/nou327.

Otte, J. *et al.* (2021) 'MYCN Function in Neuroblastoma Development', *Frontiers in Oncology*, 10. Available at: https://doi.org/10.3389/fonc.2020.624079.

Ouyang, J.F. *et al.* (2021) 'ShinyCell: simple and sharable visualization of single-cell gene expression data', *Bioinformatics*, 37(19), pp. 3374–3376. Available at: https://doi.org/10.1093/bioinformatics/btab209.

Pagin, M. *et al.* (2021) 'Sox2 Controls Neural Stem Cell Self-Renewal Through a Fos-Centered Gene Regulatory Network', *Stem Cells*, 39(8), pp. 1107–1119. Available at: https://doi.org/10.1002/stem.3373.

Papachristou, P. *et al.* (2014) 'Transgenic increase of Wnt7b in neural progenitor cells decreases expression of T-domain transcription factors and impairs neuronal differentiation', *Brain Research*, 1576, pp. 27–34. Available at: https://doi.org/10.1016/j.brainres.2014.06.015.

Park, S.H.E., Kulkarni, A. and Konopka, G. (2023) 'FOXP1 orchestrates neurogenesis in human cortical basal radial glial cells', *PLOS Biology*, 21(8), p. e3001852. Available at: https://doi.org/10.1371/journal.pbio.3001852.

Pearson, C.A. *et al.* (2020) 'Foxp1 Regulates Neural Stem Cell Self-Renewal and Bias Toward Deep Layer Cortical Fates', *Cell Reports*, 30(6), pp. 1964-1981.e3. Available at: https://doi.org/10.1016/j.celrep.2020.01.034.

Phillips, N.E. *et al.* (2016) 'Stochasticity in the miR-9/Hes1 oscillatory network can account for clonal heterogeneity in the timing of differentiation', *eLife*. Edited by A. Sánchez Alvarado, 5, p. e16118. Available at: https://doi.org/10.7554/eLife.16118.

Phipson, B. *et al.* (2022) 'propeller: testing for differences in cell type proportions in single cell data', *Bioinformatics*, 38(20), pp. 4720–4726. Available at: https://doi.org/10.1093/bioinformatics/btac582.

Pilz, G.-A. *et al.* (2013) 'Amplification of progenitors in the mammalian telencephalon includes a new radial glial cell type', *Nature Communications*, 4(1), p. 2125. Available at: https://doi.org/10.1038/ncomms3125.

Pinto, L. *et al.* (2008) 'Prospective isolation of functionally distinct radial glial subtypes—Lineage and transcriptome analysis', *Molecular and Cellular Neuroscience*, 38(1), pp. 15–42. Available at: https://doi.org/10.1016/j.mcn.2008.01.012.

Platzer, K. *et al.* (2017) 'GRIN2B encephalopathy: novel findings on phenotype, variant clustering, functional consequences and treatment aspects', *Journal of medical genetics*, 54(7), pp. 460–470. Available at: https://doi.org/10.1136/jmedgenet-2016-104509.

Pollen, A.A. *et al.* (2015) 'Molecular identity of human outer radial glia during cortical development', *Cell*, 163(1), pp. 55–67. Available at: https://doi.org/10.1016/j.cell.2015.09.004.

Prasad, R. *et al.* (2021) 'Hypermethylation of Mest promoter causes aberrant Wnt signaling in patients with Alzheimer's disease', *Scientific Reports*, 11(1), p. 20075. Available at: https://doi.org/10.1038/s41598-021-99562-9.

Pratt, H.E. *et al.* (2022) 'Factorbook: an updated catalog of transcription factor motifs and candidate regulatory motif sites', *Nucleic Acids Research*, 50(D1), pp. D141–D149. Available at: https://doi.org/10.1093/nar/gkab1039.

Prieto-Colomina, A. *et al.* (2021) 'MiRNAs in early brain development and pediatric cancer', *BioEssays*, 43(7), p. 2100073. Available at: https://doi.org/10.1002/bies.202100073.

Prüfer, K. *et al.* (2014) 'The complete genome sequence of a Neanderthal from the Altai Mountains', *Nature*, 505(7481), pp. 43–49. Available at: https://doi.org/10.1038/nature12886.

Pruunsild, P. and Timmusk, T. (2005) 'Structure, alternative splicing, and expression of the human and mouse KCNIP gene family', *Genomics*, 86(5), pp. 581–593. Available at: https://doi.org/10.1016/j.ygeno.2005.07.001.

Puelles, L. *et al.* (2000) 'Pallial and subpallial derivatives in the embryonic chick and mouse telencephalon, traced by the expression of the genes Dlx-2, Emx-1, Nkx-2.1, Pax-6, and Tbr-1', *The Journal of Comparative Neurology*, 424(3), pp. 409–438. Available at: https://doi.org/10.1002/1096-9861(20000828)424:3<409::aid-cne3>3.0.co;2-7.

Rakic, P. (1972) 'Mode of cell migration to the superficial layers of fetal monkey neocortex', *The Journal of Comparative Neurology*, 145(1), pp. 61–83. Available at: https://doi.org/10.1002/cne.901450105.

Rakic, P. (2009) 'Evolution of the neocortex: a perspective from developmental biology', *Nature Reviews Neuroscience*, 10(10), pp. 724–735. Available at: https://doi.org/10.1038/nrn2719.

Rao, S. *et al.* (2017) 'LIN28A, a sensitive immunohistochemical marker for Embryonal Tumor with Multilayered Rosettes (ETMR), is also positive in a subset of Atypical Teratoid/Rhabdoid Tumor (AT/RT)', *Child's Nervous System*, 33(11), pp. 1953–1959. Available at: https://doi.org/10.1007/s00381-017-3551-6.

Rao, S. *et al.* (2021) 'Central nervous system high grade neuroepithelial tumor with BCOR immunopositivity: Is there a molecular heterogeneity?', *Brain Tumor Pathology*, 38(1), pp. 41–49. Available at: https://doi.org/10.1007/s10014-020-00381-z.

Rauch, T.A. *et al.* (2009) 'A human B cell methylome at 100−base pair resolution', *Proceedings of the National Academy of Sciences*, 106(3), pp. 671–678. Available at: https://doi.org/10.1073/pnas.0812399106.

Rauluseviciute, I. *et al.* (2024) 'JASPAR 2024: 20th anniversary of the open-access database of transcription factor binding profiles', *Nucleic Acids Research*, 52(D1), pp. D174–D182. Available at: https://doi.org/10.1093/nar/gkad1059.

Reddy, A.T. *et al.* (2020) 'Efficacy of High-Dose Chemotherapy and Three-Dimensional Conformal Radiation for Atypical Teratoid/Rhabdoid Tumor: A Report From the Children's Oncology Group Trial ACNS0333', *Journal of Clinical Oncology*, 38(11), pp. 1175–1185. Available at: https://doi.org/10.1200/JCO.19.01776.

Reillo, I. *et al.* (2011) 'A role for intermediate radial glia in the tangential expansion of the mammalian cerebral cortex', *Cerebral Cortex (New York, N.Y.: 1991)*, 21(7), pp. 1674–1694. Available at: https://doi.org/10.1093/cercor/bhq238.

Robertson, K.D. (2001) 'DNA methylation, methyltransferases, and cancer', *Oncogene*, 20(24), pp. 3139–3155. Available at: https://doi.org/10.1038/sj.onc.1204341.

Robinson, M.D., McCarthy, D.J. and Smyth, G.K. (2010) 'edgeR: a Bioconductor package for differential expression analysis of digital gene expression data', *Bioinformatics (Oxford, England)*, 26(1), pp. 139–140. Available at: https://doi.org/10.1093/bioinformatics/btp616.

Ronjat, M. *et al.* (2013) 'Nuclear life of the voltage-gated Cacnb4 subunit and its role in gene transcription regulation', *Channels (Austin, Tex.)*, 7(2), pp. 119–125. Available at: https://doi.org/10.4161/chan.23895.

Ruan, X. *et al.* (2021) 'Progenitor cell diversity in the developing mouse neocortex', *Proceedings of the National Academy of Sciences*, 118(10), p. e2018866118. Available at: https://doi.org/10.1073/pnas.2018866118.

Ryall, S., Tabori, U. and Hawkins, C. (2020) 'Pediatric low-grade glioma in the era of molecular diagnostics', *Acta Neuropathologica Communications*, 8, p. 30. Available at: https://doi.org/10.1186/s40478-020-00902-z.

Saha, S. *et al.* (2022) 'Tipping the balance toward stemness in trophoblast: Metabolic programming by Cox6B2', *The FASEB Journal*, 36(11), p. e22600. Available at: https://doi.org/10.1096/fj.202200703RR.

Sakaki-Yumoto, M. *et al.* (2006) 'The murine homolog of SALL4, a causative gene in Okihiro syndrome, is essential for embryonic stem cell proliferation, and cooperates with Sall1 in anorectal, heart, brain and kidney development', *Development (Cambridge, England)*, 133(15), pp. 3005–3013. Available at: https://doi.org/10.1242/dev.02457.

Salamon, I. and Rasin, M.-R. (2021) 'Evolution of the Neocortex Through RNA-Binding Proteins and Post-transcriptional Regulation', *Frontiers in Neuroscience*, 15, p. 803107. Available at: https://doi.org/10.3389/fnins.2021.803107.

Sansom, S.N. *et al.* (2009) 'The Level of the Transcription Factor Pax6 Is Essential for Controlling the Balance between Neural Stem Cell Self-Renewal and Neurogenesis', *PLOS Genetics*, 5(6), p. e1000511. Available at: https://doi.org/10.1371/journal.pgen.1000511.

Sato, K., Akiyama, M. and Sakakibara, Y. (2021) 'RNA secondary structure prediction using deep learning with thermodynamic integration', *Nature Communications*, 12, p. 941. Available at: https://doi.org/10.1038/s41467-021-21194-4.

Scardigli, R. *et al.* (2003) 'Direct and concentration-dependent regulation of the proneural gene Neurogenin2 by Pax6', *Development (Cambridge, England)*, 130(14), pp. 3269–3281. Available at: https://doi.org/10.1242/dev.00539.

Schmahl, W. *et al.* (1993) 'Defects of neuronal migration and the pathogenesis of cortical malformations are associated with Small eye (Sey) in the mouse, a point mutation at the Pax-6-locus', *Acta Neuropathologica*, 86(2), pp. 126–135. Available at: https://doi.org/10.1007/BF00334879.

Schöler, J. *et al.* (2015) 'The Intracellular Domain of Teneurin-1 Induces the Activity of Microphthalmia-associated Transcription Factor (MITF) by Binding to Transcriptional Repressor HINT1', *Journal of Biological Chemistry*, 290(13), pp. 8154–8165. Available at: https://doi.org/10.1074/jbc.M114.615922.

Scopa, C. *et al.* (2023) 'JUN upregulation drives aberrant transposable element mobilization, associated innate immune response, and impaired neurogenesis in Alzheimer's disease', *Nature Communications*, 14(1), p. 8021. Available at: https://doi.org/10.1038/s41467-023-43728-8.

Seki, M. *et al.* (2010) 'Human ROBO1 is cleaved by metalloproteinases and γ-secretase and migrates to the nucleus in cancer cells', *FEBS Letters*, 584(13), pp. 2909–2915. Available at: https://doi.org/10.1016/j.febslet.2010.05.009.

Seki, M. *et al.* (2014) 'Biallelic DICER1 Mutations in Sporadic Pleuropulmonary Blastoma', *Cancer Research*, 74(10), pp. 2742–2749. Available at: https://doi.org/10.1158/0008-5472.CAN-13-2470.

Sena, R.M. *et al.* (2021) 'The RNA-Binding Protein HuD Regulates Alternative Splicing and Alternative Polyadenylation in the Mouse Neocortex', *Molecules (Basel, Switzerland)*, 26(10), p. 2836. Available at: https://doi.org/10.3390/molecules26102836.

Seo, P.J., Kim, S.-G. and Park, C.-M. (2008) 'Membrane-bound transcription factors in plants', *Trends in Plant Science*, 13(10), pp. 550–556. Available at: https://doi.org/10.1016/j.tplants.2008.06.008.

Serdar, L.D. *et al.* (2025) 'mRNA stability fine-tunes gene expression in the developing cortex to control neurogenesis', *PLoS biology*, 23(2), p. e3003031. Available at: https://doi.org/10.1371/journal.pbio.3003031.

Sessa, A. *et al.* (2008) 'Tbr2 directs conversion of radial glia into basal precursors and guides neuronal amplification by indirect neurogenesis in the developing neocortex', *Neuron*, 60(1), pp. 56–69. Available at: https://doi.org/10.1016/j.neuron.2008.09.028.

Sévenet, N. *et al.* (1999) 'Constitutional Mutations of the hSNF5/INI1 Gene Predispose to a Variety of Cancers', *American Journal of Human Genetics*, 65(5), pp. 1342–1348.

Shaulian, E. and Karin, M. (2001) 'AP-1 in cell proliferation and survival', *Oncogene*, 20(19), pp. 2390–2400. Available at: https://doi.org/10.1038/sj.onc.1204383.

Sherman, B.T. *et al.* (2022) 'DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update)', *Nucleic Acids Research*, 50(W1), pp. W216–W221. Available at: https://doi.org/10.1093/nar/gkac194.

Shi, W. *et al.* (2019) 'Unravel the molecular mechanism of XBP1 in regulating the biology of cancer cells', *Journal of Cancer*, 10(9), pp. 2035–2046. Available at: https://doi.org/10.7150/jca.29421.

Shibata, M., Gulden, F.O. and Sestan, N. (2015) 'From Trans to Cis: Transcriptional Regulatory Networks in Neocortical Development', *Trends in genetics : TIG*, 31(2), pp. 77–87. Available at: https://doi.org/10.1016/j.tig.2014.12.004.

Shimojo, H. and Kageyama, R. (2016) 'Oscillatory control of Delta-like1 in somitogenesis and neurogenesis: A unified model for different oscillatory dynamics', *Seminars in Cell & Developmental Biology*, 49, pp. 76–82. Available at: https://doi.org/10.1016/j.semcdb.2016.01.017.

Shimojo, H., Masaki, T. and Kageyama, R. (2024) 'The Neurog2-Tbr2 axis forms a continuous transition to the neurogenic gene expression state in neural stem cells', *Developmental Cell*, 59(15), pp. 1913-1923.e6. Available at: https://doi.org/10.1016/j.devcel.2024.04.019.

Shitamukai, A., Konno, D. and Matsuzaki, F. (2011) 'Oblique Radial Glial Divisions in the Developing Mouse Neocortex Induce Self-Renewing Progenitors outside the Germinal Zone That Resemble Primate Outer Subventricular Zone Progenitors', *The Journal of Neuroscience*, 31(10), pp. 3683–3695. Available at: https://doi.org/10.1523/JNEUROSCI.4773-10.2011.

Sidman, R.L. and Rakic, P. (1973) 'Neuronal migration, with special reference to developing human brain: a review', *Brain Research*, 62(1), pp. 1–35. Available at: https://doi.org/10.1016/0006-8993(73)90617-3.

Siegenthaler, J.A., Tremper-Wells, B.A. and Miller, M.W. (2008) 'Foxg1 Haploinsufficiency Reduces the Population of Cortical Intermediate Progenitor Cells: Effect of Increased p21 Expression', *Cerebral Cortex*, 18(8), pp. 1865–1875. Available at: https://doi.org/10.1093/cercor/bhm209.

Silver, D.L. (2016) 'Genomic divergence and brain evolution: How regulatory DNA influences development of the cerebral cortex', *BioEssays : news and reviews in molecular, cellular and developmental biology*, 38(2), pp. 162–171. Available at: https://doi.org/10.1002/bies.201500108.

Sin-Chan, P. *et al.* (2019) 'A *C19MC*-LIN28A-MYCN Oncogenic Circuit Driven by Hijacked Super-enhancers Is a Distinct Therapeutic Vulnerability in ETMRs: A Lethal Brain Tumor', *Cancer Cell*, 36(1), pp. 51-67.e7. Available at: https://doi.org/10.1016/j.ccell.2019.06.002.

Singh, A. *et al.* (2024) 'Gene regulatory landscape of cerebral cortex folding', *Science Advances*, 10(23), p. eadn1640. Available at: https://doi.org/10.1126/sciadv.adn1640.

Slowikowski, K. (2024) *ggrepel: Automatically Position Non-Overlapping Text Labels with 'ggplot2'*.

Sobhan, P.K. and Funa, K. (2017) 'TLX—Its Emerging Role for Neurogenesis in Health and Disease', *Molecular Neurobiology*, 54(1), pp. 272–280. Available at: https://doi.org/10.1007/s12035-015-9608-1.

Sokpor, G. *et al.* (2024) 'H3 Acetylation-Induced Basal Progenitor Generation and Neocortex Expansion Depends on the Transcription Factor Pax6', *Biology*, 13(2), p. 68. Available at: https://doi.org/10.3390/biology13020068.

Sousa, A.M.M. *et al.* (2017) 'Evolution of the Human Nervous System Function, Structure, and Development', *Cell*, 170(2), pp. 226–247. Available at: https://doi.org/10.1016/j.cell.2017.06.036.

Stancik, E.K. *et al.* (2010) 'Heterogeneity in ventricular zone neural precursors contributes to neuronal fate diversity in the postnatal neocortex', *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 30(20), pp. 7028–7036. Available at: https://doi.org/10.1523/JNEUROSCI.6131-09.2010.

Stauffer, W.T. *et al.* (2020) 'Sledgehammer to Scalpel: Broad Challenges to the Heart and Other Tissues Yield Specific Cellular Responses via Transcriptional Regulation of the ER-Stress Master Regulator ATF6α', *International Journal of Molecular Sciences*, 21(3), p. 1134. Available at: https://doi.org/10.3390/ijms21031134.

Stern, D.L. (2000) 'PERSPECTIVE: EVOLUTIONARY DEVELOPMENTAL BIOLOGY AND THE PROBLEM OF VARIATION', *Evolution*, 54(4), pp. 1079–1091. Available at: https://doi.org/10.1111/j.0014-3820.2000.tb00544.x.

Stillman, B. (2018) 'Histone Modifications: Insights into Their Influence on Gene Expression', *Cell*, 175(1), pp. 6–9. Available at: https://doi.org/10.1016/j.cell.2018.08.032.

Stocker, A.M. and Chenn, A. (2009) 'Focal Reduction of αE-catenin Causes Premature Differentiation and Reduction of β-catenin Signaling During Cortical Development', *Developmental biology*, 328(1), pp. 66–77. Available at: https://doi.org/10.1016/j.ydbio.2009.01.010.

Subramanian, A. *et al.* (2005) 'Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles', *Proceedings of the National Academy of Sciences*, 102(43), pp. 15545–15550. Available at: https://doi.org/10.1073/pnas.0506580102.

Subramanian, S. and Ahmad, T. (2024) 'Childhood Brain Tumors', in *StatPearls*. Treasure Island (FL): StatPearls Publishing. Available at: http://www.ncbi.nlm.nih.gov/books/NBK535415/ (Accessed: 6 December 2024).

Sudmant, P.H. *et al.* (2010) 'Diversity of Human Copy Number Variation and Multicopy Genes', *Science*, 330(6004), pp. 641–646. Available at: https://doi.org/10.1126/science.1197005.

Sugino, H. *et al.* (2023) 'High-grade neuroepithelial tumor with EP300::BCOR fusion and negative BCOR immunohistochemical expression: a case report', *Brain Tumor Pathology*, 40(2), pp. 133–141. Available at: https://doi.org/10.1007/s10014-023-00451-y.

Suzuki, I.K. *et al.* (2018) 'Human-Specific NOTCH2NL Genes Expand Cortical Neurogenesis through Delta/Notch Regulation', *Cell*, 173(6), pp. 1370-1384.e16. Available at: https://doi.org/10.1016/j.cell.2018.03.067.

Szklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R., Gable, Annika L, *et al.* (2023) 'The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest', *Nucleic Acids Research*, 51(D1), pp. D638–D646. Available at: https://doi.org/10.1093/nar/gkac1000.

Szklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R., Gable, Annika L., *et al.* (2023) 'The STRING database in 2023: protein-protein association networks and functional

enrichment analyses for any sequenced genome of interest', *Nucleic Acids Research*, 51(D1), pp. D638–D646. Available at: https://doi.org/10.1093/nar/gkac1000.

Takahashi, T., Nowakowski, R.S. and Caviness Jr., V.S. (1996) 'Interkinetic and Migratory Behavior of a Cohort of Neocortical Neurons Arising in the Early Embryonic Murine Cerebral Wall', *The Journal of Neuroscience*, 16(18), pp. 5762–5776. Available at: https://doi.org/10.1523/JNEUROSCI.16-18-05762.1996.

Takahashi, T., Nowakowski, R.S. and Caviness, V.S. (1993) 'Cell cycle parameters and patterns of nuclear movement in the neocortical proliferative zone of the fetal mouse', *Journal of Neuroscience*, 13(2), pp. 820–833. Available at: https://doi.org/10.1523/JNEUROSCI.13-02-00820.1993.

Takahashi, T., Nowakowski, R.S. and Caviness, V.S. (1995) 'The cell cycle of the pseudostratified ventricular epithelium of the embryonic murine cerebral wall', *Journal of Neuroscience*, 15(9), pp. 6046–6057. Available at: https://doi.org/10.1523/JNEUROSCI.15-09-06046.1995.

Takahashi, T., Nowakowski, R.S. and Caviness, V.S. (1996) 'The Leaving or Q Fraction of the Murine Cerebral Proliferative Epithelium: A General Model of Neocortical Neuronogenesis', *The Journal of Neuroscience*, 16(19), pp. 6183–6196. Available at: https://doi.org/10.1523/JNEUROSCI.16-19-06183.1996.

Taverna, E., Götz, M. and Huttner, W.B. (2014) 'The cell biology of neurogenesis: toward an understanding of the development and evolution of the neocortex', *Annual Review of Cell and Developmental Biology*, 30, pp. 465–502. Available at: https://doi.org/10.1146/annurev-cellbio-101011-155801.

The UniProt Consortium (2021) 'UniProt: the universal protein knowledgebase in 2021', *Nucleic Acids Research*, 49(D1), pp. D480–D489. Available at: https://doi.org/10.1093/nar/gkaa1100.

Thul, P.J. *et al.* (2017) 'A subcellular map of the human proteome', *Science (New York, N.Y.)*, 356(6340), p. eaal3321. Available at: https://doi.org/10.1126/science.aal3321.

Tirosh, I. *et al.* (2016) 'Single-cell RNA-seq supports a developmental hierarchy in human oligodendroglioma', *Nature*, 539(7628), pp. 309–313. Available at: https://doi.org/10.1038/nature20123.

Toda, T. *et al.* (2016) 'An essential role of SVZ progenitors in cortical folding in gyrencephalic mammals', *Scientific Reports*, 6(1), p. 29578. Available at: https://doi.org/10.1038/srep29578.

Tomasello, U. *et al.* (2022) 'miR-137 and miR-122, two outer subventricular zone non-coding RNAs, regulate basal progenitor expansion and neuronal differentiation', *Cell Reports*, 38(7), p. 110381. Available at: https://doi.org/10.1016/j.celrep.2022.110381.

Torchia, J. *et al.* (2015) 'Molecular subgroups of atypical teratoid rhabdoid tumours in children: an integrated genomic and clinicopathological analysis', *The Lancet. Oncology*, 16(5), pp. 569–582. Available at: https://doi.org/10.1016/S1470-2045(15)70114-2.

Torchia, J. *et al.* (2016) 'Integrated (epi)-Genomic Analyses Identify Subgroup-Specific Therapeutic Targets in CNS Rhabdoid Tumors', *Cancer cell*, 30(6), pp. 891–908. Available at: https://doi.org/10.1016/j.ccell.2016.11.003.

de la Torre-Ubieta, L. *et al.* (2018) 'The Dynamic Landscape of Open Chromatin during Human Cortical Neurogenesis', *Cell*, 172(1), pp. 289-304.e18. Available at: https://doi.org/10.1016/j.cell.2017.12.014.

Tosches, M.A. *et al.* (2018) 'Evolution of pallium, hippocampus, and cortical cell types revealed by single-cell transcriptomics in reptiles', *Science (New York, N.Y.)*, 360(6391), pp. 881–888. Available at: https://doi.org/10.1126/science.aar4237.

Tourigny, M.R. *et al.* (2002) 'CDK Inhibitor p18INK4c Is Required for the Generation of Functional Plasma Cells', *Immunity*, 17(2), pp. 179–189. Available at: https://doi.org/10.1016/S1074-7613(02)00364-3.

Trujillo, C.A. *et al.* (2021) 'Reintroduction of the archaic variant of NOVA1 in cortical organoids alters neurodevelopment', *Science*, 371(6530), p. eaax2537. Available at: https://doi.org/10.1126/science.aax2537.

Tsai, M.-H. and Lee, C.-K. (2018) 'STAT3 Cooperates With Phospholipid Scramblase 2 to Suppress Type I Interferon Response', *Frontiers in Immunology*, 9, p. 1886. Available at: https://doi.org/10.3389/fimmu.2018.01886.

Tsuboi, M. and Gotoh, Y. (2020) 'Endfoot regrowth for neural stem cell renewal', *Nature Cell Biology*, 22(1), pp. 3–5. Available at: https://doi.org/10.1038/s41556-019-0448-5.

Tsunekawa, Y. *et al.* (2012) 'Cyclin D2 in the basal process of neural progenitors is linked to non-equivalent cell fates', *The EMBO journal*, 31(8), pp. 1879–1892. Available at: https://doi.org/10.1038/emboj.2012.43.

Tuoc, T.C. *et al.* (2009) 'Selective Cortical Layering Abnormalities and Behavioral Deficits in Cortex-Specific Pax6 Knock-Out Mice', *Journal of Neuroscience*, 29(26), pp. 8335–8349. Available at: https://doi.org/10.1523/JNEUROSCI.5669-08.2009.

Uebbing, S. *et al.* (2021) 'Massively parallel discovery of human-specific substitutions that alter enhancer activity', *Proceedings of the National Academy of Sciences of the United States of America*, 118(2), p. e2007049118. Available at: https://doi.org/10.1073/pnas.2007049118.

Uemura, A. *et al.* (2009) 'Unconventional splicing of XBP1 mRNA occurs in the cytoplasm during the mammalian unfolded protein response', *Journal of Cell Science*, 122(16), pp. 2877–2886. Available at: https://doi.org/10.1242/jcs.040584.

Ule, J. and Blencowe, B.J. (2019) 'Alternative Splicing Regulatory Networks: Functions, Mechanisms, and Evolution', *Molecular Cell*, 76(2), pp. 329–345. Available at: https://doi.org/10.1016/j.molcel.2019.09.017.

Ung, D.C. *et al.* (2024) 'GRID1/GluD1 homozygous variants linked to intellectual disability and spastic paraplegia impair mGlu1/5 receptor signaling and excitatory synapses', *Molecular Psychiatry*, 29(4), pp. 1205–1215. Available at: https://doi.org/10.1038/s41380-024-02469-w.

Uziel, T. *et al.* (2006) 'The CDK Inhibitor p18Ink4c is a Tumor Suppressor in Medulloblastoma', *Cell Cycle*, 5(4), pp. 363–365. Available at: https://doi.org/10.4161/cc.5.4.2475.

Vaid, S. *et al.* (2018) 'A novel population of Hopx-dependent basal radial glial cells in the developing mouse neocortex', *Development (Cambridge, England)*, 145(20), p. dev169276. Available at: https://doi.org/10.1242/dev.169276.

Vasistha, N.A. *et al.* (2015) 'Cortical and Clonal Contribution of Tbr2 Expressing Progenitors in the Developing Mouse Brain', *Cerebral Cortex (New York, N.Y.: 1991)*, 25(10), pp. 3290–3302. Available at: https://doi.org/10.1093/cercor/bhu125.

Venteicher, A.S. *et al.* (2017) 'Decoupling genetics, lineages, and microenvironment in IDH-mutant gliomas by single-cell RNA-seq', *Science (New York, N.Y.)*, 355(6332), p. eaai8478. Available at: https://doi.org/10.1126/science.aai8478.

Versteege, I. *et al.* (1998) 'Truncating mutations of hSNF5/INI1 in aggressive paediatric cancer', *Nature*, 394(6689), pp. 203–206. Available at: https://doi.org/10.1038/28212.

Viaene, A.N. (2023) 'Pediatric brain tumors: A neuropathologist's approach to the integrated diagnosis', *Frontiers in Pediatrics*, 11, p. 1143363. Available at: https://doi.org/10.3389/fped.2023.1143363.

Visvanathan, A. *et al.* (2024) 'Early rhombic lip Protogenin+ve stem cells in a human-specific neurovascular niche initiate and maintain group 3 medulloblastoma', *Cell*, 187(17), pp. 4733-4750.e26. Available at: https://doi.org/10.1016/j.cell.2024.06.011.

Viswanathan, S.R., Daley, G.Q. and Gregory, R.I. (2008) 'Selective blockade of microRNA processing by Lin28', *Science (New York, N.Y.)*, 320(5872), pp. 97–100. Available at: https://doi.org/10.1126/science.1154040.

Vladoiu, M.C. *et al.* (2019) 'Childhood cerebellar tumours mirror conserved fetal transcriptional programs', *Nature*, 572(7767), pp. 67–73. Available at: https://doi.org/10.1038/s41586-019-1158-7.

Vogel, C. and Marcotte, E.M. (2012) 'Insights into the regulation of protein abundance from proteomic and transcriptomic analyses', *Nature Reviews. Genetics*, 13(4), pp. 227–232. Available at: https://doi.org/10.1038/nrg3185.

Vorontsov, I.E. *et al.* (2024) 'HOCOMOCO in 2024: a rebuild of the curated collection of binding models for human and mouse transcription factors', *Nucleic Acids Research*, 52(D1), pp. D154–D163. Available at: https://doi.org/10.1093/nar/gkad1077.

Wang, C. *et al.* (2011) 'Identification and characterization of neuroblasts in the subventricular zone and rostral migratory stream of the adult human brain', *Cell Research*, 21(11), pp. 1534–1550. Available at: https://doi.org/10.1038/cr.2011.83.

Wang, J. *et al.* (2021) 'Epigenomic landscape and 3D genome structure in pediatric high-grade glioma', *Science Advances*, 7(23), p. eabg4126. Available at: https://doi.org/10.1126/sciadv.abg4126.

Wang, Junbao *et al.* (2022) 'A Ctnnb1 enhancer regulates neocortical neurogenesis by controlling the abundance of intermediate progenitors', *Cell Discovery*, 8(1), pp. 1–17. Available at: https://doi.org/10.1038/s41421-022-00421-2.

Wang, L. *et al.* (2025) 'Molecular and cellular dynamics of the developing human neocortex', *Nature*, pp. 1–10. Available at: https://doi.org/10.1038/s41586-024-08351-7.

Wang, X. and Hayes, J.J. (2008) 'Acetylation mimics within individual core histone tail domains indicate distinct roles in regulating the stability of higher-order chromatin structure', *Molecular and Cellular Biology*, 28(1), pp. 227–236. Available at: https://doi.org/10.1128/MCB.01245-07.

Wang, Z., Gerstein, M. and Snyder, M. (2009) 'RNA-Seq: a revolutionary tool for transcriptomics', *Nature reviews. Genetics*, 10(1), pp. 57–63. Available at: https://doi.org/10.1038/nrg2484.

Waterhouse, A.M. *et al.* (2009) 'Jalview Version 2—a multiple sequence alignment editor and analysis workbench', *Bioinformatics*, 25(9), pp. 1189–1191. Available at: https://doi.org/10.1093/bioinformatics/btp033.

Wegiel, B. *et al.* (2008) 'Interleukin-6 activates PI3K/Akt pathway and regulates cyclin A1 to promote prostate cancer cell survival', *International Journal of Cancer*, 122(7), pp. 1521–1529. Available at: https://doi.org/10.1002/ijc.23261.

Weiss, L.A. and Nieto, M. (2019) 'The crux of *Cux* genes in neuronal function and plasticity', *Brain Research*, 1705, pp. 32–42. Available at: https://doi.org/10.1016/j.brainres.2018.02.044.

Wendt, K.S. *et al.* (2008) 'Cohesin mediates transcriptional insulation by CCCTC-binding factor', *Nature*, 451(7180), pp. 796–801. Available at: https://doi.org/10.1038/nature06634.

Westermark, U.K. *et al.* (2011) 'The MYCN oncogene and differentiation in neuroblastoma', *Seminars in Cancer Biology*, 21(4), pp. 256–266. Available at: https://doi.org/10.1016/j.semcancer.2011.08.001.

Wickham, H. (2016) *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. Available at: https://ggplot2.tidyverse.org.

Wolf, F.A. *et al.* (2019) 'PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells', *Genome Biology*, 20(1), p. 59. Available at: https://doi.org/10.1186/s13059-019-1663-x.

Wong, F.K. *et al.* (2015) 'Sustained Pax6 Expression Generates Primate-like Basal Radial Glia in Developing Mouse Neocortex', *PLOS Biology*, 13(8), p. e1002217. Available at: https://doi.org/10.1371/journal.pbio.1002217.

Wong, Y.-H. *et al.* (2010) 'Protogenin Defines a Transition Stage during Embryonic Neurogenesis and Prevents Precocious Neuronal Differentiation', *Journal of Neuroscience*, 30(12), pp. 4428–4439. Available at: https://doi.org/10.1523/JNEUROSCI.0473-10.2010.

Wu, H. *et al.* (2022) 'Single-Cell RNA Sequencing Unravels Upregulation of Immune Cell Crosstalk in Relapsed Pediatric Ependymoma', *Frontiers in Immunology*, 13, p. 903246. Available at: https://doi.org/10.3389/fimmu.2022.903246.

Wullimann, M.F., Puelles, L. and Wicht, H. (1999) 'Early postembryonic neural development in the zebrafish: a 3-D reconstruction of forebrain proliferation zones shows their relation to prosomeres', *European Journal of Morphology*, 37(2–3), pp. 117–121. Available at: https://doi.org/10.1076/ejom.37.2.117.4739.

Xiang, T. *et al.* (2021) 'The novel ZEB1-upregulated protein PRTG induced by Helicobacter pylori infection promotes gastric carcinogenesis through the cGMP/PKG signaling pathway', *Cell Death & Disease*, 12(2), p. 150. Available at: https://doi.org/10.1038/s41419-021-03440-1.

Xiong, X. *et al.* (2014) 'Ribosomal protein S27-like is a physiological regulator of p53 that suppresses genomic instability and tumorigenesis', *eLife*, 3, p. e02236. Available at: https://doi.org/10.7554/eLife.02236.

Xu, J. *et al.* (2023) 'Ccdc85c-Par3 condensates couple cell polarity with Notch to control neural progenitor proliferation', *Cell Reports*, 42(7), p. 112677. Available at: https://doi.org/10.1016/j.celrep.2023.112677.

Yamashita, T. *et al.* (2018) 'Diverse Long-Range Axonal Projections of Excitatory Layer 2/3 Neurons in Mouse Barrel Cortex', *Frontiers in Neuroanatomy*, 12, p. 33. Available at: https://doi.org/10.3389/fnana.2018.00033.

Yang, Y. *et al.* (2020) 'Intracellular Trafficking and Imaging Methods of Membrane-Bound Transcription Factors in Plants', *Critical Reviews in Plant Sciences*, 39(5), pp. 418–430. Available at: https://doi.org/10.1080/07352689.2020.1813922.

Yoon, K.-J. *et al.* (2017) 'Temporal Control of Mammalian Cortical Neurogenesis by m6A Methylation', *Cell*, 171(4), pp. 877-889.e17. Available at: https://doi.org/10.1016/j.cell.2017.09.003.

Yoon, K.-J. *et al.* (2018) 'Epigenetics and epitranscriptomics in temporal patterning of cortical neural progenitor competence', *The Journal of Cell Biology*, 217(6), pp. 1901–1914. Available at: https://doi.org/10.1083/jcb.201802117.

Yoshida, H. *et al.* (2001) 'XBP1 mRNA Is Induced by ATF6 and Spliced by IRE1 in Response to ER Stress to Produce a Highly Active Transcription Factor', *Cell*, 107(7), pp. 881–891. Available at: https://doi.org/10.1016/S0092-8674(01)00611-0.

Yost, C.S. *et al.* (2008) 'Knockout of the gene encoding the K2P channel KCNK7 does not alter volatile anesthetic sensitivity', *Behavioural brain research*, 193(2), pp. 192–196. Available at: https://doi.org/10.1016/j.bbr.2008.05.010.

Young, M.D. and Behjati, S. (2020) 'SoupX removes ambient RNA contamination from droplet-based single-cell RNA sequencing data', *GigaScience*, 9(12), p. giaa151. Available at: https://doi.org/10.1093/gigascience/giaa151.

Yu, G. *et al.* (2012) 'clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters', *OMICS : a Journal of Integrative Biology*, 16(5), pp. 284–287. Available at: https://doi.org/10.1089/omi.2011.0118.

Yu, G., Wang, L.-G. and He, Q.-Y. (2015) 'ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization', *Bioinformatics*, 31(14), pp. 2382–2383. Available at: https://doi.org/10.1093/bioinformatics/btv145.

Yuan, J. *et al.* (2023) 'SALL1 promotes proliferation and metastasis and activates phosphorylation of p65 and JUN in colorectal cancer cells', *Pathology, Research and Practice*, 250, p. 154827. Available at: https://doi.org/10.1016/j.prp.2023.154827.

Zega, K. *et al.* (2017) 'Dusp16 Deficiency Causes Congenital Obstructive Hydrocephalus and Brain Overgrowth by Expansion of the Neural Progenitor Pool', *Frontiers in Molecular Neuroscience*, 10. Available at: https://doi.org/10.3389/fnmol.2017.00372.

Zhang, M. *et al.* (2018) 'Insulin-like growth factor 1/insulin-like growth factor 1 receptor signaling protects against cell apoptosis through the PI3K/AKT pathway in glioblastoma cells', *Experimental and Therapeutic Medicine*, 16(2), pp. 1477–1482. Available at: https://doi.org/10.3892/etm.2018.6336.

Zhang, M. *et al.* (2020) 'Targeting the Wnt signaling pathway through R-spondin 3 identifies an anti-fibrosis treatment strategy for multiple organs', *PLOS ONE*, 15(3), p. e0229445. Available at: https://doi.org/10.1371/journal.pone.0229445.

Zhang, Y. *et al.* (2008) 'Model-based Analysis of ChIP-Seq (MACS)', *Genome Biology*, 9(9), p. R137. Available at: https://doi.org/10.1186/gb-2008-9-9-r137.

Zhang, Y.E. *et al.* (2011) 'Accelerated Recruitment of New Brain Development Genes into the Human Genome', *PLOS Biology*, 9(10), p. e1001179. Available at: https://doi.org/10.1371/journal.pbio.1001179.

Zhao, T. *et al.* (2008) 'Genetic mapping of Foxb1-cell lineage shows migration from caudal diencephalon to telencephalon and lateral hypothalamus', *European Journal of Neuroscience*, 28(10), pp. 1941–1955. Available at: https://doi.org/10.1111/j.1460-9568.2008.06503.x.

Zheng, G.X.Y. *et al.* (2017) 'Massively parallel digital transcriptional profiling of single cells', *Nature Communications*, 8(1), p. 14049. Available at: https://doi.org/10.1038/ncomms14049.

Zheng, M. and Wang, P. (2021) 'Role of insulin receptor substance-1 modulating PI3K/Akt insulin signaling pathway in Alzheimer's disease', *3 Biotech*, 11(4), p. 179. Available at: https://doi.org/10.1007/s13205-021-02738-3.

Zhou, Q. *et al.* (2005) 'Phospholipid Scramblase 1 Binds to the Promoter Region of the Inositol 1,4,5-Triphosphate Receptor Type 1 Gene to Enhance Its Expression *', *Journal of Biological Chemistry*, 280(41), pp. 35062–35068. Available at: https://doi.org/10.1074/jbc.M504821200.

Zhu, Y. *et al.* (2024) 'HDAC1 and HDAC2 orchestrate Wnt signaling to regulate neural progenitor transition during brain development', *iScience*, 27(9), p. 110600. Available at: https://doi.org/10.1016/j.isci.2024.110600.

Ziffra, R.S. *et al.* (2021) 'Single-cell epigenomics reveals mechanisms of human cortical development', *Nature*, 598(7879), pp. 205–213. Available at: https://doi.org/10.1038/s41586-021-03209-8.

Zweifel, S. *et al.* (2018) 'HOPX Defines Heterogeneity of Postnatal Subventricular Zone Neural Stem Cells', *Stem Cell Reports*, 11(3), pp. 770–783. Available at: https://doi.org/10.1016/j.stemcr.2018.08.006.

APPENDIX: Author's published scientific contribution

Secondary loss of *miR-3607* reduced cortical progenitor amplification during rodent evolution. *Sci. Adv.* 8, eabj4010 (2022)

DOI:10.1126/sciadv.abj4010