Toward Precision Bioengineering of Human-Induced Pluripotent Stem Cells: A Modular Platform for Transcriptional Reprogramming and Targeted Genomic Integration

Roy Hwang

This dissertation is submitted to the University of Ottawa
in partial fulfillment of the requirements for the

Doctorate in Philosophy degree in Cellular and Molecular Medicine

Department of Cellular and Molecular Medicine
Faculty of Medicine
University of Ottawa

© Roy Hwang, Ottawa, Canada, 2026

Abstract

Cellular reprogramming marks a new era of regenerative medicine and immunotherapy, enabling stem cell researchers to develop new models, therapeutics, and technologies with human-induced pluripotent stem cells (hiPSCs) at their core. However, hiPSC translational research is often limited by slow differentiation kinetics, heterogeneity within differentiated populations, transgene silencing during lineage commitment and unpredictable genomic profiles. Current methods for bioengineering hiPSCs lack the precision and modularity required to develop regenerative medicine-based therapeutics and complex modelling applications. This thesis describes the development of a modular bioengineering platform that demonstrates accelerated transcriptional reprogramming and precise, stable targeted genomic integration in human hiPSCs. By characterizing the feasibility of these two approaches, a framework and toolkit for downstream precise stem cell-based bioengineering were established.

Here, we designed a multiplexed CRISPR activation system for the directed differentiation of hiPSCs by targeting coordinated transcriptional programs and genes. This modular approach generated human iPSC-derived brain endothelial-like cells (iBECs), shortening conventional differentiation timelines by 61% while maintaining and expanding key phenotypic markers. Moreover, to achieve stable and safe transgene expression, we describe a site-specific engineering platform for hiPSCs across four genomic safe harbours (GSH): AAVS1, CCR5, hRosa26 and the novel hRogi1 locus. We evaluated the performance of modular genetic cassettes, benchmarking the stability of traditional (CAG, EFS) and unique (CBX3.1) promoters to mitigate gene silencing during lineage differentiation into germ layers. Finally, by using recombinase-mediated cassette

exchange (RMCE), we demonstrate a proof-of-principle “plug-and-play” pipeline that enables the swapping of synthetic cargoes, providing a modular platform for downstream translational applications. Collectively, this research delivers a comprehensively optimized bioengineering toolkit and blueprint for hiPSCs that improves differentiation fidelity and the reliability of genomic modifications. The modularity of this toolkit positions this research uniquely and lays the basis for the future advancement of next-generation cell therapies and complex models.

Lay Summary

Human induced pluripotent stem cells (hiPSCs) are cells that can transform into any specialized cell type and offer immense potential for studying diseases and developing novel therapeutics. However, conventional methods for programming and translating them into therapeutics and models are often complex, slow, inconsistent, and challenging. This dissertation aims to equip bioengineers who want to use hiPSCs for these practices by creating a genetic toolkit to safely, precisely, and rapidly engineer stem cells across a variety of contexts. First, by using a targeted gene activation technology (CRISPRa), I successfully designed and developed a system capable of activating key genes to program stem cells into functional brain endothelial cells, compressing a multi-week developmental process in under five days. These engineered cells can form tight barriers that mimic the human blood-brain barrier and provide a rapid and reliable platform to screen neurological drugs. The second component of this toolkit solves a safety hurdle found with genomic engineering practices. Stem cells tend to naturally turn off (silence) newly inserted therapeutic genes as they mature or reprogram. To overcome this, I've designed genomic landing pads and used a site-specific gene engineering platform into designated safe zones within the cell's DNA where they can operate without being disruptive. By systematically tracking various molecular components across different specialized cell types, this research mapped which genetic configurations remain reliably active and which are shut down during development and cell specialization. Altogether, these innovations provide a predictable blueprint that advances laboratory drug discovery pipelines and feasibility for future stem cell-based therapies.

Acknowledgments

I remember at the beginning of my graduate studies, sitting in a room full of senior scientists at the NRC as they reviewed and discussed my PhD proposal. As ideas and discussions were bouncing back and forth, I remember thinking to myself,

“How can I achieve that level of scientific rigour, knowledge and critical thinking to passionately give my take on a project like that?”

Well, the next five years really threw me into the whirlwind of highs and lows of my research endeavours and truly set me up for the next chapter of my journey as a scientist. This experience has shaped me to embody the ideal I envisioned as a new student, becoming the more critical and independent researcher I am today. Nonetheless, this entire journey and my growth as a scientist would not have been possible without the immense support and guidance from the university staff, colleagues, mentors, seniors, peers, family, partner and friends.

First, I express my sincere gratitude to my academic supervisor Dr. Mads Kaern. You were the first person I turned to at the start of my graduate studies and laid the groundwork for what it means to do science. Your guidance during those initial stages shaped my research interests, starting from biocomputing to bioinformatics and the application of synthetic biology and bioengineering frameworks. I am incredibly thankful for the wealth of transferable skills you taught me. Without your support, mentorship and your cross-collaborative networks with the National Research Council, this project wouldn’t have existed today.

Second, I want to express my deepest gratitude to my NRC co-supervisor Dr. Will J. Costain. Your mentorship over the past five years extended beyond the bench, shaping me into the independent researcher I am today. I am incredibly thankful for your mentorship in navigating science, starting from strategizing publications, networking at conferences, to your unwavering advocacy for my future career. Your support has kept me consistently motivated, and your wisdom has allowed me to actively pursue good science passionately and confidently.

Next, I want to thank my Thesis Advisory Committee members, Dr. Anna Jezierski, Dr. Maxime Rousseaux, and Dr. Ryan Russell. You all supported and challenged me to dive deeper into literature and work diligently towards my milestones every year. I am particularly grateful for how readily available every member was when support was needed. I want to extend a special thanks to Dr. Anna Jezierski at the NRC, who was always readily available to support my research. Your guidance was instrumental in shaping my stem cell-based research, culminating in our manuscript and review article. Beyond the bench, I am incredibly thankful for your grounding perspective, especially during challenging times, and your reminder that "*It's never worth losing sleep over science*".

To my peers and the entire HHT team at the NRC, thank you. Junzhuo Huang, Claudie Charlebois, Ewa Baumann, Betty Li, Marina Rukhlova, Rosa Comas, Yara Khalil, Julie Ouellette, Lea Desbiens, Yonghong Guan, and Caroline Sodja. From the hands-on training and daily support to engaging lab talks and shared weekend culture shifts, my experience as a graduate student would not have been possible without everyone. I am

incredibly fortunate to be surrounded by such a proud, supportive team with an immense wealth of knowledge and expertise.

I'd also like to thank our collaborators, senior NRC research officers, Dr. Tyler Renner, Dr. Scott McComb, Dr. Martin Loignon, Dr. Richard Marcotte, Dr. Simon Drouin, and Ziying Liu for their support and insight throughout my research.

I would also like to acknowledge the generous financial support that made this research possible. I am grateful to have received the OISB-NRC Graduate Scholarship and the University of Ottawa Doctoral Admission Scholarship. This research was funded by the NRC Collaborative R&D Initiative, the Cell and Gene Therapy Challenge Program, and by the Natural Sciences and Engineering Research Council of Canada (NSERC).

Finally, I would not be where I am today without the unwavering support of my friends, my partner, and my family. To my parents, brother in-law, my younger sister and older sister: thank you for being my constant foundation and for keeping me in check. My academic journey could have looked a lot different without you all by my side, and I thank you for every moment. I could not have reached this stage without your unconditional support. To my friends especially in the mundane city folks group chat: thank you for keeping me grounded and making sure I have a life outside of research, whether that meant dragging me to the soccer field or letting me disconnect online. And to my partner: thank you for being my anchor through the challenging moments of this degree and keeping me moving forward unconditionally. I could not have reached this milestone without you by my side, and I am excited to share our next chapter together.

This dissertation is dedicated to my parents
for the countless sacrifices they have made
to help me reach this milestone.

이 논문을 부모님께 바칩니다.
제가 이 자리에 서기까지 부모님께서 보여주신 셀 수 없이
많은 희생에 진심으로 감사드립니다.

Copyright Acknowledgments

1. Contents from the publication “Accelerated Reprogramming of hiPSCs into Functional Brain Endothelial-Like Cells Using Multiplexed CRISPR Activation” was reproduced and adapted for the thesis with permissions under Nature Portfolio:

“Authors have the right to reuse their article’s Version of Record, in whole or in part, in their own thesis. Additionally, they may reproduce and make available their thesis, including Springer Nature content, as required by their awarding academic institution. Authors must properly cite the published article in their thesis according to current citation standards.”

Roy W. Hwang et al., Scientific Reports, Accelerated Reprogramming of hiPSCs into Functional Brain Endothelial-Like Cells Using Multiplexed CRISPR Activation, 2026, Nature Portfolio.

2. Contents from the manuscript *“Programming Brain Endothelial Identity: Transcription Factor and Genome Engineering Approaches to Improve iPSC-derived BBB Models”* were reproduced and adapted for this thesis. This manuscript is currently submitted for publication and under review. As an unpublished work, the authors currently retain full copyright.

“As per Elsevier’s policy, authors of open-access articles retain the right to include their published article, in full or in part, in a thesis or dissertation. The original article is published under a Creative Commons license in Stem Cell Reports.”

Roy W. Hwang, Betty (Ying) Li, Richard Marcotte, Will Costain, Anna Jezierski, Stem Cell Reports, Programming Brain Endothelial Identity: Transcription Factor and Genome Engineering Approaches to Improve iPSC-derived BBB Models, 2026, Cell Press.

3. Contents from the publication Recent Developments in Delivering Biotherapeutics and Gene Therapy Across the Blood-Brain Barrier. Was added to the appendix with permissions from Springer Nature.

Umar Iqbal, **Roy W. Hwang**, Will J. Costain. (2026). Recent Developments in Delivering Biotherapeutics and Gene Therapy Across the Blood-Brain Barrier. BioDrugs. Springer Nature.

"Reproduced with permission from Springer Nature. As a Springer Nature journal author, the author retains the right to reuse the Version of record, in whole or in part, in their own thesis. Written consent for reuse has been obtained from all co-authors."

List of Abbreviations

3TF-SAM	Multiplexed three transcription factor CRISPRa vector
4EG-SAM	Multiplexed four endothelial gene CRISPRa vector
AAVS1	Adeno-associated virus integration site 1
ANOVA	Analysis of variance
atgRNA	Attachment guide RNA
attB	Attachment site of the bacterium
attL	Attachment site left
attP	Attachment site of the phage
attR	Attachment site right
BBB	Blood-brain barrier
BceINTa	*Bacillus cereus*-derived serine integrase
BEC	Brain endothelial cell
BFP	Blue fluorescent protein
bp	Base pairs
Bxb1	Mycobacteriophage-derived serine integrase
Caco2	Human colorectal adenocarcinoma cell line
CAG	Hybrid CMV enhancer/chicken-β actin
CAR	Chimeric antigen receptor
CBX3.1	Chromobox 3.1 Promoter
CCR5	Chemokine receptor 5 integration site
CDH5	Cadherin-5 gene
cDNA	Complementary DNA
CLDN5	Claudin-5 gene
CMFDA	5-chloromethylfluorescein diacetate
CMV	Cytomegalovirus
CNS	Central nervous system

copGFP	Copepod green fluorescent protein
CRISPR	Clustered regularly interspaced short palindromic repeats
CRISPRa	CRISPR-activation
crRNA	CRISPR RNA
Ct	Cycle threshold
CXCR4	C-X-C chemokine receptor type 4
dCas9	Dead cas9
DEST	Destination
DGE	Differential gene expression
DMEM	Dulbecco’s modified Eagle’s medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
Dox	Doxycycline
EC	Endothelial Cell
ECM	Extracellular matrix
EDTA	Ethylenediaminetetraacetic acid
EF1α	Elongation factor 1-alpha
EFS	Elongation factor 1-alpha short
EGFR	Epidermal growth factor receptor
EM	Endothelial media
Epi	Epithelial
ERG	ETS-related gene
ETS	E26 transformation-specific
ETV2	ETS variant transcription factor-2 gene
FACS	Fluorescence-activated cell sorting
FBS	Fetal bovine serum
FBS	Fetal bovine serum
FLI1	Friend leukemia integration-1 gene

FLK1	Fetal liver kinase 1
FSC	Forward scatter
GCU	Green calibration unit
GCV	Ganciclovir
gDNA	Genomic DNA
GFP	Green fluorescent protein
GII	Green integrated intensity
GJA1	Gap junction protein alpha 1
GLUT1/SLC2A1	Glucose transporter type 1 gene
GO	Gene ontology
GSEA	Gene set enrichment analysis
GSH	Genomic safe harbour
hAF	Human amniotic fluid-derived hiPSC line
hAF09	hAF line version 09
hAF12	hAF line version 12
hBEM	Human brain endothelial media
hBMEC	Primary human brain microvascular endothelial cell
hCMEC/D3	Immortalized human cerebral microvascular endothelial cell line
HDMEC	Human dermal microvascular endothelial cells
HDR	Homology-directed repair
HeLa	Henrietta Lacks immortalized cell line
hiPSC	Human induced pluripotent stem cell
hROSA26	Human reverse orientation splice acceptor integration site
HSV1-TK	Herpes simplex virus type 1 thymidine kinase
HUVEC	Human umbilical vein endothelial cells
iBEC	hiPSC-derived brain endothelial-like cell
IGF1R	Insulin-like growth factor 1 receptor gene
iPSC	Induced pluripotent stem cell

KDR	Kinase insert domain receptor
KLF4	Krüppel-like factor 4
KOEB	Knock-out pre-differentiation media
KRT7	Keratin 7
M-MLV-RT	Moloney murine leukemia virus reverse transcriptase
Mc	Minicircle
MFI	Mean fluorescence intensity
mRNA	Messenger RNA
mTESR	Modified Tenneille's secret recipe media
NaFl	Sodium fluorescein
ngRNA	Nicking guide RNA
NGS	Next-generation sequencing
NHEJ	Non-homologous end joining
NK Cells	Natural Killer Cells
OCLN	Occludin gene
OCT4	Octamer-binding transcription factor 4
P2A	Porcine teschovirus-1 2A peptide
PASTE	Programmable addition via site-specific targeting elements
PAX6	Paired box 6
Pb	PiggyBac
Pbase	PiggyBac transposase
PCA	Principal component analysis
PCR	Polymerase chain reaction
P_e	Paracellular permeability
PE	Prime editing
PECAM1	Platelet and endothelial cell adhesion molecule-1 gene
PEV	Position-effect variegation
PGK	Phosphoglycerate kinase

pMini	Parental minicircle
Poly-A	Polyadenylation
PT	Parental
RA	Retinoic acid
RGS5	Regulator of G protein signalling 5
RMCE	Recombinase-mediated cassette exchange
RNA	Ribonucleic acid
RNA-seq	RNA-sequencing
ROCK	Rho-associated protein kinase
Rogi1	Genomic safe harbour located on the DDX58 gene
RT	Reverse transcription
RT-qPCR	Quantitative reverse transcription polymerase chain reaction
SAM	Synergistic activation mediator
SEM	Standard error of the mean
sgRNA	Single guide RNA
SOX17	SRY-box transcription factor 17
SOX18	SRY-box transcription factor 18
SOX2	SRY-box transcription factor 2
SOX7	SRY-box transcription factor 7
SSC	Side scatter
SSEA3	Stage-specific embryonic antigen-3
SSEA4	Stage-specific embryonic antigen-4
TEER	Transendothelial electrical resistance
TfR1	Transferrin 1 receptor gene
TJ	Tight junction
TK	Thymidine kinase
TMEM30A	Transmembrane protein 30A gene
TRA1-60	Tumor rejection antigen 1-60

TRA1-81	Tumor rejection antigen 1-81
tracrRNA	Trans-activating CRISPR RNA
UTR	Untranslated region
VEGF	Vascular endothelial growth factor
VEGFR	Vascular endothelial growth factor receptor
Wnt	Wingless-related integration site gene
WPRE	Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element
WT	Wild type
ZO1	Zonula occludens-1 protein gene

List of Figures

Figure 1.1: Schematic Illustration of Primed Pluripotent Stem Cell (PSC) Differentiation Pathways.

Figure 2.1: Structural Features of PASTE sgRNA Constructs. (Left) pDY0209 ngRNA Plasmid (2.3 kb).

Figure 2.2: Design and Construction of 3TF-SAM and 4EG-SAM Multiplexed CRISPRa Plasmids.

Figure 2.3: Schematic Representation of 2D In Vitro Blood Brain Barrier Modelling.

Figure 3.0: Graphical Summary of Multiplexed CRISPRa Reprogramming of iPSCs.

Figure 3.1: Identifying and Screening Brain Endothelial Gene Targets and sgRNAs for CRISPRa in HeLa Cells.

Figure 3.2: Multiplexed CRISPRa Integration in hiPSCs Induces Core BEC Genes and Transcription Factors.

Figure 3.3: CRISPRa-mediated Induction of BBB Genes and Transcription Factors Initiates Directed Differentiation into iBEC-like Cells.

Figure 3.4: Quantitative Flow Cytometric Analysis Reveals Enhanced Expression of Brain Endothelial and BBB-specific Markers Following Multiplexed CRISPRa Reprogramming.

Figure 3.5: Immunofluorescence Analysis of Brain Endothelial Cell Marker Expression and Junctional Localization in CRISPRa-differentiated iBECs.

Figure 3.6: Functional Characterization of Barrier Integrity, Selective Permeability and Angiogenic Potential in iBECs.

Figure 3.7: Transcriptome Profiling of CRISPRa Differentiated iPSC-derived Brain Endothelial Cells.

Figure 4.0: Graphical Summary of Chapter 4.

Figure 4.1: Mechanism of Programmable Addition via Site-specific Targeting Elements (PASTE).

Figure 4.2: Schematic Overview of a Two-step PASTE and RMCE Platform for Modular and Site-specific Genomic Engineering in hiPSCs.

Figure 4.3: Iterative Optimization of Expression Cassettes to Enhance Transgene Expression.

Figure 4.4: Developing a Prototype PASTE-editing and Recombinase-mediated Cassette Exchange (RMCE) Platform in HeLa Cell Lines.

Figure 4.5: Investigating and Optimizing Minicircle Transgene Expression in hAF09 hiPSC Lines.

Figure 4.6: PASTE Editing hiPSCs for Pooled and Clonal Cell Line Generation.

Figure 4.7: Genomic DNA (gDNA) Validation of PASTE-edited hiPSCs.

Figure 4.8: Evaluating transgene silencing dynamics across a panel of locus and promoter dependencies in PASTE-edited hiPSCs during trilineage differentiation.

Figure 4.9: Recombinase-mediated Cassette Exchange in CBX3.1 PASTE-edited hAF09 hiPSC (AAVS1b).

Figure 5.1 Heatmap of Transgene Expression Across Germ Layer Lineages.

List of Tables

Table 1.1: Comparison of iPSC-derived brain endothelial cell (iBEC) as effective blood–brain barrier (BBB) models.

Table 1.2: Transdifferentiation and Directed Differentiation Approaches for Endothelial Cell Differentiation.

Table 1.3: Overview of CRISPR Classes and Types with Genome and RNA Engineering Capabilities.

Table 1.4: Modern Genomic Safe Harbour Site Criteria.

Table 2.1: Brain Endothelial sgRNA Chromosomal Locations.

Table 2.2: Broad Institute's CRISPick sgRNA Analysis.

Table 2.3: Genomic Safe Harbour gRNAs.

Table 2.4: List of Primers

Table 2.5: List of Human Cell Lines Used in this Thesis.

Table 2.6: List of Antibodies.

Table 2.7: Summary of Plasmids and Vectors Utilized and Generated.

Table 3.1: Genetic Markers for Brain Endothelial Cell Lineage Reprogramming.

Table 4.1: Landscape of Validated and Characterized Genomic Safe Harbour Loci.

Table 4.2: Optimized PASTE-editing Reaction Components for 96-well Nucleofection Format.

Table 4.3: Expected Amplicon Length for gDNA PCR Validation of PASTE-edited Cell Lines.

Table 4.4: Markers Assessed for Germ Layer Specification in Differentiated Cells.

List of Supplementary Figures

Figure S1: Identifying Lead Candidate sgRNA for Multiplex CRISPR Brain Endothelial Gene Induction.

Figure S2: Doubling Time of Multiplexed CRISPRa Transfected hiPSCs During Antibiotic Selection.

Figure S3: Morphology of 3TF-SAM and 4EG-SAM hiPSCs in Conventional Differentiation Schemes.

Figure S4: Directed Differentiation in hAF09 and hAF12 hiPSC Cell Lines.

Figure S5: Uncropped Western Blot of CRISPR-Cas9 in Various hiPSCs.

Figure S6: Transcriptomic PCA Analysis Between 3TF-SAM and 4EG-SAM iBECs.

Figure S7: Pearson Correlation Heatmap of Described Cell Lines.

Figure S8: Global Expression Heatmap for Top 20 Differentially Expressed Genes.

Figure S9: Capillary Brain Endothelial Genes and Choroid Plexus Genes Gene Expression Heatmap.

List of Supplementary Tables

Table S1: Top 20 Most Differentially Expressed Genes in 3TF-SAM iBECs.

Table S2: Top 20 Most Differentially Expressed Genes in 4EG-SAM iBECs.

Table S3: Data Compilation of all Described Studies for Meta-analysis.

Table of Contents

ABSTRACT II

LAY SUMMARY IV

ACKNOWLEDGMENTS V

COPYRIGHT ACKNOWLEDGMENTS IX

LIST OF ABBREVIATIONS XI

LIST OF FIGURES XVII

LIST OF TABLES XIX

LIST OF SUPPLEMENTARY FIGURES XX

LIST OF SUPPLEMENTARY TABLES XX

TABLE OF CONTENTS XXI

CHAPTER 1: INTRODUCTION 1

1.0 Declaration of Previously Published Material 1

1.1 The Bioengineering Landscape of Human iPSCs 1
1.1.1 Human Induced Pluripotent Stem Cells: From Early Discovery to Therapeutic Chassis 1
1.1.2 Pluripotent States, Lineage Commitment and Cell Fate Specification 2
1.1.3 The Ectoderm: The Antagonistic Relationship Between The Default Neuroectoderm and Non-Neural Ectoderm 4
1.1.4 Cuing the Cells to Endodermal and Mesodermal Lineages From the Mesendoderm Progenitor 5
1.1.5 Epigenetic Barriers and Limitations of Morphogen-Driven Differentiation 8
1.1.6 The Need for Precision Bioengineering 8

1.2 The Biological Challenge: Modelling the Blood-brain Barrier (BBB) 9
1.2.1 The Blood-brain Barrier 9
1.2.2 The Structure and Components of the Neurovascular Unit 10
1.2.3 Brain Endothelial Cells 11
1.2.4 Pericytes 12
1.2.5 Astrocytes 13
1.2.6 Deficiencies in Conventional BEC *In Vitro* Models 13

1.2.7 Bridging the Translational Gap in BBB Modelling: Human iPSC-derived Brain Endothelial-like Cells (iBECs) 14
1.2.8 Defining the Gap: Current Challenges with iBEC BBB Modelling 17

1.3 The Engineering Solution for iBEC Generation: Programmable Transcriptional Modulation 21
1.3.1 Shifting Away from Traditional Methods for Cell Fate Specification 21
1.3.2 Transcription Factor-guided Approaches to Lineage Programming with Directed Differentiation and Transdifferentiation Strategies 21
1.3.3 Inducing BEC Identity Through Lineage-Specific Transcription Factors 24
1.3.4 Context-dependent effects of FOXF2 and ZIC3 reprogramming 25
1.3.5 Small-molecule pathway modulation as a complementary strategy 26
1.3.6 Combinatorial transcription factor programs for BEC fate control 27
1.3.7 Programmable Control of Brain Endothelial Specification and Maturation 29
1.3.8 Temporal Control of Transcription Factor Expression in BBB-Directed Reprogramming 30

1.4 The Bioengineering Toolkit: Genetic, Epigenetic and Transcriptomic Tools for Programmable Control 34
1.4.1 The Pioneer of Genomic Engineering Tools 34
1.4.2 Modern Era of Genome Engineering: An Overview of the CRISPR/Cas Toolkit 36
1.4.2 Gene Editing Through Repair Pathways: Homology Directed Repair and Non-Homologous End Joining 39
1.4.3 Genome Editing Through Base Editors 39
1.4.4 Genome Editing Through Prime Editing 40
1.4.5 Large Transgene Integration Using NHEJ Repair Pathway: Homology-Independent Targeted Integration 41
1.4.6 Genome Editing Through Site-Specific Recombination 41
1.4.7 DSB-free Integration of Large Transgenes: Programmable Addition via Site-Specific Targeting Elements 42
1.4.8 Transcriptional Modulation: CRISPRa and CRISPRi 43
1.4.9 Stable and Heritable Epigenome Regulation: CRISPRon and CRISPRoff 45
1.4.10 RNA Regulation: Cas13 and CasRx 46

1.5 Safe and Targeted Delivery of Transgenes to the Human Genome 47
1.5.1 Survey of Transgene Integration Technologies 47
1.5.2 Mechanisms of Transgene Silencing in Differentiating hiPSCs 50
1.5.3 Genomic Safe Harbours as Targets for Expression Stability 51
1.5.4 Historical “Gold-Standard” Loci: Pseudo-Safe Harbour Sites 53
1.5.5 Modern Genomic Safe Harbours: Focusing on Intergenic Loci 54

1.6 Thesis Objectives 57
1.6.1 Rationale and Hypothesis 57
1.6.2 Objectives 58

CHAPTER 2: METHODOLOGY 60

2.1 Plasmid Design and Molecular Biology 60
2.1.1 *In Silico* sgRNA Design for CRISPRa 60
2.1.2 *In Silico* atgRNA, ngRNA, pMini Design and Safe Harbour Selection 60
2.1.3 Design and Construction of Multiplexed CRISPRa Cassette 62
2.1.4 Design and Construction of PASTE Components 65
2.1.5 Minicircle Induction and Purification 66
2.1.6 Design and Construction of RMCE Donor Cassette 67
2.1.7 Plasmid Validation and Sequencing 67

2.2 Cell Culture and Cell Engineering 67
2.2.1 Maintenance of HeLa Cell Lines 67
2.2.2 Maintenance of Human Induced Pluripotent Stem Cell (hiPSC) Lines 68

2.2.3 Cell Line Quality Control and Stability 68
2.2.4 sgRNA Validation and Prototyping CRISPRa in HeLa-SAM Cells 68
2.2.5 Generation of Multiplexed CRISPRa hiPSC Cell Lines 69
2.2.6 Validation and Prototyping PASTE Integration Efficiency in HeLa Cells 70
2.2.7 Generation of PASTE-edited HiPSC Cell Lines 70
2.2.8 Isolation of PASTE-edited Clonal and Pooled hiPSC Cell Lines with FACS 71
2.2.9 Generation and Isolation of RMCE-edited HiPSC Cell Lines 72

2.3 Differentiation of Human Induced Pluripotent Stem Cells 73
2.3.1 Standard Brain Endothelial Differentiation 74
2.3.2 CRISPR-mediated Brain Endothelial Differentiation 75
2.3.3 Trilineage Germ Layer Differentiation 76

2.4 Genomic, Transcriptomic and Protein Analysis 77
2.4.1 Genomic DNA PCR 77
2.4.2 Quantitative RT-qPCR 78
2.4.3 RNA-sequencing 79
2.4.4 Western Blotting 80

2.5 Phenotypic and Functional Characterization 81
2.5.1 Flow Cytometry, FACS and Fluorescence Microscopy Instrumentation 81
2.5.2 Assessment of Pluripotency Markers 83
2.5.3 Characterization of Brain Endothelial and Blood-Brain Barrier Markers 83
2.5.4 Trilineage Germ Layer Assessment 83
2.5.5 Small Molecular Permeability Assay (Sodium Fluorescein) 84
2.5.6 Transendothelial Electrical Resistance (TEER) Analysis 86
2.5.7 Angiogenesis Assay 87
2.5.8 Longitudinal Assessment of Promoter Stability and Transgene Silencing 88

2.6 Computational, Statistical Analysis and Figures 89
2.6.1 Differential Gene Expression and Pathway Analysis 89
2.6.2 General Statistical Methods 90
2.6.3 Figures 90
2.6.4 Data Availability 90

2.7 List of Tables 90
Table 2.1: Brain Endothelial sgRNA Chromosomal Locations. 90
Table 2.2: Broad Institute's CRISPick sgRNA Analysis. 91
Table 2.3: Genomic Safe Harbour gRNAs. Sequence identity is as follows: Spacer, Scaffold, Reverse Transcription (RT) Template, BceINTa attB, Primer binding site (PBS). 93
Table 2.4: List of Primers 94
Table 2.5: List of Human Cell Lines Used in this Thesis. 100
Table 2.6: List of Antibodies. 100
Table 2.7: Summary of Plasmids and Vectors Utilized and Generated. 101

CHAPTER 3: ACCELERATED REPROGRAMMING OF HIPSCS INTO FUNCTIONAL BRAIN ENDOTHELIAL-LIKE CELLS 104

3.0 Findings Summary 105

3.1 CRISPRa Screening of Guide RNAs for Brain Endothelial Gene Targets 106

3.2 Constructing a Multiplexed CRISPRa System Targeting BEC Genes 107

3.3 Validation of CRISPR Activation of Target Genes in hiPSCs 109

3.4 Activation of 3TF or 4EG Genes Induces and Accelerates hiPSC BEC Differentiation 113

3.5 Characterization of BBB Phenotype Induction in CRISPRa-differentiated iBECs 117

3.6 Transcriptomic analysis of 3TF-SAM and 4EG-SAM iBECs 126

CHAPTER 4: ESTABLISHING A SITE-SPECIFIC AND MODULAR GENOMIC ENGINEERING PLATFORM IN HIPSCS 131

4.0 Findings Summary 131

4.1 Design and Strategy for Site-specific Genomic Engineering 133

4.2 Iterative Optimizations to the Synthetic Reporter Transgene 139

4.3 Prototyping the Delivery and Integration of PASTE-editing 143

4.4 Optimization of the Cargo Expression Cassette in hiPSCs 149

4.5 PASTE Editing in hiPSCs and Development of a PASTE-edited Clonal Cell Line Pipeline 153

4.6 Genomic Validation of GSH Targeted Edits 158

4.7 Gene Silencing in PASTE-edited Trilineage Germ Layer Differentiations 162

4.8 Genomic Transgene Insertion by Recombinase-Mediated Cassette Exchange 168

CHAPTER 5: DISCUSSION 171

5.0 Chapter Overview 171

5.1 Accelerated Reprogramming of hiPSCs Into Functional Brain Endothelial-like Cells 171
5.1.1 Transcriptional Reprogramming Accelerates Differentiation and Induces Functional iBEC Phenotypes. 172
5.1.2 Direct Activation of Structural Genes Yields Enhanced Marker Expression. 173
5.1.3 hiPSC-derived BECs Exhibit a Hybrid Heterogeneous Cell State and Late-stage Epigenetic Silencing. 174
5.1.4 Transcriptomic Upregulation of Key Endothelial Identity and Functional Genes. 176
5.1.5 Future Adaptations and Direction: Multiplexed Modulation of the Neurovascular Unit. 177
5.1.6 Technical Limitations: Vector Integration Variability and Cell Line Generalizability. 178

5.2 Establishing a Site-specific and Modular Genomic Engineering Platform in hiPSCs 179
5.2.1 Design and Validation of Modular PASTE Integration Strategy 180
5.2.2 Gene Expression in PASTE-edited hiPSCs 182
5.2.3 Transgene Silencing in Lineage-specific Differentiations 184
5.2.4 The Oversimplification of GSH Safety and Stability Criteria 187
5.2.5 Evaluating UCOE Insulators and Promoters Across Germ Layers 189
5.2.6 Cargo Constraints and Transgene Structure 190
5.2.7 Transgene Expression at Epigenetic Checkpoints to Predict Silencing 191
5.2.8 Recombinase-mediated Cassette Exchange Function 191
5.2.9 Future Directions 193

5.3 Conclusion **197**

CHAPTER 6: BIBLIOGRAPHY **198**

CHAPTER 7: SOFTWARE AND AI USE ACKNOWLEDGEMENTS **232**

7.1 Illustration Design and Figure Generation **232**

7.2 AI and Writing Assistance **232**

7.3 Reference Management Software **232**

CHAPTER 8: APPENDIX **233**

8.1 Supplementary Figures **233**
Figure S1: Identifying Lead Candidate sgRNA for Multiplex CRISPR Brain Endothelial Gene Induction. 233
Figure S2: Doubling Time of Multiplexed CRISPRa Transfected hiPSCs During Antibiotic Selection. 233
Figure S3: Morphology of 3TF-SAM and 4EG-SAM hiPSCs in Conventional Differentiation Schemes. 234
Figure S4: Directed Differentiation in hAF09 and hAF12 hiPSC Cell Lines. 235
Figure S5: Uncropped Western Blot of CRISPR-Cas9 in Various hiPSCs. 236
Figure S6: Transcriptomic PCA Analysis Between 3TF-SAM and 4EG-SAM iBECs. 237
Figure S7: Pearson Correlation Heatmap of Described Cell Lines. 238
Figure S8: Global Expression Heatmap for Top 20 Differentially Expressed Genes. 239
Figure S9: Capillary Brain Endothelial Genes and Choroid Plexus Genes Gene Expression Heatmap. 240
Figure S10: RNA-seq TPM Bar Charts. 241

8.2 Supplementary Tables **242**
Table S1: Top 20 Most Differentially Expressed Genes in 3TF-SAM iBECs. 242
Table S2: Top 20 Most Differentially Expressed Genes in 4EG-SAM iBECs. 243
Table S3: Data Compilation of all Described Studies for Meta-analysis. 244

8.3 Manuscript: Recent Developments in Lipid Nanoparticle-mediated Delivery of Biotherapeutics and Gene Therapy Across the Blood-brain Barrier. **245**

Chapter 1: Introduction

1.0 Declaration of Previously Published Material

Portions of the introduction presented in this chapter have been adapted from the following manuscript:

Roy W. Hwang, Betty (Ying) Li, Richard Marcotte, Will Costain, Anna Jezierski. Programming Brain Endothelial Identity: Transcription Factor and Genome Engineering Approaches to Improve iPSC-derived BBB Models, 2026.

1.1 The Bioengineering Landscape of Human iPSCs

1.1.1 Human Induced Pluripotent Stem Cells: From Early Discovery to Therapeutic Chassis

With the first discovery of stem cells, the idea of programming cells into any specialized cell type brought immeasurable promise and anticipation in the biomedical research community. In 1961, Ernest McCulloch and James Till pioneered the discovery of hematopoietic stem cells in murine bone marrow[1]. In 1963, they established that multipotent stem cells – stem cells with a limited range of differentiating into lineages – formed colonies[2] and retained self-renewal properties[3], laying the foundation for downstream stem cell biology. Subsequently, James Thomson in 1998 demonstrated the first isolation of human embryonic stem cells (hESCs), capable of differentiating into any cell type, marking the first discovery and use of pluripotent stem cells. While multipotent stem cells have been thoroughly characterized and explored in clinical trials in the early 2000s[4], pluripotent stem cells (PSCs), although promising, were harshly challenged by the public due to ethical and moral barriers. More specifically, the clinical translation of hESCs was widely limited as they are derived from the inner cell mass of preimplantation

blastocysts. The Nobel Laureate Shinya Yamanaka made a breakthrough in pluripotent stem cell technology by demonstrating the reprogramming of murine somatic cells to a pluripotent state by introducing four transcription factors: *Oct3/4, Sox2, c-Myc*, and *Klf4*[5]. By 2007, this scheme was widely adopted to generate human induced pluripotent stem cells (hiPSCs), mitigating the ethical constraints of hESCs, while also providing a scalable, patient-specific platform to model complex diseases and cell therapies. As of today, the growing interest and applications of hiPSC technology can be observed across a multitude of research fields, ranging from bioengineering[6,7], differentiation[8,9], pharmacological studies[10–12], and disease modelling[13–15], yielding over a combined 38,000 publications recorded in PubMed.

1.1.2 Pluripotent States, Lineage Commitment and Cell Fate Specification

Human iPSCs rely on their self-renewing and pluripotent properties to remain locked into their undifferentiated state, at the ready for lineage commitment upon exposure to extracellular cues. Pluripotent stem cells exist in two main states: Naïve (pre-implantation) and primed (post-implantation) states[16]. In the naïve state, pluripotency is maintained independently of *ERK* signalling and utilizes the *LIF*/*STAT3* pathway[17,18]. On the other hand, hiPSCs conventionally reside in a primed state, with Activin/Nodal pathways activated to maintain pluripotency and Fibroblast Growth Factor (*FGF*), which positions the cells in a primed-like state poised for differentiation, oftentimes biasing towards the neuroectodermal fate[19–21].

The pluripotency of stem cells is governed by a network of master regulatory factors in an autoregulatory forward feedback loop (*OCT4*, *SOX2* and *NANOG*)[22,23]. When pluripotency is activated, co-activators such as acetyltransferase p300 will be recruited to

enhancer regions that regulate pluripotent networks[24]. Simultaneously, a dual function of these regulators also recruits co-repressors to early developmental programs such as the *Pax6* (Ectoderm)[22,25,26], *Brachyury/TBXT* (Mesoderm)[27–29], and *Sox17* (Endoderm)[30,31]. Once bound to these lineage-specific promoters, chromatin-modifying complexes, including Nucleosome Remodelling and Deacetylase (NuRD) complex[32] and Polycomb Repressive Complexes[26] (PRC1 and PRC2), are recruited to induce epigenetic silencing.

Evidently, PRC2 is responsible for methylation using H3K27me3, which also occurs with H3K4me3 by the Trithorax group (TrxG) complexes[33]. This bivalent domain enables the pluripotent state to silence lineage-driving genes while also keeping them readily available for rapid activation[33]. When differentiation cues are induced, demethylases rapidly remove the repressive H3K28me3 marks, resolving the bivalent domains and driving robust transcriptional activation of the directed cell fate.

The moment a stem cell receives a differentiation cue (e.g., BMP4, retinoic acid, or loss of FGF2), pluripotent regulators are transcriptomically downregulated by interfering with *OCT4* and *NANOG* expression. Additionally, other protein complexes such as E3 ubiquitin ligase (e.g., SCF, APC, CUL3, RING1)[34] act as gatekeepers that readily and swiftly degrade pluripotent master regulators, while developmental genes previously held in a repressive state can be demethylated by UTX or JMJD3[35–37]. Collectively, this initiates a rapid loss in the pluripotent feedback loop, and the versatile and “open” euchromatin shifts towards a structured and collapsed heterochromatin, responding and pushing forward to its designated lineage fate.

1.1.3 The Ectoderm: The Antagonistic Relationship Between The Default Neuroectoderm and Non-Neural Ectoderm

Previous research demonstrated that primed PSCs naturally default towards the neuroectodermal state[38–40]. The neuroectoderm gives rise to components of the central nervous system (CNS) and the peripheral nervous system (PNS)[41,42]. However, cells must actively repress specific transduction pathways driving the other two germ layers to ensure cell differentiation homogeneity. First, to prevent primitive streak and definitive endoderm differentiation, the TGF-β/Activin/Nodal pathways are inhibited, paired with the suppression of SMAD2 and SMAD3, thereby preventing the activation of endodermal transcription factors such as Brachyury and SOX17[43–47]. Simultaneously, the Bone Morphogenetic Protein 4 (BMP4) is inhibited by recombinant proteins such as Noggin, restricting additional SMAD signalling for efficient neural conversion[46,48].

A major biological factor that directs primed PSCs towards the neural ectodermal lineage is their endogenous expression of Fibroblast Growth Factor (FGF) mediated by the ERK1/2 pathway[49,50]. It is well established that the FGF and BMP pathways behave in a biphasic antagonistic feedback loop during early ectodermal signalling[51,52]. For example, endogenous FGF signalling antagonizes BMP signalling and drives the activation of the PARP-1 cascade, upregulating neuroectodermal transcription factors such as PAX6[49,53]. On the other hand, the introduction of BMP actively inhibits neural commitment in the cells and promotes epidermal genes to be expressed[54,55]. As a result, to override neural commitment within the ectodermal lineage, PSCs require suppression of FGF signalling combined with the exogenous activation of BMP and or the Wnt signalling pathways[56,57].

1.1.4 Cuing the Cells to Endodermal and Mesodermal Lineages From the Mesendoderm Progenitor

Pluripotent stem cells cannot directly transition to the mesoderm and endoderm lineages. Prior to the differentiation, primed PSCs must transition to a progenitor bipotent intermediate stage known as the mesendoderm[58,59]. Upon receiving canonical *Wnt* signalling, GSK3-β is deactivated, resulting in β-catenin accumulation in the cytoplasm[60–62]. This molecule complex translocates to the nucleus and binds to and activates TCF/LEF transcription factors[63–65]. Simultaneously, Nodal/TGF-β signalling leads to the phosphorylation of SMAD2/3, which complexes with SMAD4, and interacts with β-catenin-TCF-LEF complex to activate downstream primitive streak genes[45,66–71]. The hallmark mesendoderm transcription factors, such as Brachyury and Eomesodermin (EOMES), are expressed and prepare chromatin for either mesoderm or endoderm specification[72–75].

The mesendoderm is pushed towards the mesoderm stage by the control of the signalling gradients and master regulatory factors. While the prolonged increased expression of Nodal signalling drives the cells towards endoderm fate, the temporal downregulation of Nodal permits the cells to enter the mesoderm lineage[73]. The principal driver for mesoderm specification relies largely on the activation of *BMP4*[76,77] and *FGF*[78] signalling cascades. This initiates downstream induction of epithelial-to-mesenchymal transition (*EMT*) transcription factors such as *Snail* and *Twist*[79]. As a result, epithelial factors such as *E-cadherin* are downregulated, while mesenchymal markers such as *N-cadherin* are upregulated. Consequently, at the mesoderm-defining stage, Brachyury expression is stably activated[75], thereby repressing endodermal gene programs.

Comparatively, the endoderm specification is enabled through the persistent high expression of Nodal signalling[80]. High nodal signalling enables the expression of Forkhead Box A2 (FOXA2) transcription factor that binds to compacted chromatin and structurally remodels the chromatin[81] for endoderm-specific genetic programs[82]. For example, FOXA2 displaces linker histones to open enhancer regions specifically for the endoderm-specific master regulator such as the SRY-Box Transcription Factor 17 (SOX17)[82]. SOX17 locks in the endodermal cell fate by repressing downstream Wnt/β-catenin signalling[83] to transcriptionally shut down mesodermal programs like Brachyury and preventing the cells from diverging into the mesoderm lineage.

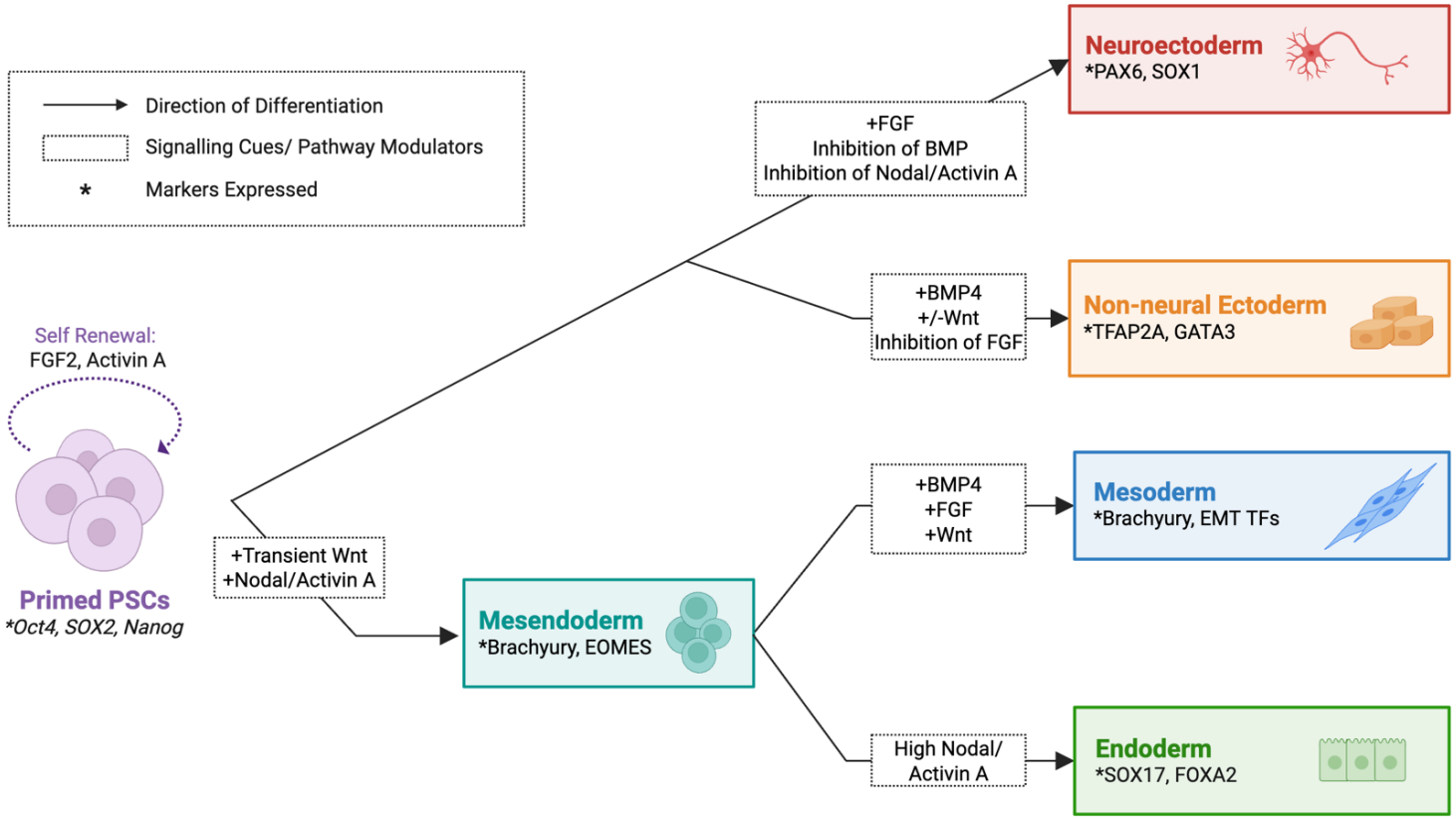


Figure 1.1: Schematic Illustration of Primed Pluripotent Stem Cell (PSC) Differentiation Pathways. Gradient and tunable signalling cues and pathway modulators guide the primed PSCs towards each germ layer lineage gate. PSCs are maintaining a pluripotent state through self-renewal factors such as FGF2 and Activin A signalling, with the core regulatory markers *Oct4, SOX2*, and *Nanog*. Subsequent commitment towards neuroectoderm, non-neural ectoderm, mesoderm, and endoderm is governed by gradient temporal activation or inhibition of signalling molecules described in this pathway. Transcriptional factors and structural markers are denoted as asterisks in each lineage (*). This schematic was synthesized from a collection of manuscripts[21,46,47,73,77,84–89].

1.1.5 Epigenetic Barriers and Limitations of Morphogen-Driven Differentiation

Historically, the mechanisms that control germ layer specification were understood through morphogen–signalling molecules that direct cell fate–gradients, such as *Nodal*, *Wnt*, *BMP* and *FGF*. However, recent research demonstrates that cell differentiation is also predetermined by its epigenetic landscape, including DNA methylation states acting as gatekeepers for the cell's ability to adopt a lineage[90]. Efforts in mitigating these challenges often looked at resetting these methylated states by reverting cells to naïve PSCs, which showed some improvement in differential potential[91]. However, a bottleneck remains in cellular reprogramming, where traditional differentiation schemes rely heavily on extracellular components to initiate signalling cascades to mimic these intracellular epigenetic changes. Exogenous growth factors and culture conditions are constantly empirically optimized and sometimes lead to stochastic and impure lineage commitments, batch variability, heterogeneity and off-target fates[92–97]. To overcome these challenges in stem cell engineering, a shift away from using extrinsic morphogens and growth factors is required and instead necessitates moving toward synthetic bioengineering of the cell's intrinsic transcriptional machinery[98–100].

1.1.6 The Need for Precision Bioengineering

As regenerative medicine and developmental biology continue to elucidate complex mechanisms and networks for these pathways, requirements for precise and stringent engineering of stem cells remain a challenge. Bioengineering efforts have historically advanced regenerative medicine through organoid research[101–103], generating biomaterials[104–106], and innovating cell line development and engineering[107–110]. However, efforts in developing methods to bypass epigenetic barriers in cell programming, and the

challenges with stochasticity observed in conventional differentiation schemes, remain to be fully addressed.

This demands implementation of the emerging field of study, synthetic biology, which applies engineering principles to design, build and optimize novel biological systems[111,112]. Within the context of stem cell engineering, synthetic biology holds potential in revolutionizing traditional differentiation methodologies by directly targeting gene networks and the epigenome[113,114]. By leveraging endogenous gene regulatory networks, this methodology allows for deliberate activation and repression of lineage-specific genes, offering researchers a precisely controlled and alternative strategy for differentiation[115].

1.2 The Biological Challenge: Modelling the Blood-brain Barrier (BBB)

1.2.1 The Blood-brain Barrier

The blood–brain barrier (BBB) is a tightly regulated vascular interface between the central nervous system (CNS) and the systemic circulation[116]. It is formed by highly specialized brain microvascular endothelial cells (BEC) characterized by continuous tight junctions, reduced transcellular and paracellular permeability, suppressed vesicular transport, and polarized transporter and receptor expression, which together maintain CNS homeostasis and restrict the entry of potentially harmful blood-borne substances into the brain[116]. Importantly, the BBB is a dynamic and adaptive interface whose properties are modulated by development, aging, disease state, and microenvironmental cues. Dysfunction of the BBB is implicated in a wide range of neurological disorders, including Alzheimer's disease, Parkinson’s disease, multiple sclerosis, and stroke[117–120], highlighting its essential role in maintaining CNS homeostasis and contributing to disease pathogenesis. Given its complexity and selective permeability, accurately modelling the BBB *in vitro* has become

a key objective in neurovascular research, not only to advance our understanding of BBB physiology but also to provide reliable platforms for CNS drug discovery. These models are integral to the preclinical evaluation of neurotherapeutics, enabling the assessment of drug permeability, transporter interactions, and barrier disruption in human-relevant contexts, and thus serving a vital role in improving translational success in the development of CNS-targeted therapies.

1.2.2 The Structure and Components of the Neurovascular Unit

Often described as “one of the most complex and metabolically active organs of the body”[121], the neurovascular unit (NVU) constitutes the structural and functional foundations of the BBB, defined by a highly specialized and dynamic microenvironment. The complex NVU is characterized by its constant intercellular communication between both cellular and non-cellular components, including brain endothelial cells (BECs), neurons, astrocytes, mural cells (pericytes and vascular smooth muscle cells), immune cells, and the extracellular matrix of the basal lamina[121,122]. The precise interactions that comprise the NVU primarily oversee signalling pathways, gene expression, cellular maturation, formation of new tissues, and neurovascular coupling, while also dictating the regulation and structure of the semipermeable membrane of the BBB[121,122].

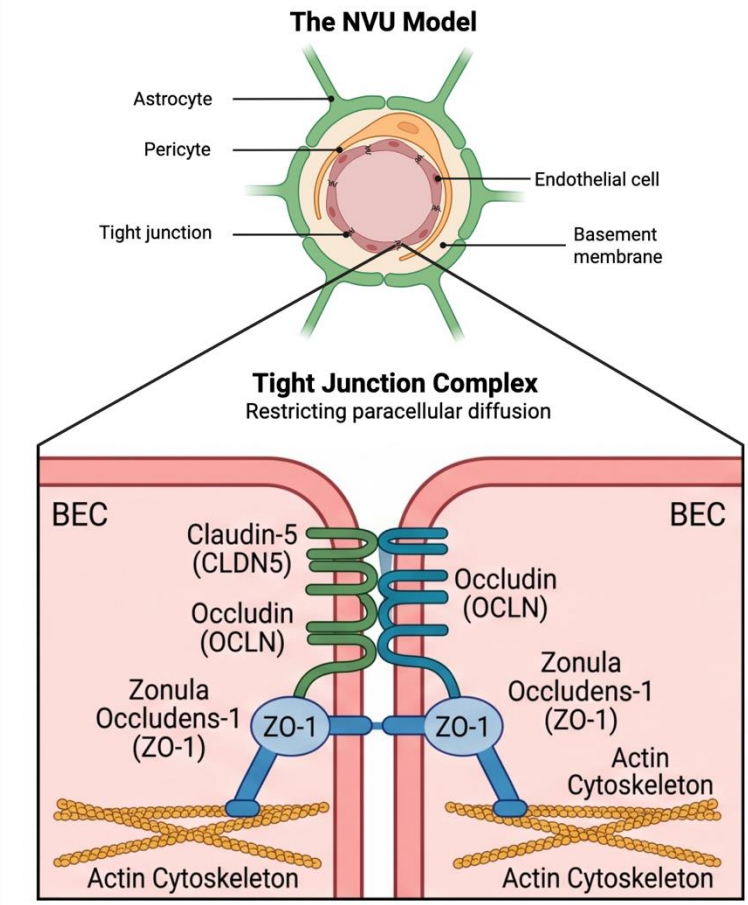


Figure 1.2: Schematic of the Neurovascular Unit (NVU) and the Blood-brain Barrier Tight Junction Complex. (Top) The NVU model reveals a cross-section of a brain capillary consisting of a wall of brain endothelial layer supported externally by pericytes and astrocytic endfeet. All components are surrounded and embedded within the basement membrane. (Bottom) Magnified view of the tight junction complex between two brain endothelial cells (BECs). This complex is critical to maintain a paracellular barrier. Certain key components include transmembrane proteins CLDN5 and OCLN, which form a seal. The underlying cytoskeleton anchors the tight junction complex and provides structural integrity.

1.2.3 Brain Endothelial Cells

BECs are the structural and foundational units of the BBB, acting as the physical interface separating the central nervous system (CNS) from the systemic capillary circulation[123]. Unlike other endothelial counterparts, BECs are uniquely specialized to stringently

regulate the bidirectional transport of molecules and factors between the blood and the brain[124]. A defining feature of BECs is the formation of hallmark tight junctions comprised of transmembrane proteins such as Claudin-5 (CLDN5), and Occludin (OCLN), anchored by intracellular actin cytoskeleton scaffolding proteins such as Zonula Occludens-1 (ZO-1) to restrict paracellular diffusion[116,125,126]. To regulate the brain metabolic requirements, BECs possess unique transport systems. For example, essential nutrients can cross the BBB through solute carrier (SLC) transporters including glucose transporter 1 (GLUT1) and monocarboxylate transporter 1 (MCT1)[116,127,128]. Furthermore, BECs also express efflux transporters such as ATP-binding cassettes (ABC), breast cancer resistance protein (BCRP) and multidrug resistance-associated proteins (MRPs) to prevent neurotoxins, foreign factors, and therapeutics from crossing into the brain[129,130]. Additionally, BECs facilitate the passage of cationic molecules through electrostatic interactions with negatively charged endothelial glycocalyx with adsorptive-mediated transcytosis (AMT)[127,131,132]. Whereas larger ligands and macromolecules can be transported across the endothelium through receptor-mediated transcytosis (RMT) such as insulin and transferrin[125,133,134].

1.2.4 Pericytes

Pericytes are mesenchymal-derived mural cells located in the basolateral side of the BBB and embedded within the endothelial basement membrane. Pericytes physically contact BECs, and through paracrine signalling they are able to induce and maintain barrier properties of the BECs. Specifically, expression of tight and adherens junctions are upregulated, restricting paracellular diffusion[135–137]. Beyond structural and barrier function, pericytes are reported to have roles in regulating angiogenesis[135,138], clearance of debris

and toxins[138,139], and modulating neuroinflammation[140,141]. Notably, pericytes express contractile proteins for neurovascular coupling to adjust capillary diameter and regulate the cerebral blood flow[135,142]. *In vitro* studies[143–147] have suggested that pericytes possess multipotent stem-cell-like capabilities to differentiate to smooth muscle, glial or neural cells; that capacity *in vivo* remains invalidated[148,149] in current literature.

1.2.5 Astrocytes

Astrocytes are specialized macroglial cells in the CNS primarily regulating homeostasis and neuronal activity. They also support the metabolic requirements of the CNS by supplying lipids, glucose and antioxidants from the blood to the BBB[150]. Within the NVU, astrocytic end-feet release paracrine signalling molecules that induce and further refine BEC tight junctions and transporters for maturation[151]. For the distribution of metabolites from the blood to other neurons in the NVU, astrocytes form intercellular networks connected by Connexin-43 gap junctions[152,153]. Through these channels, astrocytes facilitate transcellular transport of glucose and lactate over long distances from the BBB to neurons[154]. Moreover, astrocytic end-feet express polarized distribution of channels such as polarized aquaporin 4 (AQP4), and other potassium channels to maintain water, ions, and pH homeostasis across the BBB[155]. Furthermore, astrocytes have been speculated to have roles in NVU remodelling and endothelial progenitor cell recruitment for pathologies such as stroke[156].

1.2.6 Deficiencies in Conventional BEC *In Vitro* Models

A wide range of *in vitro* BBB models have been developed to support preclinical screening and prioritization of CNS-targeted drug candidates. Historically, most *in vitro* BBB models relied on primary BECs isolated from animal tissues (e.g., murine, bovine or porcine)[157–

[159]. However, evidence of cross-species differences in drug/antibody affinity towards BBB receptors and the expression of efflux transporters, has detrimentally compromised their translation to humans[160,161].

To better reflect human cerebrovascular biology, researchers transitioned to human-derived models such as biopsied brain tissues or immortalized BEC lines[157,162,163]. Although these sources have provided valuable insights into BBB function, they possess inherent limitations in drug transport and permeability studies. For example, primary human BECs are limited by tissue scarcity, poor scalability, and batch-to-batch variability. Notably, primary BECs undergo rapid dedifferentiation[164] *in vitro* when extracted from the neurovascular environment, resulting in compromised barrier fidelity and downregulation of hallmark BBB markers[157]. To mitigate issues for scalability, researchers generated immortalized human BEC lines such as the hCMEC/D3 line[163]. While this cell line offers consistency and proliferates indefinitely, immortalized cell lines undergo genetic drift and exhibit aberrant transporter expression[162,165,166]. They have been reported with high paracellular permeability and weak barrier fidelity, yielding transendothelial electrical resistance (TEER) values below 50 Ωcm^2 [157]. As a result, it is clear that both conventional primary human and immortalized BEC lines are unable to accurately model the human BBB *in vitro,* hindering biomedical research and clinical translations for complex neurological pathologies.

1.2.7 Bridging the Translational Gap in BBB Modelling: Human iPSC-derived Brain Endothelial-like Cells (iBECs)

Advances in human-induced pluripotent stem cell (iPSC) technologies have transformed BBB modelling by enabling the scalable and reproducible derivation of brain endothelial-like cells (iBECs)[167]. In particular, iPSC-derived BBB models offer substantial advantages

over other BEC sources due to their human donor origin, long-term genetic stability, self-renewal capacity, scalability, and ability to generate isogenic cellular components of the neurovascular unit (NVU). iPSC-derived BBB models exhibit hallmark features of specialized BECs, including high TEER, expression of BBB-specific tight junctions, lower caveolar-mediated transport and the expression of BBB-relevant influx and efflux transporters as well as receptor-mediated transcytosis (RMT) pathways[167–169]. As a result, these platforms have been increasingly integrated into CNS drug discovery workflows to interrogate how different therapeutic modalities, including small molecules, biologics, cell and gene therapies, neurotropic viruses, and lipid nanoparticle-based formulations, interact with the human BBB under both physiological and disease-relevant conditions[170–173]. Specifically, iPSC-derived BBB platforms have proven to be effective tools for screening and characterizing novel CNS targeting modalities. In the context of biologics, these models have successfully recapitulated antibody-triggered receptor-mediated transcytosis (RMT) crossing strategies targeting the transferrin receptor (TfR1)[174,175], insulin growth factor receptor 1 (IGF1R) and Transmembrane Protein 30A (TMEM30a)[169], showing predictive *in vitro–in vivo* correlations[169] and discriminating species differences in TfR1 crossing[176]. Similarly, in gene therapy applications, syngeneic NVU models are used to screen libraries of engineered AAV capsids to select variants with enhanced crossing efficiency, neuronal transduction profiles and species tropism[177–179]. Recent advancements have also extended to non-viral delivery systems that employ iPSC-derived BBB models to evaluate the transport efficacy of lipid nanoparticles (LNPs) and other polymeric carriers for optimized delivery of mRNA and other transgenes to the CNS[180–182].

Beyond delivery, these platforms have emerged as robust disease models for investigating BBB dysfunction induced by infectious and immune-mediated insults[183–185]. iPSC-derived BBB models have been used to study neurotropic viral infections[186], including Zika virus[186,187], West Nile virus[186], Japanese encephalitis virus[186], alphaviruses and flaviviruses[186], revealing mechanisms of brain endothelial barrier disruption, inflammatory activation, and neurotoxicity. In addition, these systems have been applied to model CNS-associated complications of non-neurotropic pathogens, such as SARS-CoV-2[188–190], where endothelial inflammation and barrier destabilization predominate, as well as cerebral malaria caused by *Plasmodium falciparum*[191], which induces BBB breakdown through parasite-mediated endothelial sequestration rather than direct neural infection. Beyond infectious disease, iPSC-derived BBB models are increasingly being applied to evaluate the safety, trafficking, and anti-tumour activity of immunotherapies, including chimeric antigen receptor (CAR)-T and CAR-NK cell–based therapies[170,192,193]. These platforms enable controlled assessment of immune cell adhesion, transendothelial migration, and cytokine-induced inflammatory responses and barrier perturbation, providing a human-relevant framework to investigate off-target neurotoxicity associated with cytokine release syndrome (CRS) and immune effector cell–associated neurotoxicity syndrome (ICANS)[194–196].

Recent studies demonstrate that BECs can reactivate developmental pathways, immune responses, and angiogenesis pathways in either vascular pathologies[197] or neuroinflammatory CNS diseases[198]. Importantly, the use of patient-derived and genetically engineered iPSCs extends these platforms to rare and monogenic neurological disorders, where disease-specific alterations in BBB integrity, transporter

function and genotype–phenotype relationships can be directly investigated in a human disease-relevant model[180,199,200]. iBECs derived from patient-iPSC lines harbouring neurodegenerative disease-associated mutations also provide a human-relevant platform for investigating how Alzheimer's disease, Parkinson's disease, ALS, and Huntington's disease impair barrier integrity, transport function, and neurovascular signalling[201–204]. These models are particularly valuable for dissecting patient-specific mechanisms of neurovascular dysfunction, including altered tight junction organization, inflammatory activation, mitochondrial dysfunction, and impaired transport or clearance of neurotoxic protein[205–211]. Lastly, the use of iPSC lines carrying *BRCA1* mutations has also been used to investigate how hereditary cancer-associated variants may influence BBB biology. These models may help determine whether BRCA1-linked genetic susceptibility contributes to BBB dysfunction and elevated risk of brain metastasis[212]. Integration of iPSC-derived BBB systems into CNS discovery and development workflows therefore supports early de-risking and optimization of therapeutic strategies, enhances translational predictability, reduces late-stage attrition, and accelerates the development of CNS therapeutics[213].

1.2.8 Defining the Gap: Current Challenges with iBEC BBB Modelling

While iPSC-derived BBB models have proven to be invaluable for CNS drug discovery, further refinement of the BBB phenotype and genotype *in vitro,* particularly in achieving adult-like BEC identity and stable long-term functional maturity, is still needed. Current iPSC-derived BBB models often display incomplete maturation, primarily because conventional differentiation strategies do not fully replicate the complex and dynamic microenvironment that instructs BBB development and specialization *in vivo*[214–216]. During

development, BECs acquire BBB identity through tightly coordinated and temporally regulated cues, including canonical *WNT* and *Notch* signalling that initiate BBB specification[217,218], with epigenetic regulators such as HDAC2 and PRC2 helping to establish the BBB by controlling endothelial gene expression and shaping the response to *WNT*/β-catenin signalling during development, while in adult CNS endothelial cells, these same mechanisms suppress developmental *WNT* target genes to preserve barrier homeostasis[219]. Reciprocal interactions with pericytes and astrocytes that reinforce barrier integrity and endothelial quiescence[220,221], dynamic extracellular matrix remodelling that shapes junctional organization and mechanotransduction[222,223], metabolic cues that support specialized transport and energy demands[224], and hemodynamic shear stress generated by blood flow that promotes endothelial maturation and functional stabilization[225]. Most *in vitro* iPSC BBB differentiation systems capture only a fraction of these stimuli, and as a result, iBECs frequently show partial and heterogeneous brain endothelial transcriptional signatures and incomplete expression of endothelial-specific attributes, such as immune adhesion molecules[226,227] and VEGF-induced angiogenic responses[228,229](**Table 1**). Furthermore, the very high TEER values reported in some iBEC models have been attributed to the acquisition of epithelial-like characteristics rather than a bona fide endothelial BBB phenotype[230]. Nishihara *et al*. describe an alternative differentiation strategy that yields iBECs with a mature immune adhesion molecule profile, enabling robust modelling of immune cell–BBB interactions and more closely recapitulating endothelial identity[231]. However, these cells do not achieve the high TEER values characteristic of conventional iBEC models, underscoring the need for further refinement of differentiation strategies that unify authentic endothelial

identity with high-fidelity BBB barrier function. In part, this limitation reflects the absence of an instructive CNS microenvironment, which not only constrains iBEC maturation but also promotes cellular heterogeneity and incomplete lineage specification, ultimately resulting in the persistence of fetal-like endothelial phenotypes[214,229,232–234]. Although the molecular mechanisms underlying BEC specification have been extensively characterized *in vivo*[235], comparatively little is known about how these cues can be induced in human iPSC-based differentiation protocols *in vitro*. This knowledge gap limits the ability to faithfully recapitulate BEC identity, function, and signalling, representing a key bottleneck in the rational optimization of iBEC differentiation protocols. Although key pathways such as *Wnt*/*β-catenin*[236,237] and retinoic acid signalling[220,238–240] have been identified as important regulators of BEC specification *in vitro*, the full gene regulatory network that distinguishes BECs from peripheral endothelial cell lineages in iPSC models remains undefined[241]. **Table 1** summarizes how current iPSC-derived BECs recapitulate key features of the *in vivo* BBB.

Table 1.1: Comparison of iPSC-derived brain endothelial cell (iBEC) as effective blood–brain barrier (BBB) models. Models are evaluated based on key BBB phenotypic and functional features, including tight junctions (TJ, CLDN5, OCLN, ZO-1), adherens junctions (AJ, CDH5), efflux transporter expression and activity (ABCB1/P-gp, MRPs), receptor-mediated transcytosis (RMT: targeting TfR1, LRP-1, IGF1R, TMEM30A receptors), and inflammatory leukocyte adhesion responses.

Study	TJ/AJ Expression	Efflux Transporter	Angiogenic Potential	Small Molecule Transport	RMT	Leukocyte Adhesion Response	Other[a]	Max TEER ($\Omega \cdot cm^2$)	BBB Fidelity Score[b]
Lippmann et al. (2012)[167]	●	●	-	●	●	-	-	Mono. ~ 222 Co. ~ 1450	4
Lippmann et al. (2014)[168]	●	●	-	●	-	-	-	Mono. ~ 2940 Co. ~ 5350	3
Wilson et al. (2015)[242]	●	●	-	-	-	-	-	Mono. ~ 2000	2
Katt et al. (2016)	●	●	-	●	-	-	-	Mono. ~ 1920	3
Hollmann et al. (2017)[243]	●	●	-	●	-	-	-	Mono. ~ 4678 Co. ~ 6635	3
Canfield et al. (2017)[221]	●	●	-	●	●	-	-	Mono. ~ 153 Co. ~ 886	4
Ribecco-Lutkiewicz M et al. (2018)[169]	●	●	-	●	●	-	-	Mono. ~ 500 Co. ~ 5000	4
Neal et al. (2019)[244]	●	●	-	●	-	-	-	Mono. ~ 8000 Co. ~ 10500	3
Park et al. (2019)[245]	●	●	-	●	●	-	●	Co. ~ 25000	5
Praca et al. (2019)[246]	●	●	-	●	●	-	-	Mono. ~ 60 Co. ~ 60	4
Linville et al. (2019)[247]	●	●	-	●	-	●	-	Mono. ~2260	4
Faley et al (2019)[248]	●	●	●	●	-	-	-	Mono. ~ 600	4
Nishihara et al. (2020)[231]	●	●	-	●	-	●	-	Mono ~ 60 Co. ~ 60	4
Selvakumaran et al. (2023)[249]	●	-	-	●	-	-	●	Mono. ~ 6000	3
Ozgür et al. (2023)[250]	●	●	-	●	●	-	-	Mono. ~5474	4
González-Gallego et al. (2025)[180]	●	●	●	●	●	-	-	Mono. ~100	5

a Other section features analysis and phenotypes not conventionally tested (e.g., Antibodies, LDL uptake, crossing efficiencies, etc.)

b The BBB fidelity score represents a qualitative combined score of listed features out of 7 and is intended for a comparative overview rather than a quantitative assessment of purity or functional magnitude.

1.3 The Engineering Solution for iBEC Generation: Programmable Transcriptional Modulation

1.3.1 Shifting Away from Traditional Methods for Cell Fate Specification

Historically, the conventional differentiation scheme relied on extrinsic signalling cues and molecules to guide pluripotent cells through each developmental stage. This approach is often limited by the stochasticity and heterogeneity in the resulting populations and requires extensive optimization of the differentiation pipeline. The current field is increasingly investigating leveraging programmable transcriptional modulation as an alternative method for programming cell fate. This engineering-centred strategy uses intrinsic gene regulatory networks to write and guide cell identity through differentiation. By directly targeting master regulators of endothelial phenotype, transcription-based reprogramming offers promise in not only refining the differentiation scheme but also for direct lineage conversions.

1.3.2 Transcription Factor-guided Approaches to Lineage Programming with Directed Differentiation and Transdifferentiation Strategies

Despite the development and refinement of diverse iBEC differentiation strategies[251], iPSC-derived BBB models continue to exhibit incomplete and variable brain endothelial specification. This limitation was systematically highlighted by Lu *et al*. through a comprehensive transcriptomic meta-analysis of previously published iPSC-derived BECs, which revealed that iBECs lack key functional endothelial attributes[214]. In particular, these cells exhibit reduced expression of core vascular lineage genes, including *KDR, VWF, ERG, TAL1, CLDN5, SOX18, SOX17, ESAM, S1PR1,* and *PECAM1,* while retaining gene clusters associated with a neuroectodermal epithelial identity such as *TRPV6, CDH1, CDH3, EPCAM, CLDN4, CLDN6, FREM2, ELF3, ESRP1,* and *ERBB3*.

To address these challenges, recent studies have pivoted towards forward transcriptional reprogramming strategies, or directed differentiation, in which lineage-defining transcription factors are expressed in iPSCs to more precisely induce BEC differentiation. These directed reprogramming or directed differentiation strategies focus on modulating the expression of transcription factors in iPSCs that control developmental programs governing lineage commitment, specification, and maturation[252,253]. Unlike conventional BEC differentiation protocols, which depend on defined media formulations, growth factor supplementation, and extracellular matrix cues, directed reprogramming and differentiation approaches primarily leverage activating gene regulatory networks to initiate and drive differentiation.

This conceptual shift builds on early demonstrations that endothelial identity can be induced through transcription factor–driven programs, as exemplified by initial transdifferentiation studies, defined as the direct conversion of one somatic cell type into another without passing through a pluripotent intermediate, in which overexpression of endothelial-specific transcription factors (such as *ETV2, FLI1, ERG*) was sufficient to directly convert somatic cells, such as amniotic cells[254] and dermal fibroblasts[255,256] , into endothelial cells. These strategies were further refined to use other cell lines, such as adipose-derived stem cells and adult fibroblasts, revealing the implicit role of *ETV2* as a master pioneer factor of broad endothelial lineage commitment[255,257].

In addition, extending these directed differentiation transcription factor–based principles to iPSCs leverages the scalability and high amenability of iPSCs to genome engineering to enable efficient lineage programming. For instance, Ding *et al*.[258] generated iPSC-derived *ETV2*-induced endothelial cells (*iETV2* ECs) that exhibited key endothelial

functions, including angiogenic potential, LDL uptake, and responsiveness to inflammatory cytokines. Transcriptomic analyses revealed that their gene expression closely resembled that of conventionally differentiated endothelial cells. Importantly, subsequent work has shown that activation of *Wnt* signalling in hPSC-derived naïve endothelial progenitors, but not in more mature endothelial cells, drives robust acquisition of canonical BBB traits, including increased *GLUT1* and *CLDN5* expression, reduced Plasmalemma vesicle-associated protein (PLVAP), and decreased permeability[236]. RNA-seq further demonstrated induction of a CNS-like endothelial transcriptional program, including *Wnt*-responsive upregulation of *LEF1*, *APCDD1*, and *ZIC3*[236]. Consistent with this developmental-stage-dependent competence, treatment of iETV2 ECs with a *Wnt* agonist in combination with TGFβ inhibition promoted the acquisition of brain endothelial-like features[258]. However, the reported TEER values remained lower than those achieved with previously established BBB differentiation protocols, and comprehensive characterization of BBB-specific properties was limited. Using a similar strategy, Lu *et al.* employed key endothelial transcription factors (*ETV2, ERG, and FLI1*) to initiate endothelial specification, generating cells that aligned with bona fide endothelium at the transcriptomic level and in functional assays; however, barrier induction, as measured by TEER, remained comparatively low[214]. These findings highlight that activation of generic endothelial programs alone does not induce acquisition of BBB-specific properties, yielding endothelial phenotypes that more closely resemble peripheral rather than brain endothelium.

1.3.3 Inducing BEC Identity Through Lineage-Specific Transcription Factors

Building on these advances, only a few studies have deliberately applied directed reprogramming strategies to induce BEC-specific phenotypes and BBB-relevant molecular and functional properties. In these studies, BEC-specific transcription factors such as *SOX18* and, more recently, *FOXF2* and *ZIC3,* were selected based on their established roles in cerebrovascular development and BBB maturation[240,259,260]. *SOX18* can upregulate *Wnt*-related signalling and contribute to endothelial lineage stabilization and regulation of junctional programs that control barrier formation [260–262], while *FOXF2* is critical to maintaining BBB tight junction integrity and low transcytosis[240,263,264]. *FOXF2* and *ZIC3*, acting downstream of canonical *Wnt*/*β-catenin* signalling[265], are upregulated during development, suggesting a potential role in BBB establishment[240,265].

In an effort to identify BBB-instructive transcription factors, 3 candidate regulators, including *ZIC3A*, *ZIC3B*, and *SOX18,* were first transduced in HUVECs[266]. Whereas *ZIC3A* and *ZIC3B* primarily increased *PLVAP* expression, *SOX18* overexpression induced both P-gp and PLVAP in a pattern more closely resembling that of brain microvascular endothelial cells (BMECs), suggesting a greater capacity to promote BBB-like endothelial features[266]. Given that *SOX18* is transiently expressed in the endothelial compartment of developing blood vessels and is linked to canonical Wnt signalling[267,268], its transient induction is developmentally relevant for promoting BBB maturation in iPSC-derived models. Building on this, a doxycycline (Dox)-inducible *SOX18* iPSC system was used to transiently express *SOX18* during BEC differentiation, resulting in higher TEER together with increased expression of endothelial markers (*CDH5* and *PECAM1*) and BBB-associated transporters (*ABCB1 and MFSD2A*)[266]. These findings support *SOX18*

as a promising developmental regulator for enhancing BBB maturation and functional barrier properties in iPSC-derived BEC models.

1.3.4 Context-dependent effects of FOXF2 and ZIC3 reprogramming

Although *ZIC3* alone did not produce a clear BBB-inductive effect in HUVECs, its established enrichment in BBB endothelium, regulation by *β*-catenin signalling, and predicted role in broader BBB gene regulatory networks[263,269,270] suggest that it may still contribute to BBB specification in more developmentally relevant iBEC differentiation contexts. Since *FOXF2* is also a key developmental regulator of BBB maturation and maintenance[234,264], examining these factors together provided a logical strategy to test whether combinatorial programming could more effectively reinforce brain endothelial identity than either factor alone. Lentiviral induction of *FOXF2* and *ZIC3* in iPSC-derived iBECs resulted in higher levels of key BBB transcripts and proteins compared with primary HBMECs[234]. They consistently exhibit higher TEER, lower transcellular transport, and equivalent efflux transport compared with HBMECs at the functional level, and respond to angiogenic stimuli, showing that these factors can induce brain-specific EC identity[234]. However, transcriptomic profiling of these iBECs confirmed a developmental brain endothelial identity resembling that of fetal BECs. Furthermore, overexpression of either *FOXF2* or *ZIC3* through adenoviral transduction in generic ECs showed that *FOXF2* enhanced some BBB-associated properties but also induced inflammatory gene signatures[260,271], whereas *ZIC3* overexpression had no measurable effect on barrier properties[260]. However, simultaneous overexpression of *FOXF2* and *ZIC3* in iPSC-derived iBECs did not induce inflammatory signatures and instead promoted BBB-relevant molecular and functional properties, suggesting that the differentiation protocol

can strongly influence the trajectory and outcome of directed reprogramming strategies. Lastly, the overexpression of *FOXF2* and *ZIC3* individually in human umbilical vein endothelial cells (HUVECs) increased the expression of some BBB-specific genes such as *ABCB1*, *SLC02B1* and *TNFRSF19*[263]. Therefore, *FOXF2* and *ZIC3* reprogramming appears to be more effective in ECs that already exhibit partial BBB-like features yet retain sufficient plasticity to acquire a more brain endothelial fate[234]. This may be particularly relevant for *ZIC3* activity, as its ability to induce BBB programs likely depends on whether the target cells are already primed to interpret BBB-specifying signals. Overall, these findings support the idea that successful directed reprogramming toward a BBB fate depends not only on transcription factor choice but also on the developmental context and baseline identity of the endothelial cells being programmed.

Consistent with the importance of *FOXF2* in BBB regulation, *FOXF2*-deficient iPSC-derived iBECs exhibited hallmark features of barrier dysfunction, including impaired cell adhesion, disrupted junctional integrity, and dysregulated endocytosis-associated pathways[180]. In a 3D human BBB model, LNP-mediated delivery of *FOXF2* mRNA largely rescued barrier impairment in *FOXF2*-knockout cultures, underscoring the essential role of *FOXF2* in BBB homeostasis while also highlighting the promise of LNP-based mRNA delivery as a therapeutic strategy and a potential platform for directed transient reprogramming approaches[180].

1.3.5 Small-molecule pathway modulation as a complementary strategy

Recent work has shown that small-molecule cocktail–based approaches can synergistically enhance BBB specification through coordinated pathway modulation. In particular, simultaneous activation of cAMP and *Wnt*/β-catenin signalling combined with

inhibition of TGF-β robustly induces BBB properties in iBECs[272]. This small-molecule cocktail, termed cARLA, promotes Claudin-5–dependent barrier tightening, shifts transcriptional programs toward a more *in vivo* BEC signature, and improves functional attributes including TEER and efflux transporter activity[272]. When combined with transcription factor-based reprogramming, these pathway-level cues may offer complementary means to improve not only BBB-associated functional properties but also the persistent gap in transcriptomic fidelity that remains a major limitation of iPSC-derived BBB models.

1.3.6 Combinatorial transcription factor programs for BEC fate control

In this context, increasing attention has turned to identifying the master and pioneer transcriptional regulators that more directly specify BEC identity and drive a more faithful brain endothelial gene expression program. To gain more insights into how and which TFs direct BBB formation, Roudnicky *et al* identified 17 candidate TFs that could direct naïve ECs into BBB-like cells[260]. Transducing combinations of these TFs in naïve iPSC-derived ECs identified a network of TFs (ETS1, SOX7, SOX18 and either TAL or LEF1) that significantly induced barrier resistance and induced expression of tight junction proteins and transporters characteristic of BBB ECs[260].

These findings helped lay the groundwork for subsequent efforts to define the master and pioneer transcriptional regulators that more directly specify BEC identity within iPSC differentiation programs. By integrating transient endothelial induction with large-scale transcriptomic meta-analysis and single-cell profiling of barcoded transcription factor combinations, Ramezankhahni *et al.* identified 12 candidate TFs that might fate map immature ECs towards BECs[273]. Following combinatorial overexpression of these 12 TFs,

a core set of regulators, *ZIC3*, *TFAP2C*, *TFAP2A*, and *DLX2*, was identified that drives immature iPSC-derived ECs toward cells with transcriptional signatures resembling mid-gestational human BECs[273].

More recently, a transcriptomics-guided forward-programming study in iPSC-derived endothelial progenitor cells showed that lentiviral expression of multiple BBB-enriched TFs, including *KLF2, KLF4, FOXF1, FOXF2, ZIC2, ZIC3, NR4A1, NR4A2, FOXC1*, and *FOXQ1*, each induced distinct but overlapping subsets of BBB-associated genes rather than a complete BBB program on their own[274]. *KLF2* and *KLF4* emerged as particularly strong inducers of BBB-like transcriptional programs and also upregulated other BBB-enriched TFs such as *FOXF2, FOXQ1, LEF1*, and *ZIC3*, suggesting that they may function upstream in the regulatory hierarchy controlling BBB specification. Importantly, *KLF2* and *KLF4* acted synergistically with *Wnt* activation, a response associated with a shift from inhibitory *TCF7L1/TCF7L2* expression toward *LEF1/TCF7*-driven canonical *Wnt* signalling, thereby reinforcing developmental BBB gene programs. Other TFs appeared to control more specialized BBB modules: *FOXF1* and *FOXF2* preferentially induced transporter and lipid-metabolism genes; *FOXQ1* and *FOXC1* reduced caveolin-1 expression, consistent with suppression of vesicular trafficking; *ZIC2* and *ZIC3* enhanced expression of several BBB-enriched solute carriers; and *NR4A1/NR4A2* modestly promoted transporter-associated signatures. Notably, simultaneous combinatorial expression of *KLF2* or *KLF4* with *FOXF2, ZIC3, NR4A2*, and *FOXQ1* enhanced expression of the majority of assessed BBB-enriched genes, increased tight junction protein expression, reduced endocytosis, improved P-gp and BCRP efflux activity, and modestly strengthened barrier integrity. These effects also extended beyond iPSC-

derived iBECs, as the same TF combinations partially induced BBB-like features in HUVECs and rescued aspects of tight junction gene expression in hCMEC/D3 cells, further supporting the idea that combinatorial TF programming can impose core elements of BBB phenotype across endothelial backgrounds.

Together, these findings underscore the growing potential of combinatorial transcription factor programming as a strategy to reconstruct core BBB features, while also highlighting the need to more precisely define the developmental regulators that establish and stabilize *bona fide* brain endothelial identity. Emerging literature, particularly large-scale developmental single-cell and spatial transcriptomic BEC atlases[234,263,273,275], will be instrumental in identifying unique BEC pioneer and master transcription factors that govern BEC identity, expanding the repertoire of putative targets for efficient directed reprogramming[197,263,273,275].

1.3.7 Programmable Control of Brain Endothelial Specification and Maturation

Collectively, these described factors represent promising synergistic and additive targets for further exploitation in directed differentiation strategies. Given the highly specialized nature of BECs and the strong influence of the microenvironment on BBB phenotype, it is unlikely that a single pioneer or master transcription factor is sufficient to fully specify brain endothelial identity. Instead, combinatorial and multiplex expression strategies leveraging these targets will likely be required to more faithfully recapitulate the developmental window and coordinated gene regulatory programs that underlie BBB specification, maturation, and stabilization. Such approaches may also help define the minimal transcription factor combinations, timing, and dosage needed to achieve a transcriptome recapitulating a mature adult BBB phenotype. Lastly, differences in the

starting cell state (e.g., iPSCs, naïve ECs, or endothelial progenitors) are likely to influence the outcome of forward reprogramming strategies, as the developmental maturity of the input population can shape epigenetic accessibility and thereby alter responsiveness to BBB-inductive transcriptional programs. Furthermore, inherent differences between iPSC lines, including donor-specific genetic variation, residual epigenetic memory, and differences in baseline developmental priming, may bias differentiation trajectories and alter the efficiency and fidelity of BBB-directed reprogramming. Collectively, these studies demonstrate the feasibility of directed reprogramming to guide iPSC differentiation toward brain endothelial–like states (**Table 2**) and establish a foundation for applying genome engineering approaches to prime BEC lineage commitment, maturation, and functional fidelity through cell-autonomous control.

1.3.8 Temporal Control of Transcription Factor Expression in BBB-Directed Reprogramming

A key caveat to the current direct reprogramming studies is that they have relied on viral transduction driven by constitutive promoters, an approach that, although useful for initial screening and factor discovery, is less well suited for controlled, stage-specific directed differentiation. Constitutive transcription factor overexpression through integrating strategies can disrupt temporal hierarchy, leading to inappropriate persistence of early developmental programs[276], interference with later maturation cues, and incomplete integration of extrinsic signals such as *Wnt*/β-catenin activation, shear stress, and NVU interactions. As a result, prolonged transcription factor activity may constrain endothelial plasticity, disrupt the temporal chromatin remodelling and gene regulatory transitions required for maturation and limit acquisition of adult BBB features, and ultimately compromise barrier fidelity and long-term functional stability.

To circumvent these issues, Dox-inducible *ETV2* systems allow more temporal control over the mesoderm-to-endothelial transition, so that *ETV2* expression can be switched off once endothelial identity is established. This design more closely recapitulates embryonic vasculogenesis, where *ETV2* is expressed transiently to initiate endothelial commitment and is subsequently downregulated to allow vascular maturation[277]. Transient *ETV2* expression may be particularly important in iPSC-BBB differentiation protocols because it can strengthen early endothelial lineage commitment while allowing timely downregulation for subsequent BBB-specifying programs to emerge[277]. After transient *ETV2* expression, a secondary endothelial maintenance program involving *ERG, FLI1, SOX7, SOX17, TAL1, ETS1, KLF2*, and *FOXO1* must normally consolidate vascular identity before BBB-specific signals such as Wnt/β-catenin, *FOXF2, ZIC3*, and *SOX18* can refine the phenotype toward brain endothelium; failure to fully establish this intermediate endothelial stabilization phase may be one reason current iBEC protocols continue to show weak endothelial identity and persistent epithelial-like features. This may help address a major limitation of current iBEC models by more effectively establishing the core vascular transcriptional program before introducing cues that promote brain endothelial maturation. These limitations may now be more directly addressed through directed differentiation strategies, which offer the ability to install early endothelial commitment programs in a more controlled manner and then sequentially activate the downstream regulators needed to stabilize vascular identity and promote BBB maturation. In this framework, transient induction of factors such as *ETV2* can be used to initiate endothelial specification, followed by stage-appropriate engagement of

endothelial maintenance and BBB-specifying programs that are often inadequately established in conventional iBEC protocols.

However, Dox-inducible systems are also known to exhibit basal leakiness[278], which can reduce temporal precision and complicate the interpretation of stage-specific transcription factor effects during differentiation. Emerging transient gene modulation strategies, such as LNP-mediated delivery of CRISPR-based differentiation programs, may offer an attractive alternative by enabling tighter, more tunable control over developmental gene expression without imposing sustained or constitutive activation[279]. Such approaches could be especially valuable for modelling BBB specification, where precise timing, dosage, and combinatorial activity of developmental regulators are likely critical for achieving a more faithful brain endothelial phenotype.

Table 2: Transdifferentiation and Directed Differentiation Approaches for Endothelial Cell Differentiation.

Study	Cell Source	Target Cell Type	Target TFs	Additional Cues	Reference
Transdifferentiation of somatic cells					
Ginsberg *et al.*, 2012	Human amniotic cell	Endothelial cell	*ETV2* *FLI1* *ERG1*	TGFβ suppression	254
Morita *et al.*, 2015	Human dermal fibroblasts	Endothelial cell	*ETV2*	-	255
Lee *et al.*, 2017	Human dermal fibroblasts	Endothelial cell	*ETV2*	VEGF supplementation	256
Cheng *et al.*, 2018	Human adipose-derived stem cells	Endothelial cell	*iETV2*	Dox induced ETV2 and TGFβ suppression	257

Chen *et al*., 2019	Human fibroblast	Endothelial cell	*DKK3*	-	[280]
Han *et al*., 2021	Adult human fibroblast	Endothelial cell	*ETV2, FLI1, ERG*	-	[281]
Directed Differentiation of iPSCs					
Wang *et al.,* 2020	iPSCs	Endothelial cell	*ETV2, SOX17*	-	[282]
Qian *et al*., 2017	iPSCs	Brain endothelial cell	*Wnt/β-catenin*	Used Wnt agonist and RA for maturation	[283]
Lu *et al*., 2021	iPSC	Brain endothelial cells	*ETV2 FLI1 ERG*	-	[214]
Zhang et al., 2022	iPSC	Endothelial cell	*ETV2*	-	[277]
Zhang *et al.,* 2023	iPSC	Brain endothelial cell	*iSOX18*	Dox-induced SOX18	[266]
Luo et al., 2025	iPSC	Endothelial cell	*iETV2*	Dox-induced ETV2	[284]
Ding et al., 2025	iPSC	Endothelial cell	*iETV2*	Dox-induced ETV2	[258]
Erzar *et al*., 2025	iPSC	Endothelial cell	*iETV2*	Dox- induced ETV2	[285]
Cui *et al*., 2026	iPSC	Brain endothelial cell	*FOXF2, ZIC3*	-	[234]
Zhang et al., 2026	iPSC	Endothelial cell	*iETV2*	Dox-induced ETV2	[286]
Tamhankar et al., 2026	iPSC	Brain endothelial cell	*KLF2, KLF4, FOXF1, FOXF2, ZIC2, ZIC3, NR4A1, NR4A2, FOXC1 or FOXQ1*	-	[274]

Hwang et al., 2026	iPSC	Brain endothelial cell	*ETV2, ERG, FLI1, CLDN5, PECAM1, CDH5, KDR*	-	[287]

1.4 The Bioengineering Toolkit: Genetic, Epigenetic and Transcriptomic Tools for Programmable Control

1.4.1 The Pioneer of Genomic Engineering Tools

The evolution of genomic engineering tools has advanced rapidly and has evolved from research into naturally occurring enzymes to developing and deploying programmable platforms for gene manipulation. Foundationally, this field was first established in the 1970s with the discovery of restriction endonucleases from *Haemophilus influenzae,* providing the first demonstration that proteins can recognize and cleave specific DNA sequences[288,289]. This laid the foundation for early molecular biology and recombinant DNA cloning; however, its implementation in cellular systems was still limited by poor translation into mammalian cell lines. Subsequently, the first example of genetic alterations in mammalian cells was done through homologous recombination (HR)[290,291]. However, spontaneous HR is markedly inefficient, occurring at a rate below 1 in 1,000,000[292]. A following study in the 1990s used the yeast meganuclease I-SceI and revealed that targeted double-stranded breaks (DSB) in the genome initiate and improve the HR efficiency drastically[293,294]. These early studies established the foundation and direction for modern genome engineering, enabling efficient and site-specific gene perturbation technologies.

The first breakthrough towards programmable genome engineering involved introducing DSBs using recombinant proteins. The first iteration was in the early 2000s with Zinc

Finger Nucleases (ZFNs)[295]. ZFNs are chimeric proteins with modifiable DNA-binding domains fused to a cleavage domain (FokI restriction endonuclease)[295]. The DNA-binding domain contains a motif that recognizes three base pairs of DNA, which can also be assembled in tandem motifs to target larger, highly specific genomic sequences. However, ZFNs require dimerization for their catalytic cleavage activity to occur[296,297]. As a result, this requires researchers to design and optimize two ZFNs that target both the top and bottom strands to dimerize the FokI endonuclease and enable efficient cleavage[297]. Moreover, ZFNs lack modularity, as the binding affinity and specificity of their motifs are highly context-dependent and easily influenced by the conformations of neighbouring ZFNs, hindering their function. As a result, although ZFNs were among the first iterations of a programmable genome-modification tool, they are laborious, unpredictable, and extremely expensive to engineer, limiting their broad applicability to more complex engineering needs.

Similar to ZFNs, Transcription Activator-Like Effector Nucleases (TALENs) are chimeric DNA-binding proteins that utilize FokI endonuclease to generate DSB[298]. Originally derived from *Xanthomonas* bacteria, TALENs consist of 33-35 amino acid repeats[299]. Researchers identified a critical programmable region within these repeats (amino acids 12 and 13) known as the repeat variable diresidue (RVD)[299]. Residue 12 acts as a stabilizer, forming hydrogen bonds with the DNA backbone, anchoring the protein. Residue 13 makes direct base-specific contact with the target nucleotide. Researchers can arrange these RVDs to target genomic sequences with the pre-characterized residue-based interactions (e.g., NI for adenine, HD for cytosine, NG for thymine, and NN for Guanine).

FokI catalyzes exclusively as a dimer[300], two TALEN monomers must be engineered to flank the sense and antisense strands of the target DNA. Each binding site also requires a thymine at the 5' position to anchor the protein, and the two binding sites are required to have a 12-21 base pair region. When these conditions are met, the TALEN arms bind, and FokI dimerizes to induce a DSB.

TALENs represent an early platform for genome engineering in site-directed mutagenesis; however, they pose certain challenges. First, TALENs are large and require two proteins per reaction, limiting viral delivery systems with restricted payload capacity (e.g., adeno-associated viruses, AAVs)[301,302]. Additionally, TALENs are limited in large-scale, high-throughput experiments by the complexity of protein engineering. For example, a large-scale project with 10 mutations requires researchers to build, clone, and individually validate over 20 unique repetitive TALEN plasmids[303,304].

1.4.2 Modern Era of Genome Engineering: An Overview of the CRISPR/Cas Toolkit

The Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and CRISPR-associated (Cas) proteins are derived from a natural adaptive immune response in bacteria[305]. The biological process involves Cas endonucleases guided by RNA to cleave viral DNA to prevent infection and integration into the host genome[305,306]. The discovery and characterization of this biological process led to a modern era of genome engineering, where researchers hijacked and re-engineered the system to a programmable tool for targeted mutagenesis[307,308].

The original CRISPR/Cas9 complex is comprised of a Cas9 endonuclease for enzymatic activity and a single guide RNA (sgRNA) for nucleotide base recognition[309]. The structure of the sgRNA is a fusion of two RNA molecules: CRISPR RNA (crRNA) containing a 20

bp nucleotide spacer sequence complementary to the target site, and a trans-activating crRNA (tracrRNA) containing a scaffold for the Cas9 protein to bind to[309]. A prerequisite for target recognition for the CRISPR complex is the protospacer adjacent motif (PAM), which is a short sequence (e.g., 5'-NGG-3' for *Streptococcus pyogenes* Cas9) downstream of the target sequence[310]. The PAM sequence can range from 2-6 bp, depending on the CRISPR class and type that is used, and acts as a checkpoint to prevent the cleaving of the self-genome.

The conventional mechanism for targeted cleavage of the DNA is enacted when the Cas9 recognizes and anchors to the PAM and unwinds the DNA duplex[311]. The crRNA spacer binds to the target strand and forms an R-loop structure. The hybridization of the sgRNA initiates a structural conformation change, activating the HNH domain and the RuvC domain, which cleaves both the complementary target strand and the non-complementary strand, respectively[309,312]. A resulting blunt DSB occurs 3 bp upstream of the PAM, and researchers can initiate homology-directed repair (HDR) or non-homologous end joining (NHEJ) using a template donor DNA for site-specific gene editing[306]. Currently, the scope of the CRISPR-based genome engineering toolkit continues to expand, with highly efficient base editing systems and programmable transcriptome modulators as examples of transformative capabilities that have been developed and validated (**Table 1.3**).

Table 1.3: Overview of CRISPR Classes and Types with Genome and RNA Engineering Capabilities. All types are organized by their class identity, targeting pre-requisites, target molecule, requirement of a tracrRNA, and their derived species. Class 1 CRISPRs typically contain multi-protein complexes fused or paired with a nuclease endonuclease, whereas Class 2 CRISPRs involve a single nuclease paired with an effector molecule. The anomaly Type III-E nuclease architecture bypasses the multi-protein complex formation as *D. ishimotonii* stitches the genes in one continuous sequence, mechanistically.

CRISPR Class	CRISPR Type	Nuclease Architecture	Targeting pre-requisites	Target	Requires tracrRNA	Derived Species
Class 1	Type I	Multi-protein complex and Cas 3[313]	PAM	DNA	No	*E. coli Cascade-Cas3*
	Type III-A/B	Multi-protein complex (Csm/Cmr) and Cas 10[314,315]	No PAM	DNA/RNA	No	*S. epidermidis Cas10-Csm*
	Type III-E	Single protein fusion (Cas7-11)[316]	No strict PFS requirement	RNA	No	*Desulfonema ishimotonii (DiCas7-11)*
Class 2	Type II	Cas9[306,309]	PAM (E.g., NGG)	dsDNA	Yes	*SpCas9 (S. pyogenes), SaCas9 (S. aureus)*
	Type V	Cas12a (cpf1)[317]	PAM (e.g., TTTV)	dsDNA	No	*AsCas12a (Acidaminococcus), LbCas12a (Lachnospiraceae)*
	Type V	Cas12j[318]	PAM (e.g., T-rich)	dsDNA	No	*CasΦ-2 (Biggiephage)*
	Type V	Cas14 (Cas12f)[319]	PAM (e.g., T-rich for dsDNA)	ssDNA/ dsDNA	Yes	*DPANN superphylum*
	Type VI	Cas13[320,321]	PFS required for some variants	RNA	No	*LwaCas13a (L. wadei), RfxCas13d*

1.4.2 Gene Editing Through Repair Pathways: Homology Directed Repair and Non-Homologous End Joining

The conventional CRISPR/Cas9 mechanism generates a blunt DSB, triggering the cell's endogenous repair pathways for genome repair and mutagenesis. NHEJ is an error-prone repair pathway that rapidly ligates broken ends of the genome, frequently resulting in unpredictable insertions or deletions (indels). This mechanism is mainly used to generate functional gene knockouts *in vitro* and *in vivo* for genome function studies. The second repair pathway, HDR, requires a DNA repair template containing homologous flanking regions to incorporate new sequences into the genome. However, HDR is inefficient and restricted to dividing cells in the S and G2 phases. While introducing DSB and endogenous repair pathways for genome engineering served as the first milestone, they often trigger a severe p53-mediated DNA damage response in the cell lines. As a result, relying on DSB-dependent mechanisms in more sensitive cell lines, such as hiPSCs and differentiating lineages, can often make genome engineering incredibly challenging.

1.4.3 Genome Editing Through Base Editors

To address the overreliance on DSB repair pathways, the second generation of CRISPR technologies led to fusing Cas proteins with Base Editors (BEs). To develop BEs, researchers fused a nickase Cas9 (nCas9, nicks only a single strand of DNA) or a dead Cas9 (dCas9, catalytically inactivated Cas9) to a deaminase domain. The resulting BE-Cas9 complex can alter a specific nucleotide within the targeted R-loop without inducing a DSB. The BE deaminase domain deaminates an amino group off the targeted nucleotide, facilitating conversion of base pairs by a downstream repair path. For example, Cytosine BE (CBE) converts cytosine to uracil, which the cell identifies as a mismatch U-G and "repairs" the mismatch by removing the nicked G strand. As the cell replicates,

thymine is replaced, transitioning the bp from a C-G to a T-A transition. Similarly, Adenine BE (ABE) converts adenine to Inosine, which is read as a guanine base, yielding an A-T to G-C transition. One major limitation of BEs is that they can only introduce transition mutations through deamination and cannot generate indels. Additionally, they are susceptible to bystander edits, in which unintended extra base-pair conversions can occur within the Cas9 catalytic window.

1.4.4 Genome Editing Through Prime Editing

To address this limitation of BE, prime editing (PE) was introduced in 2019 by Anzalone et al.[322] where an nCas9 was fused to a reverse transcriptase (RT) domain paired with an engineered prime editing guide RNA (pegRNA). The pegRNA contains an extended RNA template comprised of a primer binding site (PBS) and an RT template encoding the desired novel sequence[322]. Once genome nicking is achieved on the non-target strand, the liberated genome DNA hybridizes to the PBS. The hybridization enables the RT domain to catalyze and synthesize new sequence information directly on to the genomic locus through reverse transcription. The resulting product contains a 5' unedited flap and a newly edited 3' flap; the endogenous cellular machinery excises the unedited 5' flap allowing the edited strand to reincorporate[322]. This mechanism allows for 12 possible base pair substitutions as well as generating indels, without introducing DSBs and relying on DNA donor templates. PE drastically increases the versatility and functionality of genome engineering enabling precise genome editing. However, this technology is restricted to gene edits under 100 bp with decreasing efficiency as size increases[323]. Additionally, pegRNA requires extensive validation and empirical optimization of the

lengths of the PBS and RT template to achieve high efficiency optimization for efficiency[323].

1.4.5 Large Transgene Integration Using NHEJ Repair Pathway: Homology-Independent Targeted Integration

Another major advancement in the field of genome engineering, occurring in parallel to PE, is Homology-Independent Targeted Integration (HITI)[324]. HITI technology enables integration of large cassettes (approximately 8 kb) into the genome via NHEJ in post-mitotic, non-dividing cell lines[324]. This mechanism uses a donor plasmid flanked by Cas9 cleavage sites complementary to the genome. Cas9 cleaves both the genomic target in the donor plasmid and the genome, the NHEJ pathway then ligates the donor into the genome. The promise of HITI technology lies in how the cleavage sites function for orientation-locked integration. In the forward orientation integration, the Cas9 integration sites are destroyed, whereas in the reverse orientation, the Cas9 target site is repeatedly re-cleaved and ligated until the correct irreversible site is produced. HITI technology demonstrated that a large-scale transgene integration strategy for non-dividing cells was achievable. However, HITI still relies on DSB mechanisms, resulting in cell lines with genotoxic DSBs and off-target mutations[325,326].

1.4.6 Genome Editing Through Site-Specific Recombination

Site-specific recombination (SSR) leverages enzymes that recognize specific short nucleotide sequences and cleave the DNA to exchange and ligate the genome without relying on the DNA repair pathway[327,328]. Thus, SSRs offer a novel, efficient and DSB-free methodology that also pairs well with various cell types, including non-dividing cell lines. Bacteriophage-derived serine recombinases are enzymes that facilitate the transposition of genetic elements and have been widely adopted for performing site-specific

recombination. For example, phi31, Bxb1 and BceINTa are serine recombinases that catalyzes the phage attachment site (attP) to the bacterial attachment site (attB)[329–331]. This mechanism is an efficient and unidirectional reaction mimicking a cut-and-paste mechanism with resulting scars on the integrated genes (attL and attR). Researchers were able to adapt this innate system from bacteriophages and developed bioengineering tools to perform transposition by introducing attB landing pads and co-delivering serine recombinases with a donor transgene flanked by the attP recognition site[332]. Notably, researchers identified a key feature in the Bxb1 serine recombinase attachment sites, discovering that a 2 bp central dinucleotide sequence determines the orientation of the insertion[333]. From this, researchers characterized panels of optimal central dinucleotides combinations[331–334] to improve integration efficiency[335], while performing multiplexed integration applications[332,336].

1.4.7 DSB-free Integration of Large Transgenes: Programmable Addition via Site-Specific Targeting Elements

A state-of-the-art technology known as Programmable Addition via Site-Specific Targeting Elements (PASTE)[331] was developed to address the size constraints, efficiency and payload capacity limitations of PE. PASTE is the unification of two widely adopted cell engineering tools: PE and site-specific serine integrases. This system comprises a fusion protein containing an RT domain, a serine integrase (e.g., Bxb1 or BceINTa), and an nCas9. PASTE reactions occur with a highly coordinated, site-specific, DSB-free two-step mechanism to integrate synthetic transgenes into the genome. First, the protein complex uses a re-engineered pegRNA to assist the integration of a ~46 bp attB landing pad, called an attachment guide RNA (atgRNA). The nCas9 nicks the DNA, enabling reverse transcription of the template attB site to occur. PASTE leverages an enhanced version of

prime editing called PE3[322], which increases editing efficiency by introducing a secondary nick in the non-edited strand with an additional nCas9. Sequentially, the serine integrase recognizes the attB sequence and catalyzes a unidirectional recombination reaction when co-delivered with a donor cassette containing complementary attP sites. The promise of PASTE technology is that it allows for large transgenes to be integrated (reported up to 36 kilobases) while also minimizing the risks associated with conventional DSB repair mechanisms.

1.4.8 Transcriptional Modulation: CRISPRa and CRISPRi

CRISPR activation (CRISPRa) and CRISPR interference (CRISPRi) leverage catalytically inactivated Cas proteins to modulate the transcription of genes[337–339]. This process involves recruiting effector protein complexes and molecules to dCas9 or dCas12a[340], enabling downstream activation or repression of gene expression[341]. For example, first-generation CRISPRa systems directly fuse a single VP64 activation domain to dCas9[342,343]. Although promising, the single VP64 activation domain lacked consistent and potent activation responses across diverse conditions[98].

Second-generation activation systems evolved rapidly, reporting increased activation efficiency and consistency across different cell lines. For example, VPR is a tripartite fusion protein of three activation domains (VP64, p65, and Rta) tethered to the C-terminus of dCas9[98]. Comparatively, the Synergistic Activation Mediator (SAM) system alters the sgRNA scaffold with the addition of an MS2 RNA aptamer that recruits co-expressed activation molecules such as, MS2 coat protein (MCP) fused to p65 and heat shock factor 1 (HSF1)[344]. Additionally, the SunTag system fuses an array of GCN4 peptides to dCas9, which recruit co-expressed scFv-VP64 antibodies at a single promoter[345].

Benchmarking and characterization of the various CRISPRa systems revealed that recruiting multiple activation domains outperformed the first-generation systems[346]. Additionally, within the second-generation systems, the SAM system frequently resulted in the highest level of transcript induction[346]. However, based on the practical application, researchers are required to carefully consider the system choice. For example, SAM and SunTag require co-delivery of multiple molecular components, increasing payload size and limiting their implementation in viral delivery systems[347]. On the other hand, single-component systems such as VPR remain beneficial for *in vivo* applications due to their size and their stoichiometry to ensure 1:1 expression and function[348].

Conversely, while CRISPRa upregulates gene expression, the parallel development of CRISPRi enables the repression/silencing of gene expression without altering genomic sequences directly. First-generation CRISPRi technology pairs a dCas protein fused to the Krüppel-associated box (KRAB) domain, a transcriptional repressor[339,349]. The co-localization of the CRISPRi complex to a promoter or enhancer region initiates further recruitment of other co-repressors[349]. For example, KAP1 (TRIM28) coordinates heterochromatin formation[350] by recruiting histone deacetylase and histone methyltransferase SETDB1[351]. This results in chromatin condensing through H3K9 trimethylation, blocking transcriptional machinery at the targeted loci. Although promising, single-domain repressors have previously reported variable silencing efficiencies, prompting researchers to screen for optimized repressor domains[352]. Next-generation CRISPRi systems eventually included a Methyl-CpG binding protein 2 (MeCP2) fused to the KRAB domain, outperforming previous iterations for gene silencing when targeting endogenous targets[353].

Ultimately, CRISPRa and CRISPRi technology provide bioengineers with a versatile and modular method to regulate the transcriptome. Without relying on genomic sequence editing, these systems offer a safer mechanism to manipulate endogenous gene expression. Facile swapping of sgRNA sequences makes this platform inherently scalable and easy to use compared to other silencing/activating platforms. Moreover, one of the major advantages of CRISPRa/i systems is their capacity for multiplexing, enabling simultaneous orthogonal upregulation or suppression of transcriptional networks. This multiplexed toolkit is particularly attractive for regenerative medicine applications as it has the potential to rewire hiPSC transcriptional networks for lineage reprogramming without the need for genomic alterations.

1.4.9 Stable and Heritable Epigenome Regulation: CRISPRon and CRISPRoff

CRISPRa/i systems provide a method for transcriptome modulation, dependent on the repressor's efficacy and the consistent expression of the dCas9 fusion protein complex. However, these transient epigenetic modifiers can lose efficiency over longer periods of time due to silencing, and prolonged expression of Cas proteins is often associated with toxicity and immunogenicity. Thus, for cell fate engineering in regenerative medicine, the field relies on establishing a heritable transcriptional program to establish stable transcriptional modification.

This led to the development of CRISPRon and CRISPRoff technologies, enabling heritable stable epigenomic re-landscaping[354]. The CRISPRoff system is composed of a KRAB domain, DNA methyltransferase 3A (DNMT3A) and a catalytically inactive DNMT3L. CRISPRoff localizes to promoter regions to catalyze methylation on cytosine residues in CpG dinucleotides. As a result, heterochromatin formation occurs through

KRAB, while the endogenous methylation locks in this structural change into the epigenetic memory. Upon cell proliferation and division, the CpG dinucleotides that are methylated propagate, thus silencing the targeted gene throughout hundreds of cell divisions.

On the other hand, CRISPRon technology aims to erase epigenetic marks on these CpG dinucleotides[354,355]. This is mediated by a dCas9 fused to a Ten-Eleven Translocation 1 (TET1) domain that enables demethylation[356]. CRISPRon oxidizes 5mC into the 5-hydroxymethylcytosine and eventually results in cytosine by leveraging base repair pathways[357,358]. As a result, heterochromatin formed by methylation marks is relaxed, and allows for transcriptional machinery to access promoters and transcription start sites at the gene target.

1.4.10 RNA Regulation: Cas13 and CasRx

Apart from DNA targeting systems, researchers can also target RNA to modulate the translation and expression of proteins. Unlike Cas9, Cas13 systems[359] exclusively target single-stranded RNA (ssRNA). When paired with crRNA, Cas13 targets the target RNA molecule and undergoes a conformational change, activating Higher Eukaryotes and Prokaryotes Nucleotide-binding (HEPN) domains[320]. These domains act as an RNase and degrade the targeted RNA transcript. Of the Cas13 family, Cas13d, also denoted as CasRx, is small (~930 amino acids), allowing it to be packaged easily alongside a crRNA to a single AAV vector[360]. Moreover, CasRx has demonstrated highly efficient knockdown in human cell lines when compared to conventional RNA interference (RNAi) methods[361]. The sole advantage of utilizing CasRx is that this system does not rely on the conventional RISC/Dicer complex that RNAi systems use. Oftentimes, RISC complex formation and

Dicer function can cause cellular toxicity and off-target effects[362], which Cas13 can circumvent by avoiding the use of endogenous machinery. Recently, researchers have expanded the tool by deactivating Cas13 (dCas13) and fusing it to an RNA deaminase[363]. Researchers fused an Adenosine Deaminase Acting on RNA (ADAR2) to dCas13b, which converts adenine molecules to inosine, which is then translated as a guanine by the cell. By developing a programmable RNA editing tool, researchers can potentially reverse point mutation-based genetic diseases on the mRNA strand, avoiding risks associated with off-target editing caused by DSB DNA breaks.

1.5 Safe and Targeted Delivery of Transgenes to the Human Genome

1.5.1 Survey of Transgene Integration Technologies

As the bioengineering toolbox continues to advance, the stability of engineered cell lines is ultimately limited by where and how transgenes are integrated into the genome. Transgene integration can be performed in a multitude of ways that fall into three main categories, including: semi-random viral integration, semi-targeted transposon-mediated integration, and precision integration (programmable nucleases and recombinases).

Oftentimes, viral integration systems are described as “random”. However, in reality, these integration strategies point towards genomic biases that are innately selected by the viral integrase and host chromatin accessibility. Viral-based integration is reported to have high efficiency and is widely adopted, even achieving FDA approval for many gene therapy pipelines. For example, lentivirus derived from human immunodeficiency virus 1 (HIV-1) infects both dividing and non-dividing cells and uses reverse transcription to integrate into the host genome via viral integrase[364,365]. Similarly, γ-retroviruses are single-stranded RNA viruses that preferentially integrate proximal to transcription start

sites and active promoters[366–368]. Moreover, adeno-associated viruses (AAVs) transfer episomal DNA into the host nucleus for long-term transient expression of genes[369,370].

The major challenges associated with using viral packaging vectors include the risk of mutagenesis and genotoxicity to the host genome, which can disrupt host genes[371]. This is particularly problematic for spontaneous integrations that are biased toward transcribed and active genes, which elevates the risk of gene disruption and compromised host regulatory functions[372]. Additionally, viral vectors are limited by their packaging capacities; lentiviruses can accommodate from 8-10 kb of DNA and AAVs are restricted to 4.7-5 kb[373,374]. Thus, for complex cell-gene therapeutic cargoes requiring larger transgenes, viral-based delivery is no longer a suitable option. Viral vectors can also trigger an inflammatory response and induce immunogenicity, posing risks to patients and reducing the overall efficacy of therapeutics[375–377]. Epigenetically, viral integration systems suffer from position-effect variegation (PEV) where integration near telomeric or pericentric chromatin leads to silencing of the transgene[378,379]. This can also result in heterochromatin spreading[380], where the silencing marks can spread, downregulating both the transgene and essential neighbouring host genes.

Another transgene delivery platform involves the use of DNA transposons. Transposons utilize a semi-targeted integration approach by recognizing patterns of nucleotide bases to catalyze a cut-and-paste mechanism[381]. The two commonly used transposon-mediated transgene integration systems, piggyBac (PB)[382] and Sleeping Beauty (SB)[383,384], operate by recognizing inverted terminal repeats (ITRs). Researchers can flank cargo DNA with ITRs and co-deliver a transposase enzyme to catalyze a cut-and-paste mechanism. For example, PB recognizes and targets TTAA chromosomal sites, but it remains biased

toward accessible chromatin at transcription start sites and CpG islands. On the other hand, SB integrates into less-specific nucleotide targets through TA dinucleotides. SB demonstrates less bias towards genic regions and, compared to PB and lentiviruses, reflects a more “random” distribution of transgenes. Unlike viral integration systems, transposon systems can integrate large cassettes (up to 100 kb)[385], which is critical for the high capacity and multiplexing of transgenes. However, transposon systems suffer from reduced integration efficiency when transposase enzymes are overexpressed, requiring careful titration between cargo DNA:transposase ratios[386]. Additionally, while it is generally considered safer than retroviruses and viral integration systems, transposon systems still carry the risk for insertional mutagenesis. As transposon-mediated integration systems are considered semi-random, researchers are still limited as they are required to control the precise locus, copy number or allelic zygosity resulting in heterogeneous expression of transgenes across edited populations[387,388]. Transposase systems are also not exempt from experiencing PEV. Like viral systems, transposon-based integrations have been reported to limit therapeutic translation.

Currently, modern bioengineering tools such as programmable endonucleases and recombinases directly address the limitations seen in viral and transposon systems. These technology systems (detailed in Chapter 1.4) can leverage mechanisms like DNA repair pathways or targeted integration (such as CRISPR, HITI, PASTE, and serine recombinases) to integrate transgenes into a specified locus[324,331,389]. By controlling the coordinated integration of genes, researchers can now control the location and dose/copy number of genes, bypassing the unpredictability observed in conventional strategies. Although promising, modern tools have yet to be extensively characterized for their

impact on cell-cycle dynamics and their application in cellular reprogramming studies, thus warranting an active area of investigation.

1.5.2 Mechanisms of Transgene Silencing in Differentiating hiPSCs

hiPSCs are unique as they represent a highly dynamic and specialized epigenetic profile where they retain pluripotency but readily anticipate differentiation cues[390]. This delicate equilibrium adds a layer of additional complexity for genome engineering applications. The introduction of exogenous factors, such as transgenes, risks hiPSCs in triggering apoptotic pathways, differentiation cues, or inducing spontaneous differentiation[391–393]. Therefore, genomically engineering hiPSCs requires careful attention to minimize toxicity and avoid differentiation/cell fate specification.

PEV is commonly observed in various genome engineering studies. In hiPSCs, this poses major risks, as heterochromatin spreading is heritable through mitosis. As a result, researchers have investigated ways to mitigate this by forcing chromatin structures to remain open, such as by incorporating ubiquitous chromatin opening elements (UCOE) into transgene promoters[394,395]. It has been demonstrated that synthetic promoter design can be used to derisk PEV in stem cell-derived lines for future reprogramming studies.

Moreover, clonal heterogeneity remains a major challenge, as random and semi-random integration systems result in the transgene landing in different locations and at varying copy numbers across individual cells[232,396–399]. This results in transcriptional heterogeneity within differentiated cellular populations, producing a mixture of developmental stages and off-target lineages. In complex disease modelling, this background noise makes it especially challenging when trying to determine whether the

observed cellular defects are driven by the pathology or by the variation in differentiation efficiency.

1.5.3 Genomic Safe Harbours as Targets for Expression Stability

To circumvent the random and semi-random integration systems and to support modern bioengineering toolkits, the active characterization of genomic safe harbour (GSH) sites is imperative for synthetic biologists. A common misconception is that GSH sites are solely areas where genes can be inserted safely. While avoiding cytotoxicity is a foundational requirement, the biotechnological field has evolved beyond this definition by using a computational approach to define the criteria for a true GSH (Table **1.5**).

Previously, GSH sites were limited by sequencing and transcriptomic technologies and earlier transgene studies prioritized preventing cytotoxicity and oncogenesis[400]. As a result, the target loci were frequently considered safe simply because cell lines tolerated the disruption of endogenous genes. However, these studies failed to consider the requirements for clinical translation, neglecting the impact of gene disruption from epigenetic silencing and transcriptomic dysregulation when endogenous active gene networks are disrupted.

On the other hand, modern GSH sites were identified through epigenomic mapping and 3D chromatin landscaping[400–402]. By identifying transcriptionally permissive regions and intergenic deserts (specifically those insulated within Topologically Associating Domains, TADs), modern GSHs hold promise in cell-gene therapeutics by ensuring host endogenous machinery and regulatory networks remain untouched. Modern criteria take into account both safety and long-term expression stability, enhancing the predictability of transitioning engineered cell therapies from the bench to the clinic[400].

Table 1.4: Modern Genomic Safe Harbour Site Criteria. This table merges the historical criteria with modern computationally defined criteria, such as implementing epigenetic parameters to ensure long-term and safe expression stability. Adapted from Sadelain et al. (2012)[403], Papapetrou et al. (2011)[404], Pellenz et al. (2019)[405], Aznauryan et al. (2022)[402], Shrestha et al. (2022)[401], and Ahmed et al. (2026)[400].

Criteria	Parameters	Threshold/Condition	Biological Rationale
Genomic Distance	Distance to the 5' end of any known gene	≥ 50 kb	To prevent promoter and gene interference
	Distance to cancer-associated genes	≥ 300 kb	To minimize risk of insertional oncogenesis
	Distance to microRNAs (miRNAs)	≥ 300 kb	To preserve miRNA-related post-transcriptional regulatory function
	Distance from lncRNA and tRNA	≥ 150 kb	To prevent gene expression regulation and cellular development
	Distance from known enhancer regions	≥ 300 kb	To prevent interferences between enhancer and gene interactions
	Distance from telomeres and centromeres	≥ 20 kb	To prevent dysregulation in cell division
Sequence Integrity	Transcription units	Complete exclusion	Reside in intergenic regions to avoid truncating or disrupting endogenous transcripts
	Ultra conserved regions (UCRs)	Complete exclusion	To prevent disruption of conserved regulatory elements
Epigenetic Profile	Chromatin state	Transcriptionally permissive	To ensure transcriptional machinery remains accessible
	Topologically Associating Domains (TAD)	Insulated from endogenous promoters	To restrict aberrant chromatin looping between transgene and genome

1.5.4 Historical "Gold-Standard" Loci: Pseudo-Safe Harbour Sites

Historically, researchers heavily relied on "gold-standard" safe harbour sites for gene engineering. However, under the modern criteria for GSH sites, these loci actually violate the strict requirements and are more accurately classified as pseudo-safe harbour sites.

The widely cited Adeno-Associated Virus Integration Site 1 (*AAVS1*) disrupts the first intron of the *PPP1R12C* gene[406]. Early studies found that integrating genes into this site was safe, as there were no immediate oncogenic or cytotoxic effects[407–411]. However, transgenes integrated at this locus can be susceptible to PEV and DNA methylation, leading to expression instability[412,413]. Moreover, *AAVS1* fails to meet the modern requirement of maintaining a 50 kb distance from adjacent coding genes. For example, the *AAVS1* site is situated adjacent to *TNNT1,* which can be prone to enhancer hijacking and epigenetic dysregulation[405].

Similarly, the human Rosa26 (hRosa26) locus was originally characterized in mice for ubiquitous gene knock-in studies and was later translated for human applications[414]. This locus produces a long non-coding RNA (lncRNA), and while the integration of transgenes does not disrupt protein-coding sequences, modern transcriptomic studies now recognize that lncRNAs play critical roles in chromatin remodelling and transcriptomic regulation[415]. Thus, engineering at the hRosa26 for translational application introduces transcriptomic variables and anomalies that violate the requirements of a true GSH.

Finally, *CCR5* gained popularity after clinical studies showed that disrupting the gene conferred resistance to HIV-1 infection[416–418]. The knock-out of this gene was frequently used for engineering transgenes, such as CAR-T cell manufacturing[419]. Unlike the previous two pseudo-GSHs, *CCR5* is an actively transcribed gene encoding a chemokine

receptor involved in immune signalling that governs the trafficking of immune cells to sites of inflammation[420]. Notably, studies have revealed that *CCR5* deficiency compromises the immune system's ability to clear neurotropic flaviviruses, rendering hosts susceptible to neuroinvasive diseases such as West Nile Virus (WNV)[421,422]. Therefore, while CCR5 knockout may be advantageous for HIV-specific immunotherapies, this site cannot be universally applicable as a therapeutic GSH site, as it impairs host immune function.

1.5.5 Modern Genomic Safe Harbours: Focusing on Intergenic Loci

As researchers began to shift their focus away from intragenic safe harbours, the emphasis on validating intergenic (extragenic) safe harbours has begun[400]. The current generation of GSHs typically implement an *in silico* filtering pipeline involving multi-omics data such as RNA-sequencing, chromatin immunoprecipitation sequencing (ChIP-Seq), and high-throughput chromosome conformation capture (Hi-C)[401,402,423,424]. From this, candidate safe harbour sites are tested and validated, and researchers have begun to publish novel GSH sites that meet the new criterion. The most recent screens and validation studies have identified several intergenic GSH sites as highlighted in **Figure 1.3**.

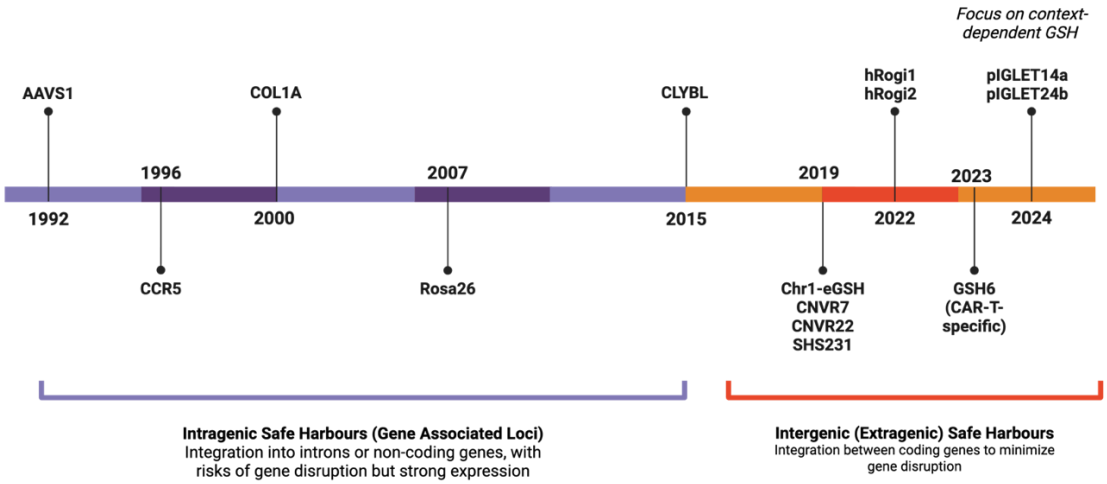


Figure 1.3: Mapping the Discovery and Characterization of Genomic Safe Harbours. For the first two decades from the initial conception of genomic safe harbour sites (GSHs), validated loci were exclusively intragenic[414,425–428]. Intragenic GSHs often resulted in robust expression; however, there was also a risk of gene disruption, mutagenesis, or inadvertent silencing of neighbouring genes. Starting in 2015, the field shifted to intergenic (extragenic) GSH sites to minimize genotoxic risk. Modern GSH sites began to prioritize extragenic loci, as they offer greater safety for clinical and therapeutic application[402,405,429–432].

The GSH sites Chr1-eGSH, SHS231, and GSH6 are located in deep intergenic regions, prioritizing the requirement to be distant from protein-coding genes and regulatory elements. By avoiding these transcription units, these integration sites can almost guarantee that host genes and RNA molecules are not impacted. For example, Chr1-eGSH was tested in hiPSCs through a CRISPR/Cas9 knock-in permitting robust and sustained transgene expression[431]. On the other hand, SHS231 was validated using an array of CRISPR/Cas9, TALENs and meganucleases across various human cell lines, demonstrating efficient target integration and sustained expression[405]. GSH6 has been functionally validated specifically for implementation in cell-based therapy, with the integration of CD19-CAR into primary human T cells[430]. Finally, sites such as CNVR7 and CNVR22 were characterized by their low copy number variation (CNV) through large-scale population-based genomic screening in humans[432]. Alternatively, a recent study has developed intergenic landing pads within *in vivo* vertebrate models called pIGLET (phiC31 Integrase Genomic Loci Engineered for Transgenesis)[429]. Specifically, they validated pIGLET14a and pIGLET24b loci in zebrafish, yielding 25-50% germline transmission rates for generating stable transgenic lines. By targeting intragenic desert areas, transgenic studies can prioritize distance from coding regions, which minimizes

the risks for aberrant chromosomal arrangements and gene disruption that typically complicates clinical translation.

The discovery of human Robust Genomic Integration (hRogi) sites relied on systematic computational filtering of the 3 billion base pair human genome[402]. Researchers developed a bioinformatic algorithm that contained an exclusion criterion detailed in **Table 1.4** in addition to inclusion criteria for transcriptionally permissive euchromatin sites. This *in silico* method generated approximately 2000 candidates, and the authors empirically validated five sites by integrating both fluorescent reporters and therapeutic constructs in HEK293T, K562, hiPSCs and primary human T cells. Researchers have identified robust and stable expression in two of the five sites (hRogi1 and hRogi2), which serve as a versatile novel GSH site for its implementation in biomanufacturing and cell therapies.

1.6 Thesis Objectives

1.6.1 Rationale and Hypothesis

[**Project 1**] The clinical translation of therapies for hard-to-treat neurodegenerative diseases is currently limited by the lack of effective BBB models. hiPSC-derived BBB models offer significant advantages over alternative models, however, their implementation remains challenged by bioengineering bottlenecks, primarily by the stochasticity and heterogeneity observed in hiPSC-derived populations. Conventional differentiation schemes to produce brain endothelial cells rely on sequential exogenous growth factor supplementation (e.g., pre-differentiation medium, bFGF, retinoic acid etc.). As a result, they often fail to yield pure populations of differentiated cells that fully reflect primary cell populations, as exemplified by the epithelial-endothelial ambiguity observed in current hiPSC-derived BBB models[214,228]. Recent literature has begun to showcase the use of exogenous transgenes to perform directed differentiation with the aim of restoring lost phenotype and leveraging intrinsic gene regulatory networks for cell fate specification. Therefore, I hypothesize that using a multiplexed CRISPR activation system targeting endogenous master transcription factors and genes can bypass traditional differentiation schemes, to accelerate and direct hiPSCs towards a brain endothelial-like cell type.

[**Project 2**] The safety and utility of genomic engineering in hiPSCs is often limited by the reliance on random integration technologies introducing mutagenesis risks, unpredictable expression profiles and gene silencing. This has led to the development of targeted gene integration approaches into GSHs to address these concerns and limitations. However, sustaining stable transgene expression in the context of differentiating cells remains a profound and complex challenge. Transgene silencing in differentiating hiPSCs cannot

be explained by a single factor but should be considered multifactorial, including (but not limited to): the integration locus, the transgene's regulatory elements, the transgene sequence, and the cell fate specification. Therefore, I hypothesize that by examining variables (locus, promoter, and germ layer) in isolation, I can characterize specific locus-promoter configurations that permit sustained transgene expression during differentiation into germ layer cell lines. Moreover, to further support downstream bioengineering of hiPSCs, I investigate the implementation of synthetic landing pads in edited hiPSCs to perform subsequent recombinase-mediated cassette exchange to swap cargo from banked hiPSCs. This improves our understanding of factors impacting gene silencing in phenotypically-changing cells, but also enhances the biotechnological toolbox for downstream regenerative medicine and bioengineering research.

1.6.2 Objectives

Objective 1: Engineer a multiplexed CRISPRa platform for accelerated brain endothelial cell fate differentiation.

This will be achieved by designing and delivering a multiplexed sgRNA array of BBB-specific master regulatory transcription factors and genes to bypass stepwise supplementation for brain endothelial differentiation. The resulting iPSC-derived brain endothelial cells will be extensively characterized for brain endothelial phenotypes such as expression of tight junction proteins and transporters, barrier properties, and angiogenesis.

Objective 2: Establish a site-specific, transgene integration platform in hiPSCs to evaluate gene expression stability across four distinct loci and validate the RMCE mechanism.

To develop a stable stem cell engineering platform, a two-tiered genomic integration pipeline will be developed. First, a novel genome integration tool called Programmable addition via site-specific targeting elements (PASTE), will be used to knock-in synthetic reporter constructs containing attP landing pads into four GSHs (e.g., AAVS1, CCR5, hRosa26, and hRogi1). Subsequently, recombinase-mediated cassette exchange (RMCE) will be performed to facilitate cargo swapping with a therapeutic transgene. To optimize the engineering platform, the entire pipeline will first be prototyped in HeLa cells, which will later be translated into hiPSCs, providing a standardized engineering pipeline for locus-specific engineering and modular cargo swapping.

Objective 3: Benchmark transgene promoters at each GSH for tunable and silencing-resistant transgene expression across germ layer differentiation.

To characterize optimal locus-promoter configurations in differentiating cell lines, a library of minicircle vectors will be synthesized, each containing a reporter gene driven by promoters with predicted strengths (EFS, CAG, CBX3.1). Leveraging the pipeline established in Objective 2, these cassettes will be integrated into GSHs (e.g., AAVS1, CCR5, hRosa26, and hRogi1) of hiPSCs, which will be subjected to differentiation into the three embryonic germ layers (mesoderm, endoderm, and ectoderm). Transgene stability and expression will be quantitatively assessed at multiple stages to identify robust locus-promoter configurations for sustained expression.

Chapter 2: Methodology

2.1 Plasmid Design and Molecular Biology

2.1.1 *In Silico* sgRNA Design for CRISPRa

To direct the CRISPRa transactivation complex to target brain endothelial-associated genes, sgRNAs were designed computationally. The transcription start sites (TSS) for the target genes (*CLDN5, PECAM1, CDH5, KDR, ETV2, ERG,* and *FLI1*) were identified using the National Center for Biotechnology Information (NCBI) gene database (**Table 2.1**). Candidate sgRNA spacer sequencers were algorithmically and hierarchically ranked based on predicted off-target and on-target efficiencies using the Broad Institute's CRISPick software, specifically utilizing the CRISPRa design parameters (**Table 2.2**).

To empirically validate the putative targets, the top five-ranked spacer sequences for each gene were selected for functional assessment in human cell lines to identify the highest-performing sgRNA. Optimal spacer sequences contained BbsI restriction site overhangs *in silico* and were synthesized as complementary nucleotides by Integrated DNA Technologies (IDT) to facilitate downstream Golden Gate cloning.

2.1.2 *In Silico* atgRNA, ngRNA, pMini Design and Safe Harbour Selection

For precise genomic integration using the PASTE system, four distinct genomic safe harbours (*AAVS1, CCR5, hRosa26* and the novel *hRogi1* locus) were selected as integration sites. To facilitate both prime editing and serine integrase-mediated cargo insertion, attachment guide RNAs (atgRNAs) and nicking guide RNAs (ngRNAs) were designed for each locus (**Table 2.3**). Spacer sequences were generated using the PASTE computational algorithm described by Yarnall et al. This algorithm was used to identify optimal protospacer adjacent motifs (PAM) sequences that define the constraints

between the prime editor nicking site and the integrase attachment site (attB/attP). A previously reported spacer sequence for Cas9-mediated prime editing at the AAVS1 locus was used as a reference point for the PASTE algorithm. The resulting gRNA spacer sequences were modified *in silico* and ordered as IDT gBlocks for downstream NEB HiFi assembly reactions.

The vector design containing the atgRNA scaffold version 2 (pDY0663) was used to achieve higher prime editing efficiency (**Figure 2.1**). The encoded attB site was modified to ensure compatibility with the serine integrase BceINTa family. Finally, the original spacer sequence targeting the ACTB locus was replaced with the respective safe harbour site sequences.

For the vector design containing the ngRNA, pDY0209 was modified by changing the spacer sequence targeting the same safe harbour site. The spacer sequence for the ngRNA plasmid was designed based on established prime editing parameters described by Yarnall et al., which dictate targeting the nicking guide 45-90 base pairs upstream or downstream of the atgRNA binding site (**Figure 2.1**). Single-stranded oligonucleotides were synthesized with 20-base pair overhangs complementary to the U6 promoter and the scaffold site of the parental ngRNA plasmid for dsDNA and ssDNA assembly.

Finally, the parental minicircle (pMini) plasmid was designed *in silico* to contain a synthetic cargo and the bacterial features required for transformation and plasmid propagation. This plasmid serves to swap the cargo for eventual integration into cell lines. Here, six modular plasmids were designed computationally, each containing a unique promoter (CAG, CBX3.1, and EFS), a fluorescent reporter (copGFP), an RMCE landing pad (Bxb1-compatible attP-GT and attP-AC), and an anti-selection gene (HSV-TK).

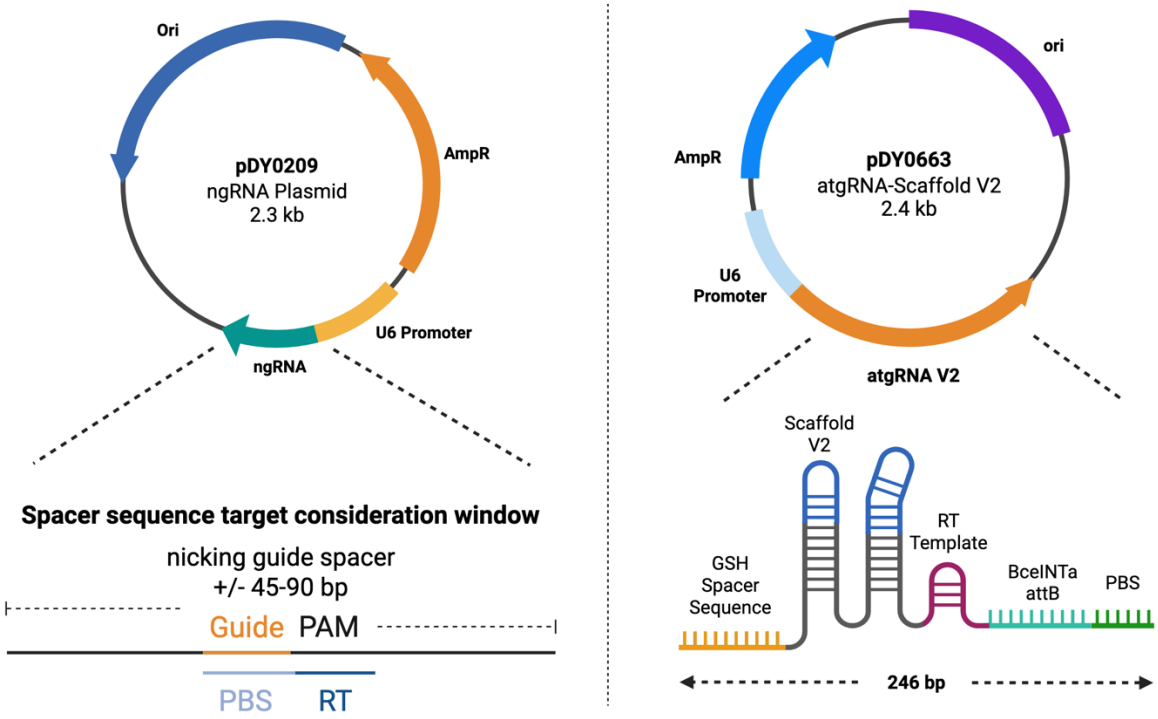


Figure 2.1: Structural Features of PASTE sgRNA Constructs. (Left) pDY0209 ngRNA Plasmid (2.3. kb). Schematic map of backbone vector of nicking guide RNA (ngRNA). The original ngRNA spacer sequence targeting the *ACTB* locus was swapped for GSH targeting spacer sequences. The enlarged schematic below shows the optimal target consideration window for the nicking guide spacer sequence. The target consideration window for optimal nicking efficiency is dictated by running the algorithm 45 bp up and downstream of cut site by the attachment guide RNA (atgRNA). This consideration window was analyzed using Broad Institute's CRISPick algorithm. (**Right**) **pDY0663-atgRNA-Scaffold V2 (2.4 kb).** Schematic map of atgRNA plasmid. Components of the atgRNA is adapted from pegRNA comprised of GSH targeting spacer sequence, scaffold, reverse transcription (RT) template, an attB site, and a primer binding site (PBS).

2.1.3 Design and Construction of Multiplexed CRISPRa Cassette

Sense and anti-sense oligonucleotides targeting brain endothelial markers containing BbsI overhangs were synthesized and cloned into PB-sgRNA (MS2) (a gift from Ying Liu, Addgene plasmid # 102560) through Golden-Gate assembly. Sequence information for

the validated sgRNAs targeting the 3-brain endothelial cell transcription factors and 4 BEC genes is listed in **Table 2.2**. To assemble arrays of multiple sgRNAs (tracrRNA and crRNA) driven by U6 promoters, each fragment was individually amplified from PB-sgRNA (MS2) containing overhangs and cloned into pGGA-Select plasmid (NEB # N0309AAVIAL) with NEBridge Ligase assembly (NEB #E1602). Two sgRNA donor plasmids, 4EG-Select and 3TF-Select, were generated, each containing compatible attL1 and attL2 sites. To generate a multiplexed CRISPRa plasmid targeting key genes involved in brain endothelial differentiation (4EG-SAM, 3TF-SAM), multiplexed sgRNA fragments were amplified from 4EG-Select and 3TF-Select and LR Cloned (Thermo Fisher #12538120) into PB-SAM-DEST (a gift from Ying Liu, Addgene plasmid # 102563), a destination vector containing piggyBac long terminal repeats (LTRs), dCas9-VP64 and MSPH SAM components (**Figure 2.2**).

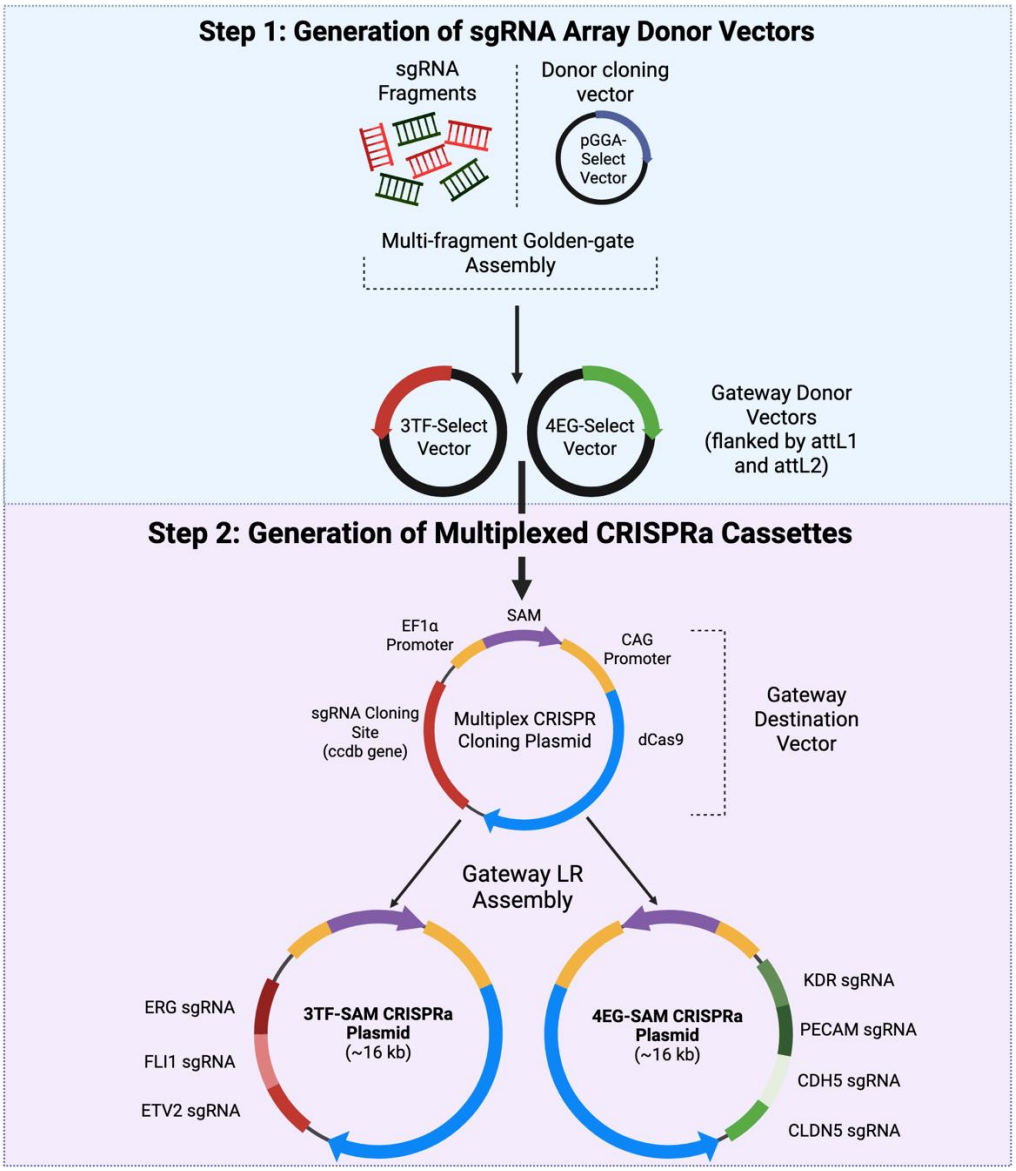


Figure 2.2: Design and Construction of 3TF-SAM and 4EG-SAM Multiplexed CRISPRa Plasmids. This schematic illustrates the two-step cloning process used to generate the multiplexed CRISPRa cassettes. Individual sgRNA fragments flanked with attL1 and attL2 sites were synthesized with restriction sites compatible with a multi-fragment Golden-Gate assembly into the pGGA-Select vector. The second step is a Gateway LR cloning reaction to transfer the sgRNA array assembled in step 1 into a Gateway-compatible destination vector (PB-SAM-DEST). The final plasmids generated

are two final ~16 kb plasmids: 3TF-SAM (targeting transcription factors ETV2, FLI1 and ERG) and 4EG-SAM (targeting brain endothelial markers CLDN5, CDH5, PECAM1, and KDR).

2.1.4 Design and Construction of PASTE Components

PASTE backbone vectors were a gift from Omar Abudayyeh (Addgene plasmid #179106, #179116, #179108, #179109) and used as starting constructs for all genome engineering components in this study. These plasmids encode the core elements required for PASTE-mediated genome integration, including Cas9, prime editor components, serine recombinase machinery, attachment guide RNA (atgRNA), nicking guide RNA (ngRNA) and parental minicircles (pMini).

To generate atgRNAs and ngRNAs targeting candidate genomic safe harbour loci, backbone atgRNA and ngRNA vectors were PCR-amplified using a high-fidelity DNA polymerase (Thermo Fisher Scientific, #12361010) to introduce 5’ and 3’ overhangs compatible with downstream cloning. In parallel, DNA fragments encoding guide RNA spacer sequences targeting safe harbour sites were PCR amplified with complementary overhangs. Spacer sequences targeting the genomic safe harbour sites were obtained using a PASTE algorithm described by Yarnall[331].

To generate pMini plasmids encoding donor cargo DNA (promoter-copGFP-TK), a pMini backbone plasmid was subcloned with a synthesized DNA cassette containing an elongation factor short (EFS) promoter flanked by Golden-gate compatible BbsI restriction sites. NEBuilder HiFi DNA assembly (NEB, #E2621S) was used to insert the EFS cassette into the parental minicircle backbone. Resulting plasmids were validated by restriction digestion and Sanger Sequencing. To generate alternative promoter variants,

Golden-gate assembly was performed using promoter fragments flanked by BbsI overhangs. Assembly reactions were carried out using NEBridge Ligase Master Mix (NEB, #M1100) to replace the EFS promoter region with alternative promoter sequences (CAG and CBX3.1).

2.1.5 Minicircle Induction and Purification

Validated pMini circle plasmids were used to generate minicircle DNA using the ZYC10P3S2T *E. coli* strain provided in the Minicircle DNA Preparation Kit (Systems Bioscience (SBI), #MN925A-1). This strain harbours inducible recombination machinery required for site-specific excision of the bacterial backbone, resulting in a minimized episomal DNA construct. 50 ng of the pMini plasmid DNA was transformed into ZYC10P3S2T competent cells and plated on Luria-Bertani (LB) agar plates supplemented with kanamycin (50 µg/mL). Individual colonies were screened by colony PCR to verify plasmid identity, and a single colony was selected for inoculation.

Selected colonies were inoculated into 5 mL of LB medium containing kanamycin (50 µg/mL) and grown overnight at 30°C at 250 rpm for 16 hours. The following day, 20 µL of overnight culture was diluted into 200 mL of 1X growth medium supplied with the kit and incubated at 30°C at 250 rpm for 16 hours. The next day, optical density at 600 nm (OD600) was measured to ensure bacterial density reached 4-6, and the culture pH was confirmed to be 7.0. Minicircle induction was initiated by adding 200 mL of 1X induction medium directly to the bacterial culture, which was incubated at 30°C and 250 rpm for 5.5 hours. Following induction, bacterial cells were pelleted by centrifugation at 3000 x g for 15 minutes. Minicircles were isolated using an endotoxin-free plasmid maxiprep kit (QIAGEN, #12362) according to the manufacturer's protocol.

2.1.6 Design and Construction of RMCE Donor Cassette

To utilize the cassette exchange with the RMCE-landing pads, a donor plasmid was constructed. The donor cassette was engineered to contain a synthetic cargo comprising an epidermal growth factor chimeric antigen receptor (EGFR1-CAR) coupled with a blue fluorescent protein (BFP) reporter. To facilitate directional and orthogonal recombination, this cargo was flanked by Bxb1-compatible attB-GT and attB-AC recombination sites, which complement the attP sites previously integrated into the safe harbours with the minicircle landing pads. The EGFR-CAR and BFP reporter sequences were PCR amplified with 20-base pair homology overhangs from the pSLQC5 plasmid (A gift from Scott McComb's lab, National Research Council Canada). The resulting fragments were cloned into the pGGA-select donor vector using the NEBuilder HiFi DNA Assembly (NEB, #E2621S).

2.1.7 Plasmid Validation and Sequencing

All assembled plasmids and minicircle constructs were validated prior to downstream applications. CRISPRa and PASTE integration cassettes were validated by PCR, restriction digests, and Sanger sequencing using the primer sequences listed in **Table 2.4**. All inserts were validated through PCR and sequencing. Furthermore, successful minicircle formation and integrity were confirmed by restriction digest and next-generation sequencing (NGS).

2.2 Cell Culture and Cell Engineering

2.2.1 Maintenance of HeLa Cell Lines

HeLa cell cultures (**Table 2.5**) were maintained in a humidified incubator at 37°C with 5% CO_2. Cells were obtained commercially from ATCC and cultured in HeLa Complete

medium, comprised of Dulbecco's Modified Eagle Medium (Wisent, #319-005 CL) supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin. HeLa cells were passaged every 3-4 days at a 1:5 split ratio with 0.025% trypsin-EDTA (Invitrogen, #25200056).

2.2.2 Maintenance of Human Induced Pluripotent Stem Cell (hiPSC) Lines

All hiPSC cell lines, including hAF, hAF09, and hAF12 (**Table 2.5**), were generated in-house at the National Research Council of Canada and reprogrammed from human amniotic fluid samples from patient donors. All experimental protocols using AF cells were conducted in accordance with the guidelines established and approved by the National Research Council Canada Research Ethics Board. HiPSCs were cultured in mTESR1 media (StemCell Technologies #85850) and washed with 1X PBS (without Ca^{2+} and Mg^{2+}) prior to passaging every 3-4 days (1:4 or 1:5 split ratios) using the enzyme-free reagent ReLeSR (StemCell Technologies #100-0483). All iPSCs were cultured on pre-coated hESC-qualified Matrix Matrigel (Corning #354277).

2.2.3 Cell Line Quality Control and Stability

To ensure genomic stability and health of cell lines, prior to bioengineering and differentiation, quality control measures were implemented at the NRC. Standard G-band karyotyping was performed to confirm diploid chromosomal euploidy in the hiPSC cell lines. Moreover, mycoplasma testing was conducted periodically on all cultured cell lines to verify the absence of contamination prior to experimental procedures.

2.2.4 sgRNA Validation and Prototyping CRISPRa in HeLa-SAM Cells

The generation of HeLa-SAM, HeLa-4EG-SAM and HeLa-3TF-SAM cell lines was accomplished using a previously described protocol by Li et al. by co-transfecting 1 μg of

CRISPRa-containing plasmid (PB-SAM plasmid) and 0.2 µg of Super PiggyBac transposase expression vector (Systems Bioscience (SBI) #PB210PA-1) in (2×10^5 cells in one 96-well nucleofector plate). All HeLa cell lines were electroporated using a 4D-Nucleofector paired with SE Cell Line Kit (#V5SC-1010) and plated on a 96-well plate containing HeLa Complete medium for 48 hours. Subsequently, recovered cell lines were cultured and selected for 2 weeks in HeLa-SAM media (HeLa Complete medium containing 15 µg/mL of blasticidin and 200 µg/mL of hygromycin). The expression of HeLa-SAM cell lines was examined by performing rt-qPCR and flow cytometry.

HeLa-SAM cell lines do not contain any sgRNA-expressing cassettes; thus, to individually assess for sgRNA performances (refer to **Table 2.2**), sgRNA-expressing cassettes were transiently transfected through electroporation using SE Cell Line Kit with 0.5 µg of sgRNA-PB-sgRNA (MS2) plasmid as described by Li et al.[433] 48 hours post-nucleofection, expression of target genes was evaluated by rt-qPCR to individually assess sgRNA performance prior to multiplexing.

2.2.5 Generation of Multiplexed CRISPRa hiPSC Cell Lines

To generate stable multiplexed hAF-SAM cell lines (hAF-4EG-SAM and hAF-3TF-SAM) 1×10^6 hAF-hiPSCs were co-transfected with 1.5 µg of the respective SAM plasmids with 0.2 µg of Super PiggyBac transposase expression vector (Systems Bioscience (SBI), #PB210PA-1). Electroporation was performed using the Nucleofector P3 Primary Cell Kit (Lonza #V4LP-3002) in a nucleocuvette with voltage parameters set to CA-137. 48 hours post-nucleofection, iPSCs were cultured with mTESR1-SAM media (containing 15 µg/mL of blasticidin and 200 µg/mL of hygromycin). The expression profiles of hAF-SAM cell

lines were examined by performing RT-qPCR, flow cytometry and immunofluorescence staining.

2.2.6 Validation and Prototyping PASTE Integration Efficiency in HeLa Cells

To validate the efficiency and feasibility of PASTE-mediated genomic integration of synthetic cargoes prior to hiPSCs, the system was first optimized and prototyped in HeLa cells. Guided by recent literature on PASTE optimization[434], the reaction components were adapted for a 96-well nucleofection format. PASTE editing was performed by co-transfecting 2 x 10^5 HeLa cells using the SE Cell Line Nucleofection Kit with 300 ng of the PASTE V4 editor, 40 ng of both the ngRNA and atgRNA and 500 ng of the cargo minicircles. All components were aliquoted into a master mix containing 16.4 µL of SE Solution and 3.6 µL of Supplement Solution. Cells were subjected to the CN-114 program for electroporation, optimized for HeLa cells. Cells were recovered in HeLa complete medium until confluency. Transfection efficiency and cell recovery were recorded and monitored using the IncuCyte Live-Cell Imaging Analysis System (Sartorius).

Upon reaching confluency, cultures were subjected to FACS. An initial sort was performed using a GFP+ gate to capture transfected cells. To eliminate cells with transient episomal expression and isolate clonal populations, a second FACS sort was performed, and the cells were cultured for approximately 2 weeks prior to downstream analysis. Genomic integrations were subsequently validated as described in Section 2.4.

2.2.7 Generation of PASTE-edited HiPSC Cell Lines

Following successful prototyping, PASTE-mediated integrations were performed in hiPSCs reflecting the optimized protocol established in Section 2.2.6. To ensure cellular viability, hiPSCs were pre-treated with 10 µM of Y-27632 ROCK inhibitor (StemCell

Technologies, #72302) and incubated for 1 hour at 37°C. Subsequently, cells were dissociated with Accutase (StemCell Technologies, #7922) and 2 x 10^5 cells were resuspended in the nucleofection buffer containing 16.4 µL of P3 Cell Primary solution and 3.6 µL of Supplement solution. The PASTE components were mixed into the P3 primary solution following the same ratios described in Section 2.2.6 (300 ng of PASTE V4 editor, 40 ng of atgRNA, 40 ng of ngRNA, and 500 ng of cargo minicircle). Electroporation was performed using the Nucleofector P3 Primary Cell Kit (Lonza #V4LP-3002) with the CA-137 program to maximize transfection efficiency. Following nucleofection, the cells were immediately plated in 12-well pre-coated Matrigel plates and recovered in eTeSR (StemCell Technologies, #100-1215) media supplemented with 10 µM Y-27632, a ROCK inhibitor, for the first 24 hours to promote survival and attachment. The following day, media changes were performed using mTeSR1 (StemCell Technologies, #85850) daily until confluency was reached.

2.2.8 Isolation of PASTE-edited Clonal and Pooled hiPSC Cell Lines with FACS

Upon reaching 70% confluency, the engineered cells were subjected to FACS for clonal and pooled cell sorting using a Sony MA900. Cells were pre-treated with 10 µM of the ROCK inhibitor Y-27632 (StemCell Technologies, #72302) for 1 hour at 37°C. Subsequently, cells were dissociated with Accutase (StemCell Technologies, #7922) and 2 x 10^5 hiPSCs were washed with 1% BSA/PBS by centrifugation at 300 x g. Cells were resuspended in a specialized sorting solution containing mTeSR1, 10µM of ROCK inhibitor, 0.5X of CloneR2 (StemCell Technologies, #100-0691) and 0.5X of antibiotic-antimycotic solution (Thermo Fisher Scientific, #15240062).

Cells were sorted using a GFP+ gate into an individual 96-well plate at a density of one cell per well, with the final well containing a pooled sort of approximately 100 cells. Sorting parameters were set to ensure minimal shear stress with a 100 µm nozzle and a set sample pressure rate of 5. The plate was centrifuged at 300 x g for 5 minutes and incubated for 48 hours undisturbed in a hypoxic incubator at 37°C, 5% CO_2 and 5% O_2. Media was changed after 48 hours with eTeSR media containing 0.5X CloneR2 for approximately two weeks until colony formation was observed. Cells were monitored with the IncuCyte Live-Cell Analysis System to track viability and stable fluorescence. Stable clones and pooled cell lines were passaged, expanded, cryopreserved in CS10 Cell Freezing Medium (StemCell Technologies, #07930), and used for downstream analysis.

2.2.9 Generation and Isolation of RMCE-edited HiPSC Cell Lines

After generating clonal and pooled PASTE-edited cell lines, RMCE reactions were performed to swap genomic cargoes. HiPSCs were pre-treated with 10 µM Y-27632, a ROCK inhibitor (StemCell Technologies #72302), and incubated for 1 hour at 37°C. Subsequently, cells were dissociated with Accutase (StemCell Technologies, #7922) and resuspended in the nucleofection buffer containing 16.4 µL of P3 Cell Primary Cell Solution and 3.6 µL of Supplement Solution. Approximately 300 ng of each RMCE plasmid (eeBxb1 and Bxb1-EFS-CAR-BFP) was aliquoted into the P3 primary master mix. Electroporation was performed using the Nucleofector P3 Primary Cell Kit (Lonza #V4LP-3002) with the CA-137 program to maximize transfection efficiency. Following nucleofection, the cells were immediately plated in 12-well pre-coated Matrigel plates and recovered in eTeSR (StemCell Technologies, #100-1215) media supplemented with 10 µM Y-27632, a ROCK inhibitor, for the first 24 hours to promote survival and attachment.

The following day, media changes were performed using mTeSR1 (StemCell Technologies, #85850) daily until 70% confluency was reached.

Upon reaching 70% confluency, the engineered cells were subjected to FACS for clonal and pooled cell sorting. Cells were pre-treated with 10 µM of the ROCK inhibitor Y-27632 (StemCell Technologies #72302) for 1 hour at 37°C. Subsequently, cells were dissociated with Accutase (StemCell Technologies, #7922) and 2 x 10^5 hiPSCs were washed with 1% BSA/PBS by centrifugation at 300 x g. Cells were resuspended in a specialized sorting solution containing mTeSR1, 10µM of ROCK inhibitor, 0.5X of CloneR2 (StemCell Technologies, #100-0691) and 0.5X of antibiotic-antimycotic solution (Thermo Fisher Scientific, #15240062).

Cells were sorted using a BFP+ gate to establish a pooled population using the Sony MA900. Sorting parameters were set to ensure minimal shear stress with a 100 µm nozzle and a set sample pressure rate of 5. The plate was centrifuged at 300 x g for 5 minutes and incubated for 48 hours undisturbed in a hypoxic incubator at 37°C, 5% CO_2 and 5% O_2. The media was changed after 48 hours with eTeSR media containing 0.5X CloneR2 for approximately two weeks until colony formation was observed. Cells were monitored with the IncuCyte Live-Cell Analysis System to track viability and stable fluorescence. Stable clones and pooled cell lines were passaged, expanded, cryopreserved in CS10 Cell Freezing Medium (StemCell Technologies, #07930), and used for downstream analysis.

2.3 Differentiation of Human Induced Pluripotent Stem Cells

2.3.1 Standard Brain Endothelial Differentiation

To perform the standard brain endothelial differentiation of hiPSCs, 100,000 iPSCs were seeded onto Matrigel (Corning #354277)-coated 6-well plates containing mTeSR1 supplemented with 10 μM Y-27632 Dihydrochloride ROCK inhibitor (StemCell Technologies, #72302). Cells were cultured for 3 days with media replenished daily until reaching approximately 80% confluency.

Upon reaching confluency, cells were transitioned to a customized KOEB pre-differentiation medium composed of Knock-out DMEM-F12 (Thermo Fisher, #10829018), 20% (v/v) Knockout Serum Replacement (Thermo Fisher, #10828028), 1X GlutaMAX (Thermo Fisher, #35050061), 1X MEM Non-Essential Amino Acids (Thermo Fisher, #11140050) and 0.2% (v/v) β-mercaptoethanol (Thermo Fisher, #21985-023). Cells were maintained in KOEB medium for 6 days with media changed daily in a 37°C incubator, 5% CO_2.

Following pre-differentiation, the media were replaced with complete Endothelial Medium (EM) to drive terminal BEC differentiation. Complete EM medium was comprised of Human Endothelial SFM media (Thermo Fisher Scientific, #11111044) supplemented with 1% fetal bovine serum (HyClone, #SH30396), 20 ng/mL of basic fibroblast growth factor (bFGF; Fisher Scientific, #13256029) and 10 μM retinoic acid (RA; Sigma, #R2625). The cells were cultured in complete EM for 48 hours.

Prior to subculturing on Day 12, 1.0 μm polyethylene terephthalate (PET) transwell inserts in 12-well companion plates (Corning, #353103, 353503) were pre-coated with sterile aqueous solution containing 80 μg/mL Collagen IV (Sigma, #C5533) and 20 μg/mL plasma fibronectin (Sigma, #F4759) for at least 1 hour at room temperature. The

differentiated cells were pre-treated with the ROCK inhibitor for 1 hour at 37°C. Cells were subsequently dissociated into single cells with Accutase for 20-30 minutes at 37°C. To ensure single-cell suspension, the dissociated cells were passed through a 40 µm cell strainer and pelleted by centrifugation at 300 x g.

The cells were resuspended and seeded into the apical chamber of the pre-coated inserts at a density of 1×10^6 cells per insert in 1 mL of complete EM medium supplemented with 20 µM of ROCK Inhibitor. The basolateral chamber was filled with 2 mL of complete EM medium. On day 13, the apical media were aspirated and replaced with standard complete EM lacking the ROCK inhibitor. On Day 14, the fully differentiated iBECs underwent downstream analyses.

2.3.2 CRISPR-mediated Brain Endothelial Differentiation

A novel abbreviated differentiation protocol was established, leveraging CRISPR-mediated transactivation of BEC genes. 6-well plates were pre-coated with hESC-qualified Matrigel and incubated at room temperature for 1 hour. All cell lines were pre-treated with 10 µM of ROCK inhibitor and incubated at 37°C for 1 hour. Subsequently, approximately 100,000 hiPSCs from the engineered and parental cell lines (Parental, Mock, 3TF-SAM and 4EG-SAM) were dissociated with Accutase and seeded on Matrigel-coated 6-well plates containing mTeSR1 medium supplemented with 10 µM of ROCK inhibitor. The following day, the media was transitioned to mTeSR1 media supplemented with 15 µg/mL blasticidin and 200 µg/mL hygromycin and cells were cultured daily for 3-4 days until they reached 50-60% confluency.

Upon reaching the desired confluency, mTeSR1 media was aspirated and the cells were transitioned to complete primary Human Brain Endothelial Cell culture medium

(CelProgen, #M36201-01) supplemented with 2% fetal bovine serum. To maintain continuous selection pressure and robust constitutive expression of integrated CRISPRa cassettes, the differentiation media was also supplemented with 15 µg/mL blasticidin and 200 µg/mL hygromycin. The media was replenished daily for 5 days.

Following five days of CRISPR-mediated differentiation, the cells were prepared for downstream analyses. The iBECs were dissociated by incubating with Accutase for 20-30 minutes at 37°C. The cell suspension was passed through a 40 µm cell strainer to ensure a single-cell suspension and pelleted by centrifugation (300 x g). The cells were resuspended in the differentiation medium supplemented with ROCK inhibitor and seeded at a density of 1 x 10^6 cells/insert onto a PET membrane insert in a 12-well companion plate that was pre-coated with Collagen IV (80 µg/mL) and fibronectin (20 µg/mL) solution. The following day, the media was replaced with differentiation medium lacking ROCK inhibitor and was maintained for downstream analyses.

2.3.3 Trilineage Germ Layer Differentiation

Prior to differentiation, hiPSC cultures were maintained in mTeSR1 on Matrigel-coated plates, as described in Section 2.2.2. To begin the differentiation process, hiPSCs were lifted for single-cell suspension using Gentle Cell Dissociation Reagent (StemCell Technologies, #100-0485) and incubated at 37°C for 10 minutes. The cell suspension was pipetted 2-3 times and centrifuged at 300 x g for 5 minutes to remove residual enzymes. Viable cells were counted using Trypan Blue (StemCell Technologies, #07050) exclusion. Cells were resuspended in mTeSR1 supplemented with 10 µM of ROCK inhibitor and seeded onto Matrigel-coated 24-well plates at densities specific to the target germ layer cell line (Mesoderm: 100,000 cells, Endoderm: 400,000 cells, Ectoderm:

400,000 cells). Each well was filled with 500 µL of mTeSR1 supplemented with 10 µM ROCK inhibitor for 24 hours to promote attachment.

On the following day, aliquots of STEMdiff Trilineage medium (StemCell Technologies, #05230) were thawed at room temperature. 1 mL of the appropriate lineage-specific STEMdiff Trilineage medium (Mesoderm, Endoderm, or Ectoderm) was aliquoted into the respective well. Differentiation media was refreshed daily for 5 days (Mesoderm and Endoderm) or for 7 days (Ectoderm), depending on the specific differentiation conditions. Upon completion of the differentiation, cells were harvested and prepared for downstream analyses.

2.4 Genomic, Transcriptomic and Protein Analysis

2.4.1 Genomic DNA PCR

PASTE-mediated genomically integrated cells were harvested by washing the cells with 1X PBS three times and scraping cells them in 1 mL of 1X PBS. Cells were pelleted by centrifugation at 300 x g and resuspended in 100 µL of cold 1X PBS. Genomic DNA purification was performed using the Monarch Spin gDNA Extraction Kit (NEB, #T3010S). To lyse the cells, 1 µL of Proteinase K and 3 µL of RNase A were added to the suspension, followed by 100 µL of Cell Lysis Buffer to prevent precipitation. The samples were vortexed thoroughly and incubated for 5 minutes at 56°C in a heating block with manual vortexing every 1.5 minutes.

Subsequently, 400 µL of gDNA Binding Buffer was added, and the samples were pulse-vortexed, then transferred to a gDNA purification column. The columns were centrifuged for 3 minutes at 1000 x g to bind the gDNA, followed by one minute at 13,000 x g to clear the membrane. The column was transferred to a new collection tube, washed with 500

µL of gDNA Wash Buffer, and centrifuged for 1 minute at 13,000 x g. The column was transferred to a DNase-free microcentrifuge tube, and 60 µL of pre-heated gDNA elution buffer was applied to the centre of the matrix and incubated for 1 minute at room temperature. The columns were centrifuged at 13,000 x g to elute the purified gDNA.

Targeted genomic regions were amplified by PCR using Platinum Superfi II DNA Polymerase (Thermo Fisher Scientific, #12361010). Reaction components consisted of 10 µL of 2X DNA polymerase master mix, 0.2 µM of each primer, 50 ng of gDNA template, and nuclease-free water up to a final volume of 20 µL. Thermocycling conditions for extension time were doubled relative to the manufacturer's recommended protocol (30 seconds) to account for the structural complexity of the genomic DNA template. Primer sequences designed to span the integration junctions detailed in **Table 2.3** were used to identify integrated regions. PCR products were resolved on a 1.5% agarose gel with Biotium GelRed Prestain Plus 6X Loading Dye (Fisher Scientific, #NC1712532) and visualized using a fluorescent imaging system to confirm fragment sizes.

2.4.2 Quantitative RT-qPCR

To quantify the transcriptional upregulation of target genes mediated by the CRISPRa (SAM) system, quantitative reverse transcription PCR (RT-qPCR) was performed. To rapidly generate complementary DNA (cDNA) from cultured cells, the EZ FastLane Cell cDNA Kit (QIAGEN, #215011) was used. In a 6-well plate format, cells were washed with 2000 µL of Buffer FCW and aspirated. Next, 1000 µL of Buffer FCP was added, and the plate was incubated at room temperature for 10 minutes to generate FastLane lysates.

To remove genomic DNA contamination, an elimination reaction was prepared by aliquoting 2 µL of 7X gDNA Wipeout Buffer, 4 µL of FastLane lysate, and RNase-free

water to a volume of 14 µL. The solution was incubated at 42°C on a heating block for 5 minutes and placed on ice. For cDNA synthesis, a reverse transcription reaction containing 20 µL was prepared with 1 µL of Quantiscript Reverse Transcriptase, 4 µL of 5X Quantiscript RT Buffer, 1 µL of RT primer mix, and 14 µL of template RNA. The reaction was incubated for 30 minutes at 42°C on a heating block, followed by a 3-minute incubation at 95°C to inactivate the reverse transcriptase.

Quantitative PCR was performed using the QuantiTect SYBR Green PCR Kit (QIAGEN, # 204143). A 50 µL reaction mixture was prepared, comprising 25 µL of 2X QuantiTect SYBR Green PCR Master Mix, 500 ng of cDNA template, 0.3 µM of target-specific primers for the genes of interest (*CLDN5, PECAM1, KDR, CDH5, ETV2, ERG, FLI1*) and the housekeeping gene (*ACTB*) and nuclease-free water to a final volume of 50 µL. Amplification was conducted in a real-time thermocycler (Bio-Rad CFX96) with the following conditions: an initial HotStarTaq DNA polymerase activation step at 95°C for 15 minutes, followed by 40 cycles of denaturation at 94°C for 15 seconds, annealing at 60°C for 30 seconds and extension at 72°C for 30 seconds. Fluorescence data were acquired during the extension phase, and the results were analyzed using the $2^{-\Delta\Delta Ct}$ method, with normalization to the *ACTB* gene.

2.4.3 RNA-sequencing

Cultured cells were isolated in a microcentrifuge tube, spun at 300 x g for 5 minutes, and subsequently washed with 1X PBS. Total RNA was extracted using the Direct-zol RNA Miniprep Kit (Zymo Research, Cat. No. R2050) according to the manufacturer's protocol. Briefly, cell pellets were lysed directly in TRI reagent, mixed with an equal volume of ethanol and the mixture was transferred to a Zymo-Spin column for binding. Columns

were washed with RNA Wash Buffer, treated with DNase I, washed with RNA PreWash, washed with RNA Wash Buffer, and RNA was eluted in DNase/RNase-free water. The quality and quantity of the extracted RNA were assessed using an Agilent Bioanalyzer.

RNA-seq libraries were prepared using the NEBNext Ultra II Directional RNA Library Prep Kit (Cat. No. E7760) following the manufacturer's instructions for sequencing on the Illumina NextSeq 500/550 Mid Output Kit v2.5 (150 cycles; Cat. No. 20024904). Briefly, poly(A)+ mRNA was enriched using oligo-dT beads, followed by fragmentation and cDNA synthesis. End repair, dA-tailing, adapter ligation and USER-enzyme digestion were performed. Lastly, ligated cDNA fragments were amplified by PCR. Final libraries were quantified using a Qubit fluorometer (Thermo Fisher Scientific), qPCR with the NEBNext Library Quant Kit (#E7630), and assessed for quality using an Agilent Bioanalyzer.

Sequencing was performed on the Illumina NextSeq 550 platform using paired end (2 × 75 bp) reads. Raw sequence data were demultiplexed and converted to FASTQ format using Illumina bcl2fastq software. AdaptorRemoval was used to identify adaptor sequences, and FASTX-Toolkit was used for trims and filters. Alignments were performed using STAR 2.7.11a to GRCH38.p14, and RSEM v1.3.3 was used for TPM. Data was normalized with DESeq2, and GSEA was used for gene enrichment analysis.

2.4.4 Western Blotting

Cells were washed twice with 1X PBS (Wisent, #311-013 CL) and lifted by manual scraping. Subsequently, they were pelleted at 300 x g for 5 minutes and stored at -80°C until further processing. Cells were lysed with 1 RIPA buffer with protease inhibitors (BioRad, Sigma #20-188, Roche Diagnostics, #11697498001), and a Quantipro BCA assay (Sigma, #QPBCA) was performed to quantify the protein concentration for each

sample. Western blot assays were performed using the Wes System (Protein Simple, #004-600), and samples were diluted to 1.2 mg/mL in sample buffer and a fluorescent master mix, according to the manufacturer's protocol (Protein Simple, #PS-ST03, #042-195). All samples, blocking reagent, primary antibodies, HRP-conjugated secondary antibodies, and chemiluminescent substrate solutions were pipetted into a plate separation module (Protein Simple, # 042-203, #042-414, #043-31, #043-379, #SM-W004-1, #SM-W008-1). Electropherograms were generated, and the area under the curve (AUC) for chemiluminescent signal peaks was determined. The data and system parameters were analyzed using ProteinSimple's Compass software. Signal intensity of each band was normalized to actin (Sigma, #A3854) to account for loading differences.

2.5 Phenotypic and Functional Characterization

2.5.1 Flow Cytometry, FACS and Fluorescence Microscopy Instrumentation

HeLa cell lines were prepared for staining using staining buffer (1 X PBS, 2% FBS, 1 mM EDTA, 0.1% sodium azide) and fixed with 10% formalin (Fisher Scientific, #SF100-4). For the detection of intracellular proteins, fixed cells were permeabilized with 0.2% Triton X-100 (Sigma, #P7949). Cells were stained with fluorochrome-conjugated monoclonal antibodies or monoclonal primary antibodies and respective fluorochrome-conjugated secondary antibody following the manufacturer's recommended dilutions. For each acquisition, 10,000 events were recorded on the BD Accuri C6 Plus Flow Cytometer and the BD LSRFortessa Cell Analyzer (BD Biosciences). All data were analyzed with FlowJo software.

Similarly, iPSC cell lines were pre-treated with 10 μM Y-27632 Dihydrochloride, a ROCK inhibitor (StemCell Technologies, #72302), for at least 1 hour prior to analysis in liquid

culture. Cells were lifted using Accutase and dissociated into single cell suspensions and pelleted by spinning down at 300 x g. Subsequently, iPSCs were fixed with 10% formalin (Fisher Scientific, #SF100-4) and stained in staining buffer (1X PBS, 2% FBS, 1 mM EDTA, 0.1% sodium azide). Cells were stained with fluorochrome-conjugated monoclonal antibodies or monoclonal primary antibodies and respective fluorochrome-conjugated secondary antibody following the manufacturer's recommended dilutions. For each acquisition, 10,000 events were recorded on the BD Accuri C6 Plus Flow Cytometer and the BD LSRFortessa Cell Analyzer (BD Biosciences). Flow cytometry results were analyzed using FlowJo™ 10.10 (BD Biosciences).

For immunofluorescence imaging, hAF iPSCs and CRISPR-differentiated cell lines were cultured in 12-well plates on 15 mm platinum coated round coverslips (Chemglass, #CLS-1760) coated with Matrigel (Corning #354277). Cells were fixed and permeabilized with ice-cold methanol (Fisher Scientific, #A412P-4) for 10 minutes. All samples were blocked with DAKO Protein block Serum-Free Solution (Agilent, #X0909) for 1 hour. Cells were incubated with primary antibodies in DAKO Antibody Diluent (Agilent, # S0809) for 1 hour and subsequently incubated with secondary antibodies for another hour, as described in **Table 2.6**. All coverslips were washed 3 times with 1 X PBS (Wisent, #311-013 CL) in between and after antibody staining. 10 µg/mL of Hoechst 33342 (Invitrogen, #H3570) was added to counterstain the nucleus and incubated for 10 minutes. All cell samples were mounted in DAKO fluorescent mounting medium (Agilent, # S3023) and visualized with Axiovert 200M Microscope (Zeiss).

2.5.2 Assessment of Pluripotency Markers

To assess for the maintenance of pluripotency in the engineered and parental cell types, cells were assessed for canonical pluripotency transcription factors and surface markers (*SSEA3, SSEA4, SOX2, OCT4, TRA1-60, Nanog, TRA1-81,* and *KLF4*). Cells were processed for immunofluorescent staining as described in Section 2.5.1 with antibodies detailed in **Table 2.6**.

2.5.3 Characterization of Brain Endothelial and Blood-Brain Barrier Markers

Following the standard 14-day brain endothelial differentiation and 5-day CRISPR-mediated differentiation protocol, the iBECs were characterized for core brain endothelial and BBB-specific markers. Cells were fixed and processed as described in Section 2.5.1 for immunofluorescence staining and flow cytometry. Expression and proper localization of junctional markers were assessed using primary antibodies targeting CLDN5, CDH5, PECAM1, GLUT1, ZO-1, and OCLN (**Table 2.6**). Expression of transporters were evaluated by assessing IGF1R, TfR1, TMEM30A, and GLUT1.

2.5.4 Trilineage Germ Layer Assessment

At the designated differentiation endpoints, cell cultures were processed for downstream analysis. Cells were fixed with 10% formalin (Fisher Scientific, #SF100-4) and permeabilized with 0.2% Triton X-100 (Sigma, #P7949). Lineage-specific markers were probed with antibodies specific to germ-layer cells. For example, mesoderm was assayed with CXCR4 (Stem Cell Technologies, #60089PE.1), Brachyury (Biotechne, #IC2085G) and FKL1 (BD, #555308). Endoderm was assayed with SOX17 (BD, #562205) and CXCR4. Ectoderm was assayed with PAX6 (BD, #561552) and Nestin (Thermo Fisher, #53-9843-82) (**Table 2.6**).

2.5.5 Small Molecular Permeability Assay (Sodium Fluorescein)

In vitro BBB modelling for NaFl permeability assay and TEER were performed in 12-well companion transwell plates (**Figure 2.3**). Proceeding differentiation, 12-well transwell inserts were pre-coated with 80 µg/mL of Collagen IV and 20 µg/mL of fibronectin in filter-sterilized H_2O and incubated for 1 hour at room temperature. Differentiated iBECs were suspended with 1 mL of Accutase and incubated at 37°C for 20-30 minutes. Cells were pipetted and filtered through a 40 µm cell strainer to ensure single cell suspension. Cells were harvested by spinning at 300 x g and resuspended in EM media. Cells were counted and approximately 1×10^6 live cells were seeded onto the pre-coated transwell inserts. The media in the apical chamber was filled with EM media containing 10 µM of retinoic acid (RA) and 20 µM of ROCK inhibitor. The basolateral chamber was filled with 2 mL of EM media. The next day, transwell inserts containing a confluent monolayer of iBECs were washed twice in PBS (Wisent, #311-013 CL) and transferred to 12-well companion plates containing pre-warmed 1X transport buffer (5 mM MgCl2, 10 mM HEPES, HBSS pH 7.4). 500 µL of media was removed and replaced with 500 µL of sodium fluorescein solution (50 µg/mL), resulting in a final concentration of 25 µg/mL in each well. Inserts containing no iBECs were used as controls. 100 µL of transport buffer was aliquoted into the bottom of the well in a 96-well plate, and fresh 1X transport buffer was added to the companion plate. iBECs were incubated at 37°C, 5% CO_2, with gentle rotation (20 rpm) in a 311DS Labnet Rotating Incubator containing an orbital shaking platform. 100 µL of transport buffer at the bottom of the companion plates was collected, replaced every 15 minutes for up to 1 hour, and transferred to a 96-well plate. Sodium fluorescein standard curves were generated by formulating NaFl-1X transport buffer solutions at varying concentrations (0 µg/mL, 12.5 µg/mL, 25 µg/mL, 50 µg/mL, 100 µg/mL and 200 µg/mL).

Quantification of sodium fluorescein was performed using a Fluostar Optima microplate reader (BMG Labtech). The endothelial permeability coefficient (P_e) of sodium fluorescein was calculated from the clearance principle to isolate the barrier function from physical resistance of the transwell membrane previously described by Garberg et al[435]. The clearance volume at each timepoint was calculated using the following formula:

$$Clearance = \frac{(Counts \times 10) \times 2 + Correction}{C_0 \times 1000}$$

Where:

- Counts is the raw fluorescent signal in the 100 µL of basolateral acceptor sample.
- 10 is the multiplier to scale the 100 µL sample to 1 mL.
- 2 is the multiplier to account for the total volume of the basolateral chamber.
- Correction represents the accumulated fluorescent counts removed during the previous time-point sampling.
- C_0 is the initial fluorescent concentration in the apical insert
- 1000 is the conversion factor to standardize the clearance volume (10^{-4} cm/min).

Subsequently, the clearance values were plotted against time to generate a linear clearance slope to be used for calculating the endothelial permeability coefficient. The endothelial P_e expressed in 10^{-4} cm/min was calculated by subtracting the background permeability of the blank inserts by the following equation shown below:

$$P_e = \frac{1}{\left(\frac{1}{Slope_{Cells}/A}\right) - \left(\frac{1}{Slope_{Empty}/A}\right)}$$

Where:

- $Slope_{Cells}$ is the clearance slope of the iBEC seeded inserts
- $Slope_{Empty}$ is the clearance slope of the blank control inserts
- A is the surface area of the insert membrane (cm^2)

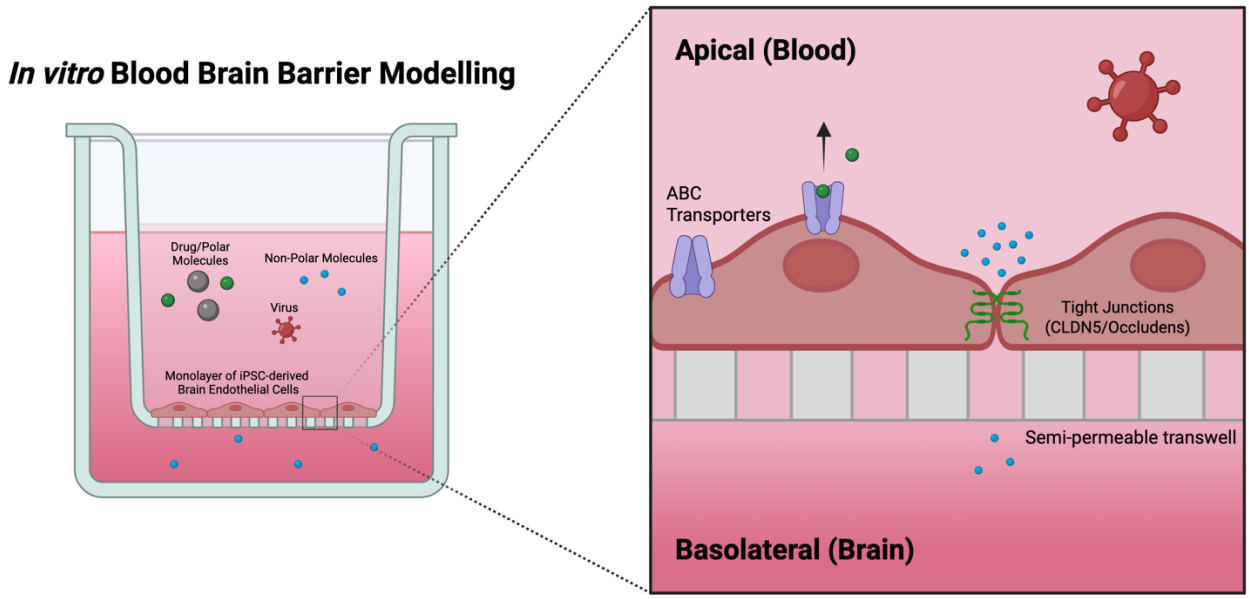


Figure 2.3: Schematic Representation of 2D *In Vitro* Blood Brain Barrier Modelling. A monolayer of iBECs cultured on a semipermeable membrane separates the environments of two distinct compartments to evaluate the permeability of various substrates. (Right) An enlarged cross-sectional view of the apical (blood) compartment and the basolateral chamber (brain).

2.5.6 Transendothelial Electrical Resistance (TEER) Analysis

In vitro BBB modelling for the NaFl permeability assay and TEER measurements was performed in 12-well companion transwell plates (**Figure 2.3**). Proceeding differentiation, 12-well transwell inserts were pre-coated with 80 µg/mL of Collagen IV and 20 µg/mL of

fibronectin in filter-sterilized H_2O and incubated for 1 hour at room temperature. Differentiated iBECs were suspended with 1 mL of Accutase and incubated at 37°C for 20-30 minutes. Cells were pipetted and filtered through a 40 µm cell strainer to ensure single cell suspension. Cells were harvested by spinning at 300 x g and resuspended in EM media. Cells were counted and approximately 1×10^6 live cells were seeded onto the pre-coated transwell inserts. The media in the apical chamber was filled with EM media containing 10 µM of retinoic acid (RA) and 20 µM of ROCK inhibitor. The basolateral chamber was filled with 2 mL of EM media. TEER measurements were performed on a confluent monolayer of iBECs in 12-well transwell inserts, 24 hours following media change. Inserts were transferred to 12-well chambers in the CellZscope apparatus (NanoAnalytics) using 1 cm-diameter electrodes. The system was operated using standard spectrum settings: frequency 1Hz-100 kHz, points per decade 9 and logarithmic spacing. Baseline resistance values were obtained from empty transwell control inserts and were subtracted from iBEC-seeded inserts to isolate for true barrier resistance, yielding the final TEER in Ωcm^2.

2.5.7 Angiogenesis Assay

To evaluate the angiogenic potential of the differentiated iBECs, a tube formation assay was performed. A 24-well plate was coated with 500 µL of fresh undiluted Corning Matrigel (Corning #354277) and incubated at 37°C for 1 hour to ensure complete and even polymerization. Approximately 100,000 iBECs were resuspended in 500 µL of complete endothelial media supplemented with 10 µM of Y-27632 ROCK inhibitor (StemCell Technologies, #72302) and seeded onto the polymerized gel. To establish a positive control condition, parallel wells containing 50 ng/mL of VEGF-165 (R&D Systems #293-

VE) were added to stimulate angiogenesis. Cells were incubated at 37°C with 5% CO_2 overnight to allow for network formation.

Subsequently, the media was aspirated, and cellular networks were stained with 1 µM CellTracker Green CMFDA dye (Thermo Fisher Scientific #C2925) for 1 hour at 37°C. Fluorescent imaging was performed using an Axiovert 200M Microscope (Zeiss). The resulting images were analyzed in ImageJ using the Angiogenesis Analyzer plugin to quantify morphological features, including total tubule length, branch number, and node count.

2.5.8 Longitudinal Assessment of Promoter Stability and Transgene Silencing

To assess the long-term stability of genomically engineered cell lines, cells were monitored for silencing and clearance of transient expression over an extended period of time. Clonal and pooled engineered hiPSCs were maintained in standard pluripotency conditions as described in Section 2.2.2 for a total of 5 passages over the course of 3 weeks. During this expansion phase, cells were cultured and monitored using an IncuCyte Live-Cell Analysis System (Sartorius) to track relative fluorescent expression over time. At the onset of differentiation, aliquots of engineered cell lines were fixed and processed for flow cytometry, as detailed in Section 2.5.1, to verify the stable and homogeneous expression of the integrated reporters.

For maintenance of CRISPRa-engineered cell lines, cells were subjected to a dual-antibiotic selection medium denoted as SAM media. SAM media consisted of standard pluripotency medium supplemented with hygromycin (200 µg/mL) and blasticidin (15 µg/mL). Following every passage, cells were subjected to continuous antibiotic selection to ensure the purity of transgene-expressing populations.

2.6 Computational, Statistical Analysis and Figures

2.6.1 Differential Gene Expression and Pathway Analysis

Differential gene expression (DGE) analysis was performed on raw counts using the DESeq2 package in R[436]. The standard DESeq2 workflow was executed, including size factor estimation for normalization, dispersion estimation and negative binomial linear model fitting. Differential expression statistics were determined using the Wald test. P-values were adjusted for multiple testing using the Benjamini-Hochberg protocol to control for false discovery rate (FDR). Significantly differentially expressed genes were defined by adjusted p-values (Padj) < 0.05 and an absolute ($|Log_2 FC|$) of > 1.0. Gene annotation was performed using org.Hs.eg.db package to map HGNC gene symbols[437]. For the visualization, volcano plots were generated using the EnhancedVolcano package.

To identify functionally enriched biological pathways, Gene Set Enrichment Analysis (GSEA) was performed using the clusterProfiler package[438]. All GSEA visualizations, including enrichment plots and dot plots, were generated using the enrichplot package[439]. Finally, to assess global transcriptional similarity between biological samples and replicates, a Pearson correlation coefficient (r) was calculated for all sample pairs. The resulting matrices were visualized as hierarchically clustered heatmaps (**Figure S7**) using the pheatmaps package. The Pearson correlation was calculated using the standard formula:

$$r = \frac{\sqrt{\mathrm{n}\sum \mathrm{xy} - (\sum \mathrm{x})(\sum \mathrm{y})}}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$$

Where x and y represent the normalized gene expression values for two distinct samples being compared, and n represents the total number of aligned genes analyzed within the dataset.

2.6.2 General Statistical Methods

All quantitative experimental data were presented as the mean ± standard error of the mean (SEM) from at least three biological replicates (n = 3), unless otherwise stated in the figure legends. Statistical analyses and data visualization was performed using GraphPad Prism software version 10.1.1.

To evaluate statistical significance, a Two-Way Repeated Measure Analysis of Variance (RM-ANOVA) was conducted. To correct for multiple testing, the RM-ANOVA was followed by Šídák multiple comparisons post hoc test to determine specific differences between groups at individual time points. Statistical significance was defined as *p < 0.05, ** p < 0.01, ***p < 0.001, and ****p < 0.0001.

2.6.3 Figures

All figures included in chapters 1-5 were designed, generated and created using BioRender.

2.6.4 Data Availability

All bulk RNA-seq data related to Chapter 3 have been deposited in GEO under the accession number GSE299978.

2.7 List of Tables

Table 2.1: Brain Endothelial sgRNA Chromosomal Locations.

Gene	**Chromosomal Location by sgRNA Rank** (Derived from NCBI GRch38)

CLDN5	(1) Chr22:19524629: q11.21 (Sense) (2) Chr22:19524624: q11.21 (Antisense) (3) Chr22:19524433: q11.21 (Sense) (4) Chr22:19524422: q11.21 (Sense) (5) Chr22:19524632: q11.21 (Sense)
PECAM1	(1) Chr17:64390990: q24.1 (Sense) (2) Chr17:64390851: q24.1 (Antisense) (3) Chr17:64390897: q24.1 (Sense) (4) Chr17:64390927: q24.1 (Sense) (5) Chr17:64391152: q24.1 (Sense)
CDH5	(1) Chr16:66366555: q21 (Sense) (2) Chr16:66366646: q21 (Sense) (3) Chr16:66366528: q21 (Sense) (4) Chr16:66366590: q21 (Sense) (5) Chr16:66366468: q21 (Sense)
KDR	(1) Chr4:55125613: q12 (Antisense) (2) Chr4:55125610: q12 (Antisense) (3) Chr4:55125761: q12 (Sense) (4) Chr4:55125762: q12 (Sense) (5) Chr4:55125695: q12 (Antisense)
ETV2	(1) Chr19:35641681: q13.12 (Sense) (2) Chr19:35641662: q13.12 (Antisense) (3) Chr19:35641353: q13.12 (Sense) (4) Chr19:35641345: q13.12 (Sense) (5) Chr19:35641352: q13.12 (Sense)
ERG	(1) Chr21: 38498761: q22.2 (Antisense) (2) Chr21: 38498705: q22.2 (Sense) (3) Chr21: 38498713: q22.2 (Sense) (4) Chr21: 38498719: q22.2 (Antisense) (5) Chr21: 38498683: q22.2 (Antisense)

Table 2.2: Broad Institute's CRISPick sgRNA Analysis.

Gene Symbol	sgRNA Identity	Spacer Sequence (5' to 3')	mRNA ID

CLDN5	sgRNA1	GGCCGCGACTAGGACAACGG	NM_001130861.1
CLDN5	sgRNA2	ACCCGCCGTTGTCCTAGTCG	NM_001130861.1
CLDN5	sgRNA3	CCGCCCCCGAACCTTCAAAG	NM_001130861.1
CLDN5	sgRNA4	CCTTCAAAGAGGGTACCCCC	NM_001130861.1
CLDN5	sgRNA5	CTCGGCCGCGACTAGGACAA	NM_001130861.1
CDH5	sgRNA1	ATAACAGGAAACCATCCCAG	NM_001136154
CDH5	sgRNA2	GGCTCCCCTCCAAAGACGGT	NM_001136154
CDH5	sgRNA3	AGCCAGCCCAGCCCTCACAA	NM_001136154
CDH5	sgRNA4	GGCCAGCTGGAAAACCTGAA	NM_001136154
CDH5	sgRNA5	TCTGTGGAGACAGCCATCCG	NM_001136154
KDR	sgRNA1	GCCGAAACTCTAGAGCGCGG	NM_002253.4
KDR	sgRNA2	GGTGCCGAAACTCTAGAGCG	NM_002253.4
KDR	sgRNA3	CAGCGCAGTCCAGTTGTGTG	NM_002253.4
KDR	sgRNA4	CCAGCGCAGTCCAGTTGTGT	NM_002253.4
KDR	sgRNA5	CCCCTCACCCGGGTACCCGG	NM_002253.4
PECAM1	sgRNA1	AGGGAGGGTCAAGAACGCCA	NM_000442.5
PECAM1	sgRNA2	GCTGGAAACCGGGAACAATG	NM_000442.5
PECAM1	sgRNA3	GGAAGGCCGGTCATTTCCTG	NM_000442.5
PECAM1	sgRNA4	AACAAAGCGGCCGCGTTCGG	NM_000442.5
PECAM1	sgRNA5	GCAAAAGCCCTCTTTGATGG	NM_000442.5
FLI1	sgRNA1	TGACTTCACTTTGCGAATCG	NM_001167681.3
FLI1	sgRNA2	CTTCACTTTGCGAATCGGGG	NM_001167681.3
FLI1	sgRNA3	ACACACTATATAAATATACG	NM_001167681.3
FLI1	sgRNA4	TAGTGTGTGATGCGAAAAGC	NM_001167681.3
FLI1	sgRNA5	CACACAGGATACTTGCTCCT	NM_001167681.3
ETV2	sgRNA1	ATTGTGACGTCAGCTGACGCTGG	NM_001300974.2
ETV2	sgRNA2	CAATCGGCAGAGTAACTGCGGGG	NM_001300974.2
ETV2	sgRNA3	TTACAGCGCGGAAGGGCATGGGG	NM_001300974.2
ETV2	sgRNA4	CTAAACTATTACAGCGCGGAAGG	NM_001300974.2
ETV2	sgRNA5	ATTACAGCGCGGAAGGGCATGGG	NM_001300974.2
ERG	sgRNA1	GACTCCGCATGGCTTCACAC	NM_004449

ERG	sgRNA2	TGTGTCACACGGCGGCACAC	NM_004449
ERG	sgRNA3	CAGTGCAGTGTGTCACACGG	NM_004449
ERG	sgRNA4	GACACACTGCACTGAGACCA	NM_004449
ERG	sgRNA5	ACTGCCACAGCAGGTGTGAC	NM_004449

Table 2.3: Genomic Safe Harbour gRNAs. Sequence identity is as follows: Spacer, Scaffold, Reverse Transcription (RT) Template, BceINTa attB, Primer binding site (PBS).

Gene Locus	gRNA Type	gRNA Sequence (5' to 3')
AAVS1-RH	atgRNA	TTCCCTTTTCCTTCTCCTTCGTTTGAGAGCTATGCTGGAAACAGCATAGCAAGTTCAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCCATCCTAAGAAACGAGAGATGGCACAGGCCCCAGAACGTACTGTGGCGGTTGTACTGATTCACTTCCCCATATCGGAGAAGGAAAAGGGAA
AAVS1-RH	ngRNA	CCATCCTAAGAAACGAGAGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC
AAVS1-MC	atgRNA	GGGGCCACTAGGGACAGGATGTTTGAGAGCTATGCTGGAAACAGCATAGCAAGTTCAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCAGGAGGCCTAAGGATGGGGCTTTTCTGTCACCAATCCGTACTGTGGCGGTTGTACTGATTCACTTCCCCATATCTGTCCCTAGTGGCCCC
AAVS1-MC	ngRNA	ACTAGGAAGGAGGAGGCCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC
CCR5	atgRNA	TATTTTATAGGCTTCTTCTCGTTTGAGAGCTATGCTGGAAACAGCATAGCAAGTTCAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCATCGATTGTCAGGAGGATGATGAAGAAGATTCCAGAGATATGGGGAAGTGAATCAGTACAACCGCCACAGTACAAGAAGCCTATAAAATA

CCR5	ngRNA	GGTACCTATCGATTGTCAGGAGGGTTTTAGAGCTAGAA ATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGA AAAAGTGGCACCGAGTCGGTGC
hRosa26	atgRNA	CTGCAATCCTGAGGGAGCTGGTTTGAGAGCTATGCTGG AAACAGCATAGCAAGTTCAAATAAGGCTAGTCCGTTATC AACTTGAAAAAGTGGCACCGAGTCGGTGCAGGGTGCG GCCTTCTCTCCACCTCCTCCACCGCAGCGTACTGTGG CGGTTGTACTGATTCACTTCCCCATATCCTCCCTCAGGA TTGCA
hRosa26	ngRNA	TTGCAGCTCGCGCCGGTTTTTGGGTTTTAGAGCTAGAA ATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGA AAAAGTGGCACCGAGTCGGTGC
hRogi1	atgRNA	TAGGCCCGAAAGGAGTCTTCGTTTGAGAGCTATGCTGG AAACAGCATAGCAAGTTCAAATAAGGCTAGTCCGTTATC AACTTGAAAAAGTGGCACCGAGTCGGTGCCAGGGACA TGGGTCCAGAGGGGGGTAGTCTCCAGAACGTACTGTG GCGGTTGTACTGATTCACTTCCCCATATGACTCCTTTCG GGCCTA
hRogi1	ngRNA	GGCCTATTACCATGCCTCAGAGGGTTTTAGAGCTAGAA ATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGA AAAAGTGGCACCGAGTCGGTGC

Table 2.4: List of Primers

Primer	Sequence 5’ to 3’	Purpose
Chapter 3 Primer Sets		
qCLDN5-F	ATTTCGCTTCCCCTCCAAGA	RT-qPCR
qCLDN5-R	CCCTGCCGATGGAGTAAAGA	RT-qPCR
qCDH5-F	ACGCCTCTGTCATGTACCAA	RT-qPCR
qCDH5-R	ACGATCTCATACCTGGCCTG	RT-qPCR
qERG-F	CAGAGGGGTGAAGAAGGAGG	RT-qPCR
qERG-R	TCCCCAAATGTCCTGAGTCC	RT-qPCR
qETV2-F	ACCCCGTAAACTTCTCCCAG	RT-qPCR
qETV2-R	TTTTCCGAGACGTCAGAGCA	RT-qPCR
qFLI1-F	GACCGAGTCGTCCATGTACA	RT-qPCR

qFLI1-R	CAGTATTGTGATGCGGCTCC	RT-qPCR
qKDR-F	TGGGGAAAGCATCGAAGTCT	RT-qPCR
qKDR-R	TTCCGGTTCCCATCCTTCAA	RT-qPCR
qPECAM-F	TCCCCTAAGAATTGCTGCCA	RT-qPCR
qPECAM-R	CTGGGGAGAGTTCTTGGTGT	RT-qPCR
hU6-F	GAGGGCCTATTTCCCATGATT	Sequencing
M13-R	CAGGAAACAGCTATGAC	Sequencing
M13-F	GTAAAACGACGGCCAG	Sequencing
T7-R	TAATACGACTCACTATAGGG	Sequencing
TetON-F	TTGATATGCTGCCTGCTGACG	Sequencing
CmR-R	TGCTCATGGAAAACGGTGTAACAAG	Sequencing
TRE-F	AGCTCGTTTAGTGAACCGTCAG	Sequencing
MS2-R	GCGGAAGGGATAGGATTACC	Sequencing
MP-Ins-F	CAAGTTTGTACAAAAAAGCAGGCTTAGA	Sequencing
MP-Ins-R	GAGCTGCTTCGATGGATTTATCAA C	Sequencing
Chapter 4 Primer Sets		
1-AAVS1-MC-Junction-FOR	AGGGAGTTTTCCACACGGAC	gDNA Amplification
2-AAVS1-MC-Junction-FOR	GAGTTTTCCACACGGACACC	gDNA Amplification
1-AAVS1-RH-Junction-FOR	TACCCAGAACCAGAGCCACA	gDNA Amplification
2-AAVS1-RH-Junction-FOR	TGCTGTCCTGAAGTGGACAT	gDNA Amplification
1-CCR5-Junction-FOR	GACAGGGAAGCTAGCAGCAA	gDNA Amplification
2-CCR5-Junction-FOR	TGCATTCATGGAGGGCAACT	gDNA Amplification
1-Rosa26-Junction-FOR	GTCCTTACAAGACGGGAGGC	gDNA Amplification
2-Rosa26-Junction-FOR	AAACTAATCGGAGTCGCCCC	gDNA Amplification
1-Transgene-Universal-REV	CCTGGCCGTCGTTTTACTGGGTTTG	gDNA Amplification

2-Transgene-Universal-REV	CTCCCTGGCCGTCGTTTTACTGG	gDNA Amplification
1-Rogi1-Junction-FOR	AGACAGGAAAAAGCCTGGTCA	gDNA Amplification
2-Rogi1-Junction-FOR	GAAGTGCAGTCCTGTGCTCG	gDNA Amplification
1-CBX-Transgene-Junction-REV	AGTACCCCTTTGGGACCACT	gDNA Amplification
2-CBX-Transgene-Junction-REV	CGGGGGTTATCATTGGCGTA	gDNA Amplification
3-Rogi1-Junction-FOR	GAAGTGCAAGCATTGATGCAACTTACTG	gDNA Amplification
3-Rosa26-Junction-FOR	AATCCGCAGTCTCGTTGCATAAAATCAG	gDNA Amplification
3-CCR5-Junction-FOR	ATGACATCAATTATTATACATCGGAGCC	gDNA Amplification
3-AAVS1-MC-Junction-FOR	AGTTAGAACTCAGGACCAACTTATTCTG	gDNA Amplification
3-AAVS1-RH-Junction-FOR	CAAAGAAAGCAAGAGGATGGAGAGGTGG	gDNA Amplification
copGFP-Junction-REV	CGGCAAATCTTTGCCAAAATGATGAGAC	gDNA Amplification
AAVS1-RH-REV	CTTTAACTGGCCCTGGCTTTGG	gDNA Amplification
AAVS1-MC-REV	GGTTAATGTGGCTCTGGTTCTGGG	gDNA Amplification
CCR5-REV	CCCAAGTGATCACACTTGTCACC	gDNA Amplification
hRosa26-REV	GGGTAGGTAGGTGGGGTGTAGA	gDNA Amplification
Rogi1-REV	CTTTAGTCCTAGTGCCATGAAGTGG	gDNA Amplification
EFS-GG-F	GGCTACGAAGACGAGGAGGGGCAGAGCGCACATCGC	Golden Gate Promoter Edit
EFS-GG-R	GGCTACGAAGACCCATGGCTGTGTTCTGGCGGCAAACC	Golden Gate Promoter Edit
CBX3.1-GG-R	GGCTACGAAGACTTATGGTCTGTGTTCTGGCGGCAAAC	Golden Gate Promoter Edit
CBX3.1-GG-F	GGCTACGAAGACATGGAGGGGAGGTGGTCCCTGCAG	Golden Gate Promoter Edit
CAG-GG-REV-V2	GGCTACGAAGACTAATGGTCTGTAGGAAAAAGAAGAAGGCATG	Golden Gate Promoter Edit

CAG-GG-FOR-V2	GGCTACGAAGACAGGGAGTCAATGTATCTTAGACATTGATTATTGAC	Golden Gate Promoter Edit
EF1A-GG-R	GGCTACGAAGACATATGGTCACGACACCTGAAATGGAAGAAAAAAACTTTGAACC	Golden Gate Promoter Edit
EF1A-GG-F	GGCTACGAAGACTGGGAGGGCTCCGGTGCCCGTCAG	Golden Gate Promoter Edit
pGK-GG-R	GGCTACGAAGACCCATGGTCGAAAGGCCCGGAGATGAG	Golden Gate Promoter Edit
CMV-GG-F	GGCTACGAAGACCTGGAGCGATGTACGGGCCAGATATAC	Golden Gate Promoter Edit
CMV-GG-R	GGCTACGAAGACCGATGGAATTTCGATAAGCCAGTAAGC	Golden Gate Promoter Edit
pGGA-PASTE-REV	ATCGGTCGATTCGTTTGATGTTATG	RMCE Cassette Cloning
pGGA-PASTE-FOR	GTCGTATTAATTTCGCGGGCGGAAC	RMCE Cassette Cloning
CAR-BFP-REV	GCTAGAGCCAGACATGATAAGATAC	RMCE Cassette Cloning
CAR-BFP-FOR	GGGCAGAGCGCACATCG	RMCE Cassette Cloning
PASTE-CAR-FOR	GTCGTGACGCGGGATCC	RMCE Cassette Cloning
PASTE-CAR-REV	CGATGTGCGCTCTGCCC	RMCE Cassette Cloning
CD28-REV	GTACCTGCGAAGACCCTCCT	RMCE Cassette Cloning
hU6-DWN-R	CGGTGTTTCGTCCTTTCCAC	PASTE Cloning and Sequencing
Scaffold-UP-F	GTTTTAGAGCTAGAAATAGCAAGTT	PASTE Cloning and Sequencing
atgRNA-DWN-F	GCTTGGGCCGCTCGAG	PASTE Cloning and Sequencing
atgRNA-DWN-R	CTCGAGCGGCCCAAGC	PASTE Cloning and Sequencing
pMini-UP-R	CGCGCCCGGGGAGCC	PASTE Cloning and Sequencing

P2A-TK-UP-F	GGTTCTGGTGCCACAAACTTC	PASTE Cloning and Sequencing
P2A-TK-DWN-R	TCAGGCGTAGTCAGGCACA	PASTE Cloning and Sequencing
pMini-DWN-F	CCCCAACTGGGGTAACCTTT	PASTE Cloning and Sequencing
sGold-P2A-R	AAGTAAGCTGAAGTTTGTGGCACCAGAACCCAAGTGCGCCTCCAATGTCT	PASTE Cloning and Sequencing
F1-Assembly-For	CCCAGCTTTCTTGTACAAAGTGGT	PASTE Cloning and Sequencing
F3-Assembly-Rev	ACCACTTTGTACAAGAAAGCTGGG	PASTE Cloning and Sequencing
pMini-UP-F	GGCTCCCCGGGCGCG	PASTE Cloning and Sequencing
sGold-UP-R	CTTGAAGGGAGTAGACGCCAT	PASTE Cloning and Sequencing
AAVS1-Mc-F	ACTAGGAAGGAGGAGGCCTA	PASTE Cloning and Sequencing
AAVS1-Mc-REV	TAGGCCTCCTCCTTCCTAGT	PASTE Cloning and Sequencing
Rosa26-REV	CCAAAAACCGGCGCGAGCTGCAA	PASTE Cloning and Sequencing
M13-REV (NEW)	GTCATAGCTGTTTCCTG	PASTE Cloning and Sequencing
pMini-EFS-BB-REV	GTTTCCTGATGGTCGTCTTC	PASTE Cloning and Sequencing
pMini-BB-FOR	GGTAACCTTTGAGTTCTCTC	PASTE Cloning and Sequencing
pMini-Insert-REV	GAGAGAACTCAAAGGTTACC	PASTE Cloning and Sequencing
pMini-Insert-EFS-FOR	GAAGACGACCATCAGGAAACCTAATACGACTCACTATAGGG	PASTE Cloning and Sequencing
pMini-All-Insert-FOR	CCATCAGGAAACAGCTACTGAGACTAATACGACTCACTATAGGG	PASTE Cloning and Sequencing
pMini-Seq-FOR	CCATCAGGAAACAGCTACTGA	PASTE Cloning and Sequencing
pMini-EFS-Seq-FOR	CCATCAGGAAACCTAATACGAC	PASTE Cloning and Sequencing

pMini-UP-V3-F	CAGGGCGTGCCCTTG	PASTE Cloning and Sequencing
pMini-Promoter-R	AGTCTCAGTAGCTGTTTCCTG	PASTE Cloning and Sequencing
pMini-Ins-UP-Kozac-F	GCCGCCACCATGGCGTCTACTCCCTTC	PASTE Cloning and Sequencing
pMini-Ins-V3-R	CCCCAACTGAGAGAACTCAAAGG	PASTE Cloning and Sequencing
pMini-All-BB-REV	TCAGTAGCTGTTTCCTGATGG	PASTE Cloning and Sequencing
pMini-LP-V3-FOR	TCTGGCTCTAGCTATCCGTG	PASTE Cloning and Sequencing
SV40P-HR-V3-REV	CAAACCACGGATAGCTAGAGCCAGAGGATAGCTAGAGCCAGACATG	PASTE Cloning and Sequencing
WPRE-Mid-REV	ACGCCACGTTGCCTGAC	PASTE Cloning and Sequencing
TK-Mid-FOR	GGGCCCTCGATGTGCTG	PASTE Cloning and Sequencing
RNAseP-F	AGATTTGGACCTGCGAGCG	PASTE Cloning and Sequencing
RNAseP-REV	GAGCGGCTGTCTCCACAAGT	PASTE Cloning and Sequencing
copGFP-CNV-FOR	TGAACGGCGTGGAGTTCG	PASTE Cloning and Sequencing
LKO.1 5'-F	GACTATCATATGCTTACCGT	PASTE Cloning and Sequencing
hU6-F	GAGGGCCTATTTCCCATGATT	PASTE Cloning and Sequencing
M13-F	GTAAAACGACGGCCAG	PASTE Cloning and Sequencing
gRNA-ori-R	GGGGGCGGAGCCTATG	PASTE Cloning and Sequencing
SP6-F	ATTTAGGTGACACTATAG	PASTE Cloning and Sequencing
T7-F	TAATACGACTCACTATAGGG	PASTE Cloning and Sequencing
MLVRT-F	TGATAGCCGTTATGCTTTTGC	PASTE Cloning and Sequencing

M13Reverse-For	CAGGAAACAGCTATGAC	PASTE Cloning and Sequencing

Table 2.5: List of Human Cell Lines Used in this Thesis.

Cell Identity	Full Name	Derived Source	Reprogramming Method
hAF	Human amniotic fluid-derived iPSC	Human amniotic fluid	Episomal vectors
hAF09	Human amniotic fluid-derived iPSC Clone 9	Human amniotic fluid	Sendai virus vector reprogramming
hAF12	Human amniotic fluid-derived iPSC Clone 12	Human amniotic fluid	Sendai virus vector reprogramming
HeLa	Henrietta Lacks Cells	Human cervical adenocarcinoma	N/A

Table 2.6: List of Antibodies.

Antibody	Supplier	Catalog
Alexa Fluor 488 Goat-Anti-Rabbit IgG	Invitrogen	A11034
Alexa Fluor 488 Goat-Anti-Mouse IgG	Invitrogen	A28175
Anti-Human CD144 VE-Cadherin	Invitrogen	53-1449-42
CLDN5-488	Invitrogen	352588+A6
ERG	AbCam	ab133264
ETV2	AbCam	ab181847
FLI1	Invitrogen	MA1-196
Cas9	AbCam	EPR18991
VEGFR2-647	AbCam	ab307529
CD31-APC Flow Antibody	Thermo Fisher	17-0319-42
ZO-1	Invitrogen	40-2200
GLUT1-AF647	BD Pharmacology	566580
CD31	AbCam	Ab28364
Rb mAb GLUT1-488	AbCam	Ab195359

VE-Cadherin ICC	AbCam	ab33168
Occludin	Invitrogen	331588
IGF1R	Cell Signalling	3027S
TFR1	Invitrogen	13-6800
TMEM30A	Thermo Fisher	BS-16576R
GLUT1	Abcam	Ab195359
Rabbit anti Oct4	Abcam	AB19857
Rabbit anti KLF4	Abcam	AB151733
Goat poly anti Nanog	R&D Systems	AF1997
Rat anti SSEA3	Cedarlane	MAB143
Mouse anti SSEA4	R&D Systems	MAB1435
Mouse anti TRA1-60-Alexa488	StemCell Technologies	60064AD.1
Anti TRA1-81-Alexa488	StemCell Technologies	60065AD.1
Goat-anti-rabbit-Alexa 568	Invitrogen	A11011
Goat-anti-mouse-Alexa 488	Invitrogen	A10680
Donkey-anti-goat-Alexa 568	Invitrogen	A11079
Donkey-anti-rat-Alexa 488	Invitrogen	A21208
Streptavidin-HRP	Biotechne	042-414
CXCR4	Stem Cell Technologies	60089PE.1
Brachyury	Biotechne	IC2085G
FLK1	BD	555308
SOX17	BD	562205
PAX6	BD	561552
Nestin	Thermo Fisher	53-9843-82

Table 2.7: Summary of Plasmids and Vectors Utilized and Generated.

Plasmid Name	Application	Source
PB-SAM	PiggyBac CRISPRa-SAM vector	Addgene, 102559
PB-sgRNA (MS2)	PiggyBac sgRNA expression cassette	Addgene, 102560

PB-SAM-DEST	PiggyBac CRISPRa-SAM cloning vector	Addgene, 102563
pGGA-Select	Cloning vector	NEB, N0309AAVIAL
3TF-Select	3TF sgRNA array cloned vector	In-house
4EG-Select	4EG sgRNA array cloned vector	In-house
3TF-SAM	PiggyBac 3TF multiplexed CRISPRa-SAM plasmid	In-house
4EG-SAM	PiggyBac 4EG multiplexed CRISPRa-SAM plasmid	In-house
pDY0663	Template atgRNA	Addgene, 79108
pDY0209	Template ngRNA	Addgene, 179109
pDY0423	Template Parental Minicircle	Addgene, 179116
pDY0863	PASTE V4 editor	Addgene, 179106
pSLQC5	EGFR-CAR-BFP vector	NRC, McComb lab
eeBxb1	Enhanced Bxb1 serine recombinase expression cassette	Addgene, 222339
RMCE-EFS-CAR-BFP	Synthetic RMCE donor cassette (bxb1 compatible)	In-house
atgRNA-AAVS1-MC	AAVS1 targeting gRNA	In-house
atgRNA-CCR5	CCR5 targeting gRNA	In-house
atgRNA-hRosa26	hRosa26 targeting gRNA	In-house
atgRNA-hRogi1	hRogi1 targeting gRNA	In-house
ngRNA-AAVS1-MC	AAVS1 targeting gRNA	In-house

ngRNA-CCR5	CCR5 targeting gRNA	In-house
ngRNA-hRosa2	hRosa26 targeting gRNA	In-house
ngRNA-hRogi1	hRogi1 targeting gRNA	In-house
pMini-CAG	Parental minicircle with transgene regulated by CAG promoter	In-house
pMini-CBX3.1	Parental minicircle with transgene regulated by CBX3.1 promoter	In-house
pMini-EFS	Parental minicircle with transgene regulated by EFS promoter	In-house
MC-CAG	Minicircle vector with transgene regulated by CAG promoter	In-house
MC-CBX3.1	Minicircle vector with transgene regulated by CBX3.1 promoter	In-house
MC-EFS	Minicircle vector with transgene regulated by EFS promoter	In-house

Chapter 3: Accelerated Reprogramming of hiPSCs into Functional Brain Endothelial-like Cells

Declaration of Previously Published Material

The results and text presented in this chapter have been adapted from the following peer-reviewed publication, reproduced with permission from the Nature Publishing Group (refer to Copyright Acknowledgements for full details):

Roy W. Hwang et al., 2026, Accelerated Reprogramming of hiPSCs into Functional Brain Endothelial-Like Cells Using Multiplexed CRISPR Activation, Scientific Reports.

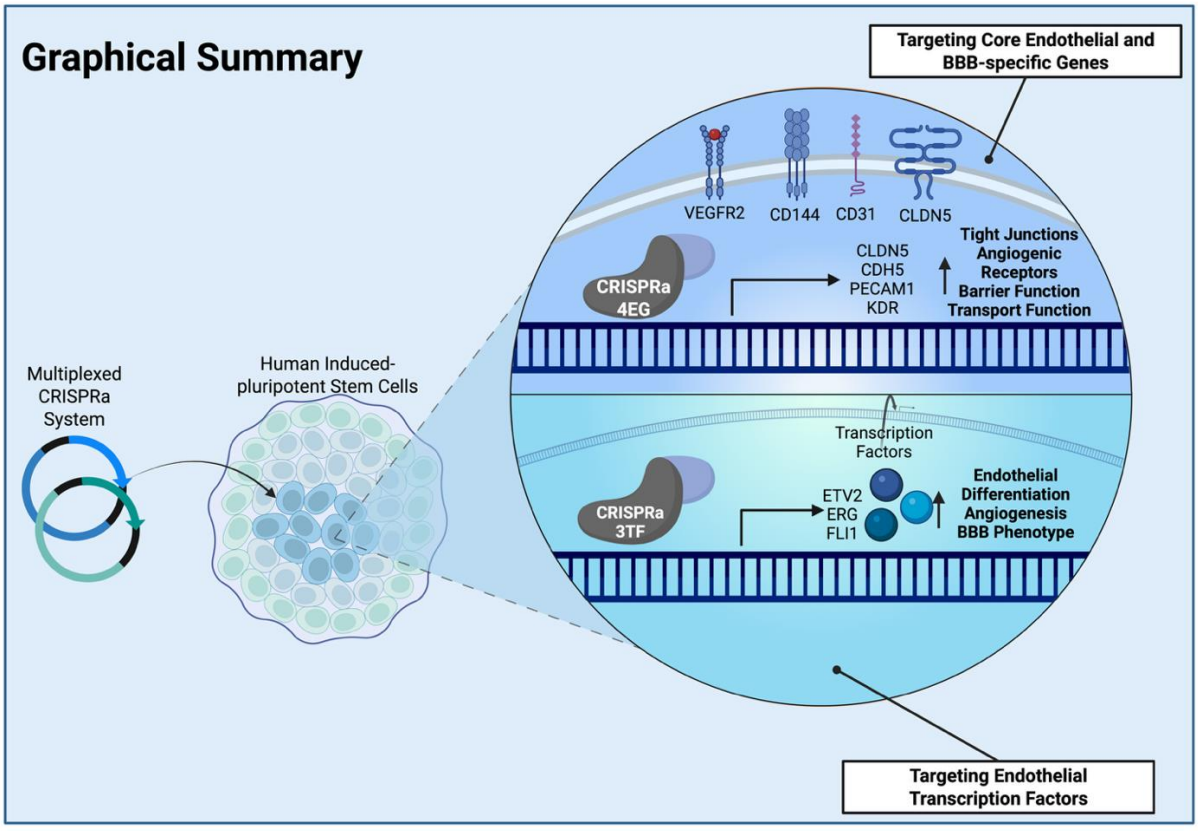


Figure 3.0: Graphical Summary of Multiplexed CRISPRa Reprogramming of iPSCs.

3.0 Findings Summary

Human induced pluripotent stem cells (hiPSCs) offer a renewable and scalable source for generating brain endothelial cells (BECs), enabling the development of *in vitro* blood-brain barrier (BBB) models that closely reflect human physiology. In this study, we demonstrate a streamlined differentiation strategy for producing hiPSC-derived BEC-like cells (iBECs) using multiplex CRISPR/dCas9 activation (CRISPRa) to transcriptionally reprogram iPSCs by selectively and simultaneously upregulating BEC transcription factors (ERG, ETV2, FLI1) and BEC genes (CLDN5, CDH5, PECAM1, KDR). This approach significantly accelerates differentiation from 13 to five days while maintaining BBB characteristics. CRISPRa-iBECs exhibit high transendothelial electrical resistance (TEER ~ 400 $\Omega \cdot cm^2$), low transcellular permeability (< 0.015×10^{-3} cm/min), and expression of key BBB markers including Claudin-5, ZO-1, CDH5, PECAM1, Occludin, and GLUT-1, as well as receptor-mediated transporters TfR1, IGF1R, and TMEM30A. This novel CRISPRa-directed differentiation strategy not only accelerates BEC differentiation but also demonstrates the utility of CRISPR gene regulation as an option in modifying the phenotype of iPSC-derived cells.

3.1 CRISPRa Screening of Guide RNAs for Brain Endothelial Gene Targets

We identified seven genes involved in brain endothelial differentiation and phenotype that are typically under-expressed in iPSC-derived BEC models: three transcription factors (ERG, ETV2, FLI1) and four genes important to BEC function (CLDN5, CDH5, PECAM1, KDR) were selected for CRISPR-mediated activation in iPSCs[440]. The four selected BEC genes function in barrier formation, membrane transport and angiogenesis, all inherent phenotypes essential to modeling the BBB. Furthermore, the panel of transcription factors are known to play an essential role in angiogenesis and BEC differentiation[440] (**Table 3.1**).

Table 3.1: Genetic Markers for Brain Endothelial Cell Lineage Reprogramming.

Data derived from Lu_2021[228].

Genetic Markers	Gene Group
ERG	ETS Transcription Factor
ETV2	ETS Transcription Factor
FLI1	ETS Transcription Factor
CLDN5	BEC Marker
PECAM1	BEC Marker
CDH5	BEC Marker
KDR	BEC Marker
[a]SOX7, SOX17, SOX18	Transcription Factor

[a] *Genes reported in subsequent studies and not included in original construct*

Among the described CRISPRa systems, we selected the Synergistic Activation Mediator (SAM) system based on its ability to be multiplexed and its described consistency in transcriptional activation performance in mammalian cell lines[344] (**Figure 3.1a**). For each gene, we used CRISPick (Broad Institute) to design five sgRNAs that target the genomic region immediately upstream of the respective transcription start sites (**Figure 3.1b**) (**Table 2.2**). The performance (induction of gene expression) of

individual sgRNAs was initially assessed in HeLa cells by determining relative mRNA expression by RT-qPCR (**Figure S1**). Optimal individual sgRNAs for each gene were selected for inclusion in the multiplexed CRISPRa construct (**Figure 3.1c)**. Interestingly, we observed a high degree of variability in gene activation among individual sgRNAs targeting a given gene in HeLa cells, as well as a lack of correlation between gene induction and the CRISPick sgRNA ranking algorithm (**Figure 3.1d**). Thus, identification of optimal sgRNAs based on observed efficacy in CRISPRa/i systems appears to be an important step to ensure functionality.

3.2 Constructing a Multiplexed CRISPRa System Targeting BEC Genes

We used the one-step piggyBac transposase (PB)-based multiplex system described by Li et al[433] to activate the expression of sets of genes in an all-in-one vector (**Figure 3.1e)**. Two multiplex sgRNA expression vectors were designed based on the highest performing sgRNA for each gene identified in **Figure 3.1c**. Here, each sgRNA expression cassette consists of a U6 promoter, a unique spacer sequence and an MS2 scaffold. Two distinct all-in-one multiplexed CRISPRa expression plasmids were assembled; 3TF (3 transcription factors; ERG, ETV2, FLI1) and 4EG (4 endothelial marker genes; CLDN5, CDH5, PECAM1, KDR) (**Figure 3.1f)**.

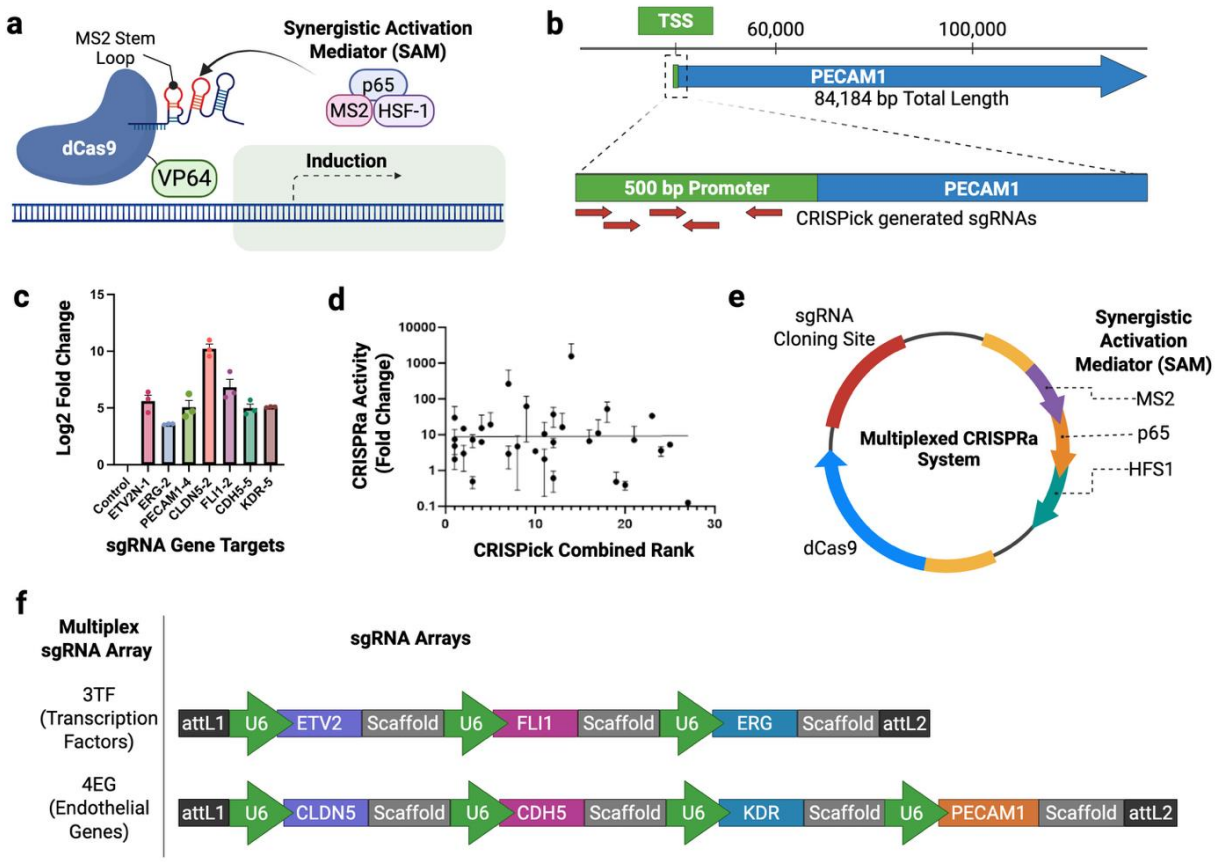


Figure 3.1: Identifying and Screening Brain Endothelial Gene Targets and sgRNAs for CRISPRa in HeLa Cells. (a) Graphical illustration of CRISPRa utilizing Synergistic Activation Mediator (SAM) system. **(b)** Schematic overview of how various sgRNAs targeting the same gene were designed with Broad Institute's CRISPick software. Candidate sgRNAs were specifically selected within 500 bp upstream of the transcription start site of each target gene. **(c)** Screening for putative highly performing sgRNA designs for each key brain endothelial gene target in HeLa cells co-transfected with CRISPRa and sgRNA vectors. RT-qPCR was performed for each HeLa transfected culture to identify which sgRNA elicited a consistently robust response. Individual Ct values were assessed in biological triplicates (n=3) and normalized to Log_2 fold change (mean ± SEM). **(d)** Correlation plot of observed CRISPRa performance (y-axis) and expected performance based on of the combined rank predicted by CRISPick algorithm (x-axis) (mean ± SEM). **(e)** Schematic of all-in-one multiplexed CRISPRa vector. **(f)** Tandem sgRNA designs targeting core BBB genes and transcription factors. Final engineered sgRNA designs are 3TF and 4EG.

3.3 Validation of CRISPR Activation of Target Genes in hiPSCs

To determine if the expression of the multiplexed CRISPRa systems would have an impact on the viability of hiPSCs, human amniotic fluid (hAF) derived iPSCs (hiPSCs) were co-electroporated with a CRISPRa plasmid (3TF-SAM, 4EG-SAM or PB-SAM-DEST) and piggyBac transposase. The transfected cells underwent hygromycin selection and growth was monitored for 10 days. Stably transfected iPSCs were obtained after 2-3 days in the presence of hygromycin, with stably transfected mock CRISPRa (PB-SAM-DEST plasmid lacking an sgRNA cassette) hiPSCs exhibiting a doubling time of ~ 1.3 days (**Figure S2**). The doubling time of stably transfected hiPSCs with an sgRNA cassette (3TF-SAM, 4EG-SAM) did not appear to be different from mock transfected hiPSCs. Expression of dCas9-SAM was verified in the transfected hiPSCs using flow cytometry and western blotting (**Figure 3.2 a, b**). The efficacy of CRISPRa-mediated induction of the sgRNA-targeted genes was evaluated using flow cytometry for BBB-specific tight junction, adherens junction and endothelial proteins (CLDN5, PECAM1, CDH5 and KDR; **Figure 3.2c**). We observed an induction of all four BBB-specific surface markers in hiPSCs expressing in both the 4EG-SAM and 3TF-SAM constructs, with higher levels of expression observed in 4EG-SAM cells. In contrast, mock and parental (unmodified hiPSCs) hiPSC cultures exhibited negligible expression of the targeted markers (**Figure 3.2c)**. Interestingly, targeting the ETS transcription factors (3TF-SAM) produced significant levels of induction of CLDN5, PECAM1 and KDR, and lesser induction of CDH5 expression. While it was not necessarily expected that the 3TF-SAM cells would exhibit upregulation of these specific BEC markers, their expression is consistent with the ETS factors governing iPSC differentiation into BECs.

To further validate the induction of the CRISPRa-targeted genes, RT-qPCR analysis of gene expression in transfected hiPSCs against mock lines revealed levels of RNA transcript induction (**Figure 3.2d**) that were similar to protein level induction in both transformed cell lines (**Figure 3.2c)**. Stable CRISPRa-integrated hiPSC lines were isolated and expanded for downstream analysis, and the retention of pluripotency marker expression was validated using immunofluorescence. **Figure 3.2e** shows that the pluripotency marker expression profile (SSEA3, SSEA4, SOX2, OCT4, Nanog, TRA1-60, TRA1-81 and KLF4) of the 3TF-SAM and 4EG-SAM hiPSCs was indistinguishable from the parental hiPSCs.

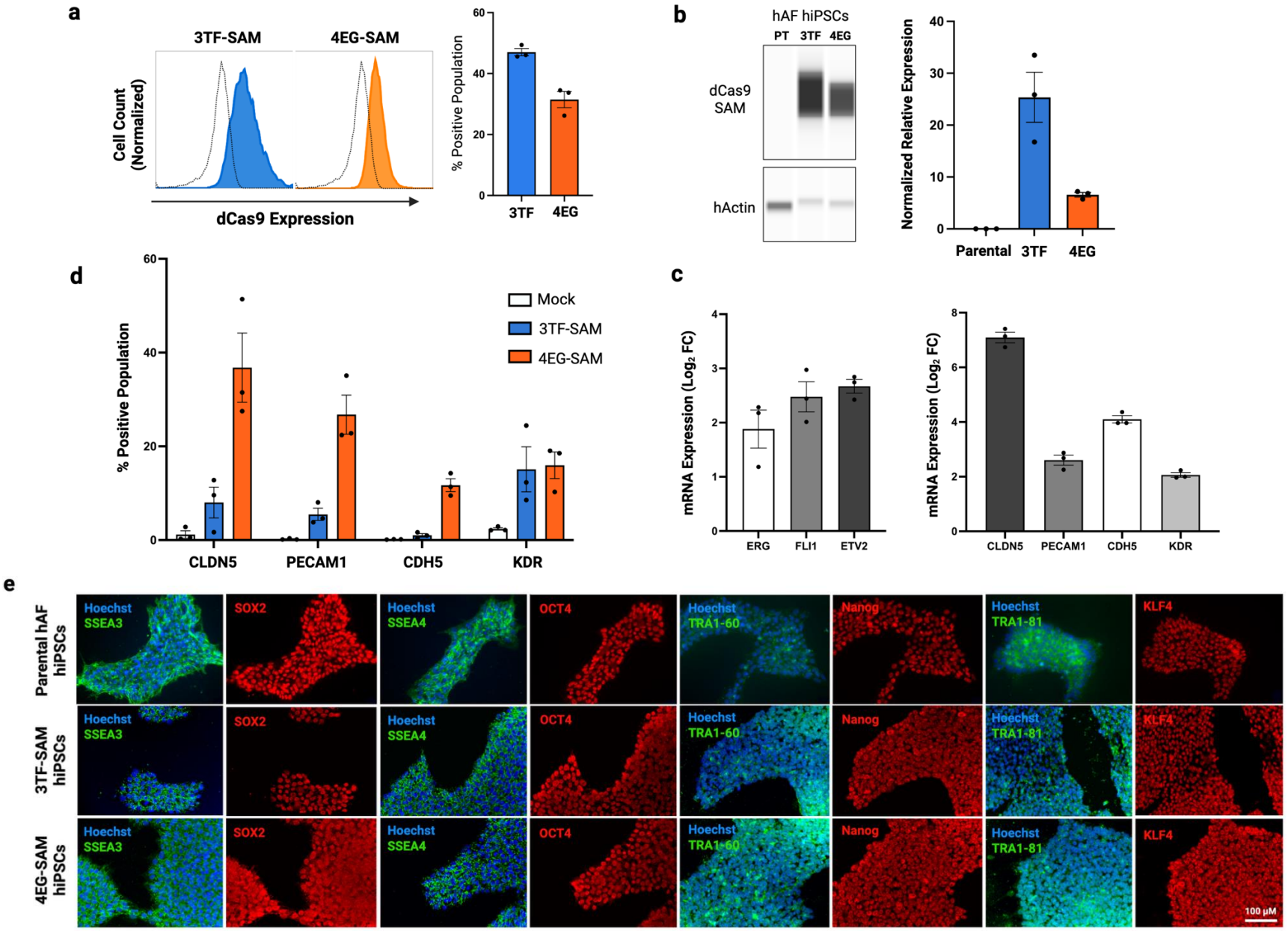

a
3TF-SAM
4EG-SAM
Cell Count (Normalized)
dCas9 Expression
% Positive Population
3TF
4EG
b
hAF hiPSCs
PT 3TF 4EG
dCas9 SAM
hActin
Normalized Relative Expression
Parental 3TF 4EG
c
mRNA Expression (Log2 FC)
ERG FLI1 ETV2
CLDN5 PECAM1 CDH5 KDR
d
% Positive Population
Mock
3TF-SAM
4EG-SAM
CLDN5 PECAM1 CDH5 KDR
e
Parental hAF hiPSCs
3TF-SAM hiPSCs
4EG-SAM hiPSCs
Hoechst SSEA3
SOX2
Hoechst SSEA4
OCT4
Hoechst TRA1-60
Nanog
Hoechst TRA1-81
KLF4
100 µM

Figure 3.2: Multiplexed CRISPRa Integration in hiPSCs Induces Core BEC Genes and Transcription Factors.

(a) Representative flow cytometry histograms for dCas9 expression in 3TF-SAM hiPSCs (blue), 4EG-SAM hiPSCs (orange) and parental hiPSCs (dotted). Viable cells were gated using forward scatter versus side scatter (FSC vs. SSC) followed by the gating of singlet populations (FSC-H vs. FSC-A). Bar graphs (mean ± SEM) present percentage positive population from 3 biological replicates (n=3).

(b) Western blot showing the expression of dCas9-SAM in CRISPRa transformed hiPSCs and parental (PT) hiPSCs. Quantification of protein bands for dCas9, normalized to hActin, showing expression levels (mean ± SEM) in 3 biological replicates (n=3). Western blots were cropped to clearly depict the expression; uncropped image can be found in the supplementary figure S5.

(c) Flow cytometric analysis for BEC tight junction (CLDN5, PECAM1), adherens junction (CDH5) and angiogenic protein (KDR) expression in CRISPRa transformed hiPSCs. Viable cells were gated using forward scatter versus side scatter (FSC vs. SSC) followed by gating of singlet populations (FSC-H vs. FSC-A). The bar graph (mean ± SEM) presents percentage positive population from 3 biological replicates (n=3).

(d) mRNA expression of transcription factors and brain endothelial genes in 3TF-SAM and 4EG-SAM hiPSCs determined by RT-qPCR. Individual Ct values were assessed in biological triplicates (n=3) and standardized with Log_2 fold change (mean ± SEM).

(e) Immunofluorescence staining of parental hAF hiPSCs, 3TF-SAM hAF hiPSCs and 4EG-SAM hAF hiPSCs stained for nuclear pluripotency markers (SOX2, Oct4, Nanog, KLF4) and membrane-associated pluripotency markers (SSEA3, SSEA4, Tra1-60 and Tra1-81). Hoechst staining (blue) marks cell nuclei.

3.4 Activation of 3TF or 4EG Genes Induces and Accelerates hiPSC BEC Differentiation

Observation of the transformed hiPSCs in culture resulted in the unexpected finding that both the 3TF-SAM and 4EG-SAM hiPSC cultures exhibited consistent spontaneous differentiation (**Figure 3.3a**). In contrast, spontaneous differentiation was not observed in mock transfected hiPSC cultures, which were indistinguishable from the parental hiPSCs.

BEC differentiation of the 3TF-SAM and 4EG-SAM cells was conducted using established protocols, as previously described[433]. Transitioning the 3TF-SAM and 4EG-SAM hiPSCs into KOEB pre-differentiation media resulted in a change to cobblestone-like morphology, consistent with endothelial cells (ECs) derived from mock and parental hiPSCs. The cells acquired a distinct, flattened elongated morphology at the end of the differentiation period in complete EM media (**Figure S3**). We subsequently assessed if BEC differentiation could be directly induced by culturing the 3TF-SAM and 4EG-SAM hiPSCs, as well as the parental and mock hiPSCs, in complete human BEC media (hBEM). CRISPRa-integrated hiPSC cultures that retained pluripotent markers were seeded for differentiation. As expected, all cell lines expanded and began differentiating within the first 24 hours (**Figure 3.3b, 3c**). Whereas complete cell death was observed for both parental and mock hiPSCs by 72 hrs, the 3TF-SAM and 4EG-SAM hiPSCs remained viable and continued to differentiate as observed by emerging colonies of flattened, elongated endothelial-like cells. At day 5 in hBEM, the morphology of 3TF-SAM and 4EG-SAM cells gave to confluent, cobblestone, elongated morphology (**Figure 3.3d**). No notable differences between the cell lines were observed in terms of

purity, quantity and efficiency. To further validate the reproducibility of our multiplexed CRISPRa platform, we repeated the transduction and differentiation using two additional hAF-iPSC lines, hAF09 and hAF12 (**Table 2.5**). The hAF09 and hAF12 lines showed similar accelerated differentiation and BEC morphology (**Figure S4**).

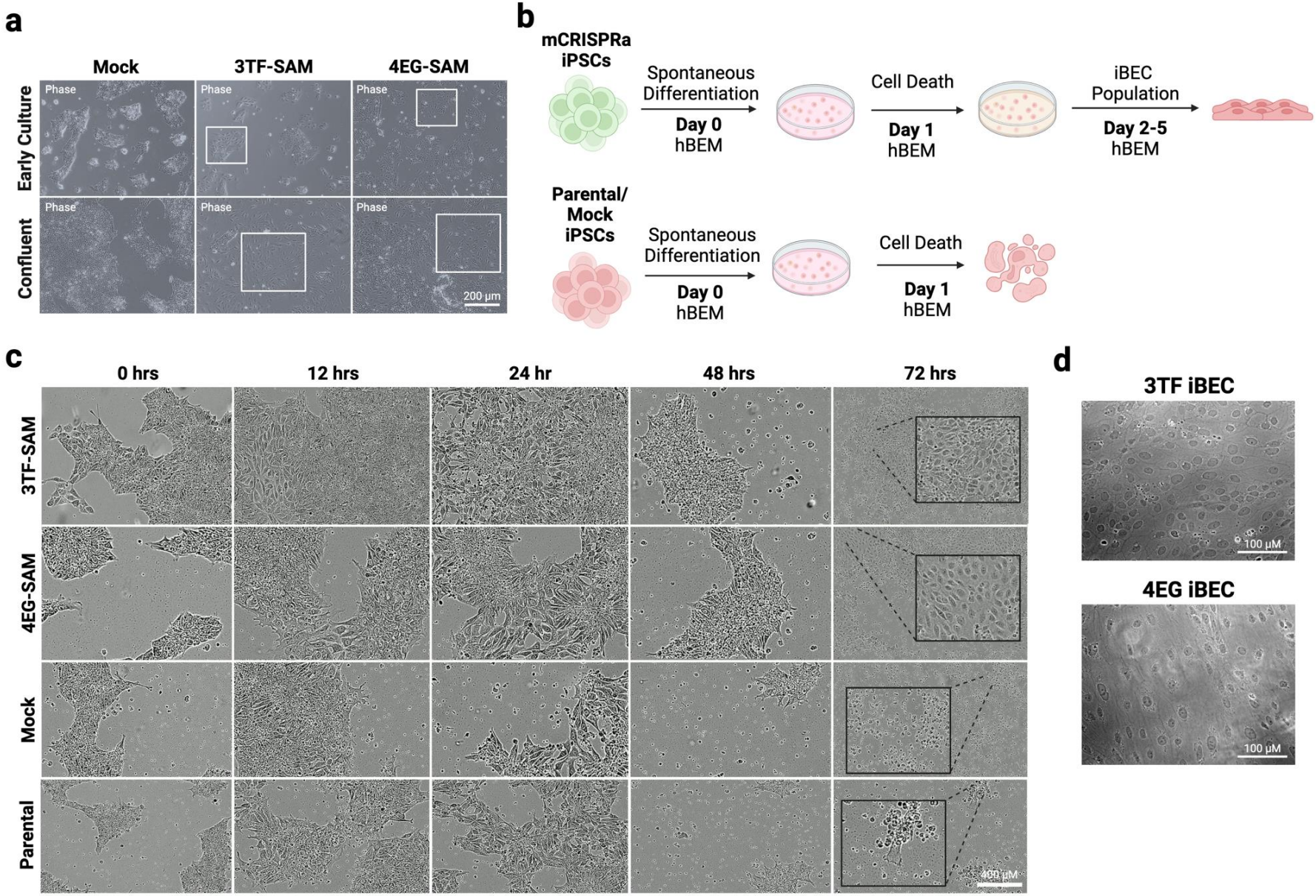

a
Mock
3TF-SAM
4EG-SAM
Early Culture
Confluent
Phase
200 µm
b
mCRISPRa iPSCs
Spontaneous Differentiation
Day 0 hBEM
Cell Death
Day 1 hBEM
iBEC Population
Day 2-5 hBEM
Parental/ Mock iPSCs
Spontaneous Differentiation
Day 0 hBEM
Cell Death
Day 1 hBEM
c
0 hrs
12 hrs
24 hr
48 hrs
72 hrs
3TF-SAM
4EG-SAM
Mock
Parental
400 µM
d
3TF iBEC
100 µM
4EG iBEC
100 µM

Figure 3.3: CRISPRa-mediated Induction of BBB Genes and Transcription Factors Initiates Directed Differentiation into iBEC-like Cells.

(a) Spontaneous differentiation (boxed regions) observed in 3TF-SAM and 4EG-SAM transfected hiPSCs and normal culture conditions observed in mock transfected hiPSCs.

(b) Schematic overview of novel differentiation scheme employed in CRISPRa-transformed hiPSCs, illustrating the culture outcomes for both parental and CRISPRa-transformed cells.

(c) Imaging of differentiation of CRISPRa-transformed hiPSCs to iBEC-like cells over 72 hours. Phase contrast images were captured at 10X every 6 hours with a Live Cell Imaging system (Incucyte).

(d) Resulting iBEC-like cells after 5 days of differentiation in hBEM exhibit endothelial-like morphology.

3.5 Characterization of BBB Phenotype Induction in CRISPRa-differentiated iBECs

The CRISPRa-differentiated iBECs were further characterized to validate BEC-specific marker expression and transcriptome profile. We first assessed dCas9 expression in CRISPRa-differentiated iBECs by western blotting to confirm the absence of transgene silencing in the terminally differentiated state. The CRISPRa iBECs derived from all three hiPSC lines (hAF, hAF09 and hAF12) retained strong dCas9-SAM expression (**Figure S5**). Moreover, flow cytometry was performed to compare BEC surface protein expression in CRISPRa iBECs (accelerated differentiation) with iBECs differentiated using the standard protocol (**Figure 3.4a**). The CRISPRa iBECs (both 3TF-SAM and 4EG-SAM) differentiated using the abbreviated protocol exhibited heterogeneous expression of CLDN5, PECAM1 and CDH5. While the vast majority of the cells were positive for CLDN-5, only 40-50% of the cells were positive for CDH5 and PECAM1 (**Figure 3.4b**). In contrast, expression of the markers in parental hiPSCs differentiated using the standard protocol was largely indistinguishable from unstained control cells (**Figure 3.4b**). The induction and expression of CLDN5, PECAM1, CDH5 and KDR in all three iPSC lines were assessed, revealing that the CRISPRa-differentiated iBECs expressed the markers regardless of the parental iPSC line of origin (**Figure 3.4b)**. Interestingly, the levels of induction were generally consistent among the three iPSC lines, with the notable exceptions of PECAM1 and KDR in hAF12-4EG, suggesting that factors affecting the efficiency of induction were relatively consistent among the iPSC lines. Furthermore, flow cytometry analysis was performed to validate the expression levels of BBB tight junction proteins (ZO1 and Occludin) and proteins involved in receptor and

carrier-mediated transcytosis and transport, such as IGF1R, TfR1, TMEM30A and Glut1. The data show that the expression of these important factors for BBB function was elevated in the CRISPRa-differentiated iBEC lines compared with control iBECs (**Figure 3.4c**). The expression and subcellular localization of BEC proteins were assessed in CRISPRa-differentiated iBECs by immunocytochemistry. **Figure 3.5** shows that CLDN5 and PECAM1 localized to the cell membrane, producing a staining pattern consistent with that of ZO-1 and GLUT1. In contrast to tight junction proteins, the adherens junction protein CDH5 exhibited a diffuse staining pattern in CRISPRa iBEC cells that was not limited to intracellular junction.

To assess the ability of CRISPRa iBECs to form a functional and intact barrier, the CRISPRa iBECs were seeded and cultured in semipermeable transwell inserts and the transendothelial electrical resistance (TEER) was measured every 24 hours to characterize barrier formation and integrity (**Figure 3.6a**). Strong barrier formation was observed as early as day 2, while peak TEER values were observed at approximately day 4-7 (377 ± 49 $\Omega \cdot cm^2$ in 3TF-SAM iBECs and 553 ± 39 $\Omega \cdot cm^2$ in 4EG-SAM iBECs). TEER values for the 4EG-iBECs were significantly greater than those observed for 3TF-iBECs. By day 8-9, TEER values declined sharply for both 3TF- and 4EG-iBECs, a trend that is consistent with our previous work on the parental iBECs[169]. We subsequently assessed sodium fluorescein (NaFl) paracellular permeability (P_e) as a measure of barrier integrity and tight junction function. In CRISPRa iBECs, lower permeability was observed in 4EG-SAM iBECs (Pe = 0.0084 ± 0.002 × 10^{-3} cm/min) compared to 3TF-SAM iBECs (0.027 ± 0.091 × 10^{-3} cm/min), consistent with their higher TEER values (**Figure 3.6b**). Both CRISPRa iBECs demonstrate highly selective paracellular diffusion as their P_e

values were substantially lower than previously reported values in hCMEC/D3 cells[163] and comparable to other hiPSC models (P_e = 0.028-0.034 × 10^3 cm/min)[163].

The ability to undergo angiogenesis (the formation of blood vessels and capillaries) is a key characteristic of primary BECs. In isolated primary BECs, this is demonstrated as the ability to form tubules in the presence of angiogenic factors such as VEGF[441,442]. However, some iPSC-derived *in vitro* BBB models do not report BECs being able to undergo angiogenesis in *in vitro* assays. We hypothesized that inducing the expression of KDR (VEGFR2) could induce an angiogenic phenotype on iPSC-derived BECs. **Figure 3.6c** shows that both 3TF-SAM and 4EG-SAM iBECs seeded in undiluted Matrigel layers containing VEGF formed networks of tubules, which is indicative of the angiogenic potential (**Figure 3.6d**). This is in sharp contrast to the absence of angiogenic phenotype in the standard differentiation with parental iPSCs (**Figure 3.6c-d**).

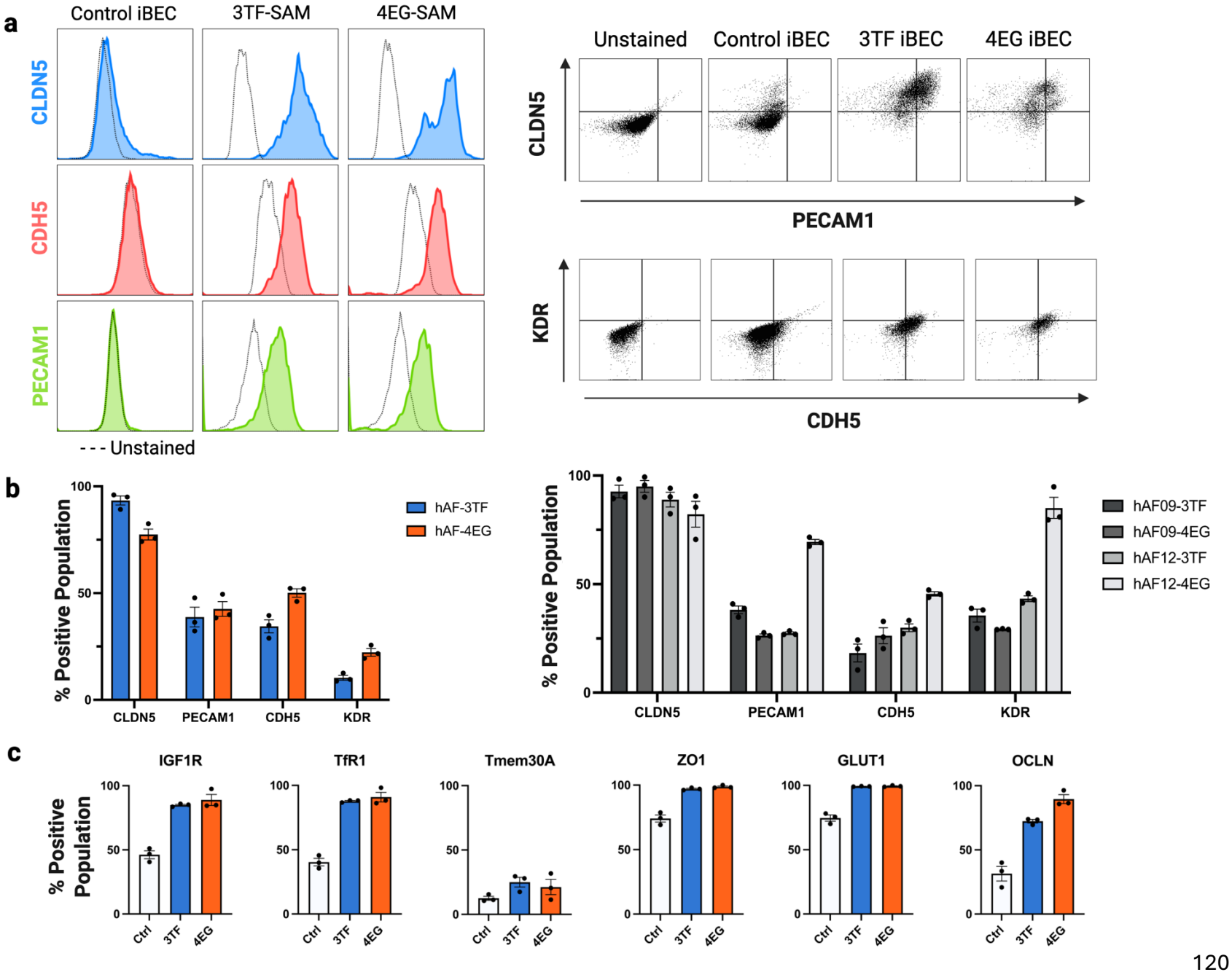

a
Control iBEC
3TF-SAM
4EG-SAM
CLDN5
CDH5
PECAM1
--- Unstained
Unstained
Control iBEC
3TF iBEC
4EG iBEC
CLDN5
PECAM1
KDR
CDH5
b
% Positive Population
hAF-3TF
hAF-4EG
CLDN5
PECAM1
CDH5
KDR
hAF09-3TF
hAF09-4EG
hAF12-3TF
hAF12-4EG
c
IGF1R
TfR1
Tmem30A
ZO1
GLUT1
OCLN
% Positive Population
Ctrl
3TF
4EG

Figure 3.4: Quantitative Flow Cytometric Analysis Reveals Enhanced Expression of Brain Endothelial and BBB-specific Markers Following Multiplexed CRISPRa Reprogramming.

(a) Flow cytometry histograms and corresponding dot plots comparing the expression of *CLDN5* (blue), *CDH5* (red) and *PECAM1* (green) in control iBECs (standard differentiation scheme) and 3TF-SAM and 4EG-SAM iBECs using the abbreviated differentiation scheme. The gating strategy illustrates the initial selection of viable hiPSCs based on forward scatter versus side scatter (FSC-A vs. SSC-A) followed by singlet discrimination (FSC-H vs, FSC-A) to exclude aggregates prior to final marker analysis.

(b) Flow cytometric analysis of BEC tight / adherens junction (*CLDN5, CDH5*), angiogenic receptor (*KDR*), and adhesion molecule (*PECAM1*). The left panel presents BEC marker analysis for CRISPRa-differentiated iBECs derived from hAF hiPSCs. The right panel demonstrates the reproducibility of the CRISPRa differentiation protocol across two additional independent parental hiPSC lines (hAF09 and hAF12). Viable cells were gated as above. Bar graphs (mean ± SEM) present percentage positive from 3 biological replicates (n=3).

(c) Flow cytometric analysis of BBB transporters and associated markers across control iBECs (white), 3TF-SAM (blue) and 4EG-SAM (orange) iBEC populations. Viable cells were gated as above. Bar graphs (mean ± SEM) present percentage positive from 3 biological replicates (n=3).

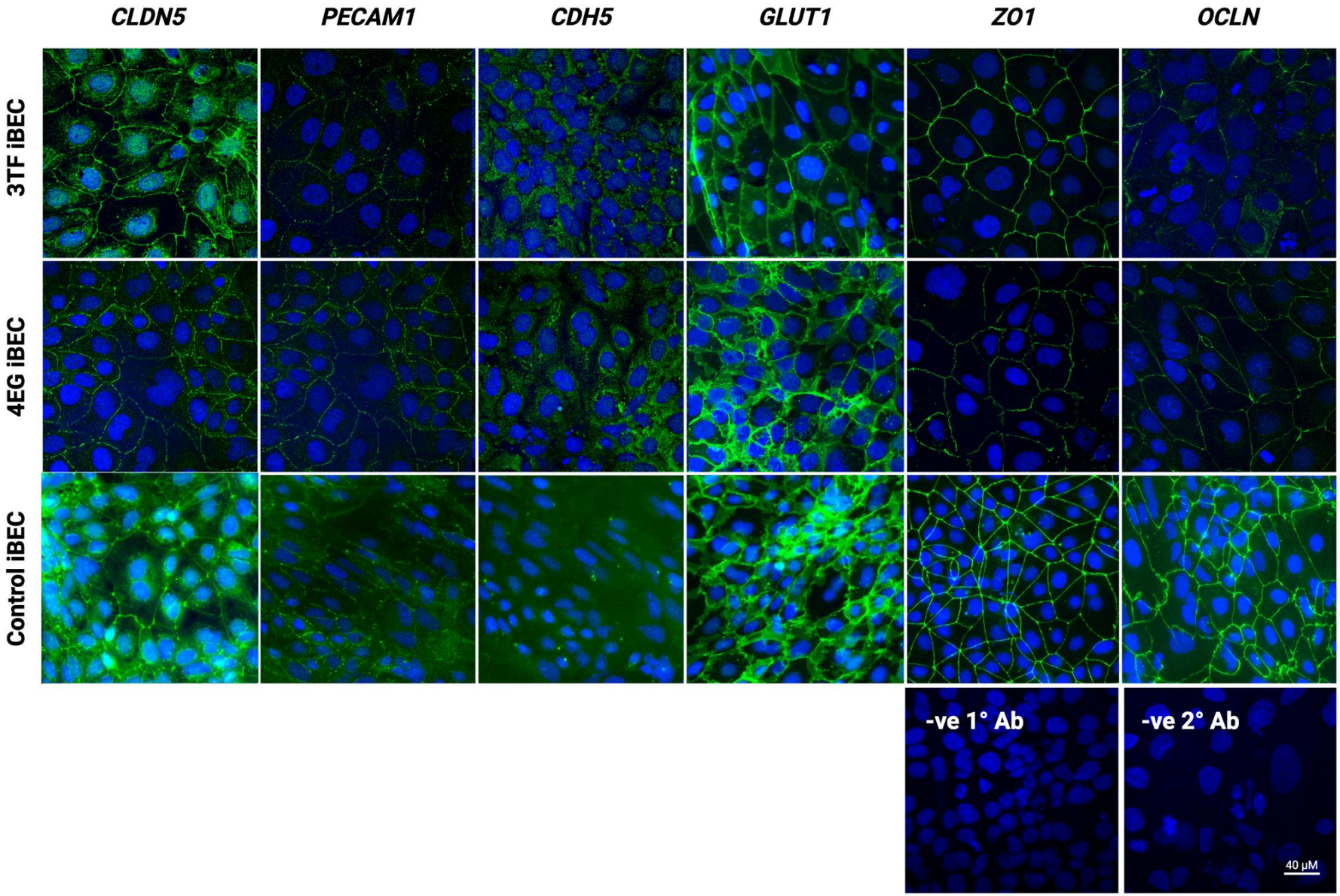

CLDN5
PECAM1
CDH5
GLUT1
ZO1
OCLN
3TF iBEC
4EG iBEC
Control iBEC
-ve 1° Ab
-ve 2° Ab
40 µM

Figure 3.5: Immunofluorescence Analysis of Brain Endothelial Cell Marker Expression and Junctional Localization in CRISPRa-differentiated iBECs. Immunofluorescence staining of iBECs derived from hAF hiPSCs. The iBECs were generated using two protocols: The abbreviated protocol (Refer to **Figure 3.3b**) was used to generate 3TF-SAM and 4EG-SAM, and the standard differentiation protocol (refer to section 2.3.1) was used to generate control iBEC cells. Immunostaining (green) evaluated the expression and localization of key blood-brain barrier (BBB) hallmarks such as tight junction proteins (*CLDN5, ZO1, OCLN*), adherens junction protein (*CDH5*), the endothelial adhesion molecule (*PECAM1*), and the BBB-specific glucose transporter (*SLC2A1; GLUT1*). Nuclei were counterstained with Hoechst 33342 (blue) dye. Negative controls omitting either the primary antibody (-ve 1° Ab) or the secondary antibody (-ve 2° Ab) were included to confirm reagent specificity and to rule out autofluorescence. Scale bar = 40 μM.

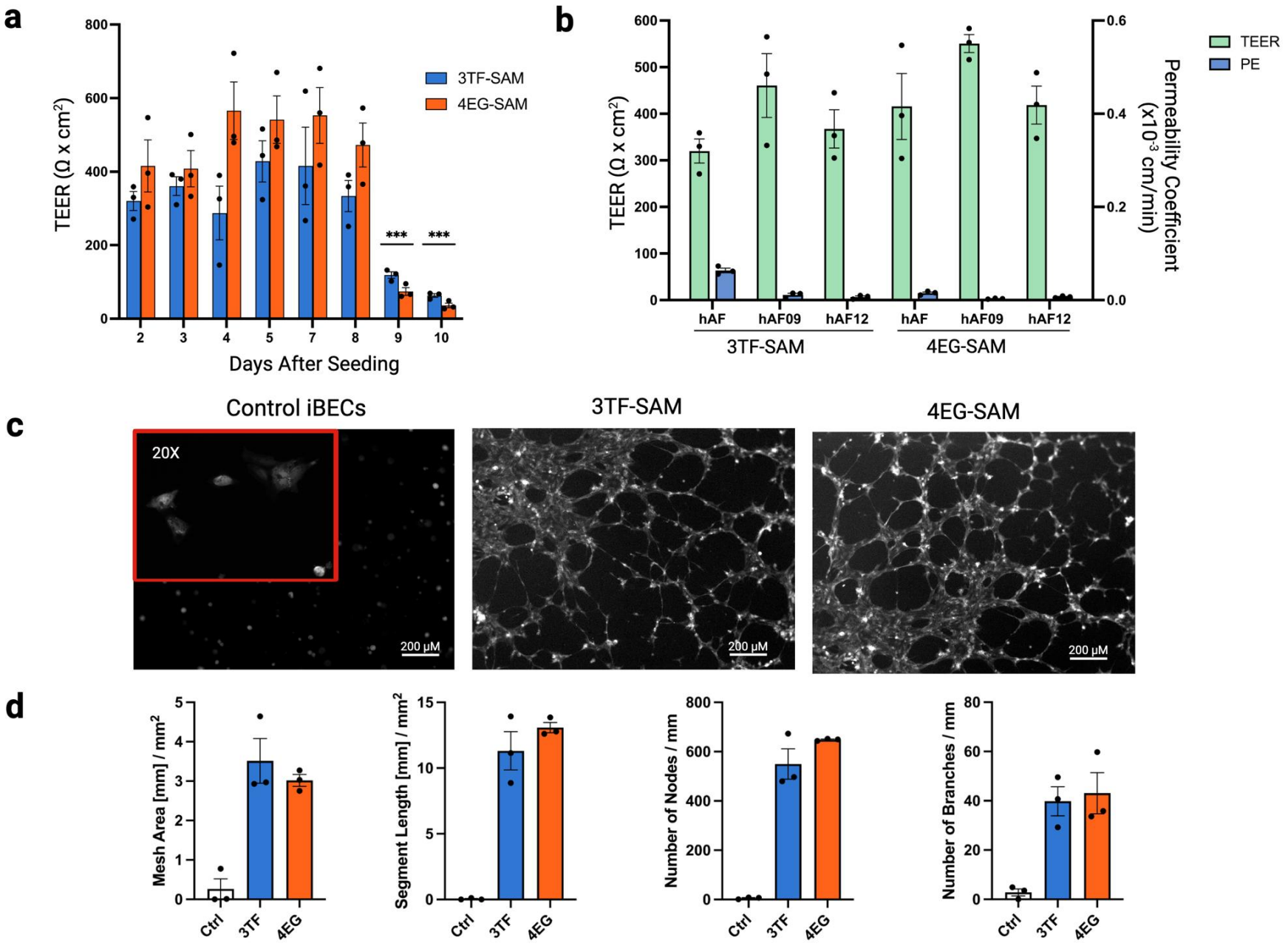

a
TEER (Ω x cm²)
800
600
400
200
0
3TF-SAM
4EG-SAM
***
***
2
3
4
5
7
8
9
10
Days After Seeding
b
TEER (Ω x cm²)
600
500
400
300
200
100
0
0.6
0.4
0.2
0.0
Permeability Coefficient (x10⁻³ cm/min)
TEER
PE
hAF
hAF09
hAF12
hAF
hAF09
hAF12
3TF-SAM
4EG-SAM
c
Control iBECs
3TF-SAM
4EG-SAM
20X
200 µM
200 µM
200 µM
d
Mesh Area [mm] / mm²
Segment Length [mm] / mm²
Number of Nodes / mm
Number of Branches / mm
Ctrl
3TF
4EG

Figure 3.6: Functional Characterization of Barrier Integrity, Selective Permeability and Angiogenic Potential in iBECs.

(a) Transendothelial electrical resistance (TEER) measurements over a 10-day post-seeding period demonstrates the dynamics of barrier integrity in 3TF-SAM and 4EG-SAM iBECs derived from hAF hiPSCs. Stable TEER values were observed from days 2 through 8, with TEER values declining significantly on days 9 and 10. Statistical significance was determined using a two-way ANOVA with Sidak's multiple comparison test, comparing each time point to day 2. Asterisks denote statistical significance (*** $p < 0.0001$). No significant differences were observed between iBECs derived from the 3TF- and 4EG-SAM iPSCs. Thus, optimal barrier characteristics are observed from approximately day 4-7 post-seeding. Bar graphs (mean ± SEM) present percentage positive from 3 biological replicates (n=3).

(b) Comparison of TEER values and paracellular diffusion coefficient (P_e) determined using a small molecule tracer sodium fluorescein (NaFl). The graph shows the formation of a tight, highly selective barrier, characterized by high TEER and low P_e. Data includes iBECs derived from three distinct hiPSC cell lines (hAF, hAF09, hAF12). Bar graphs (mean ± SEM) were generated from three biological replicates (n=3).

(c) Angiogenesis assay comparing tubule formation capability of iBECs derived from the standard differentiation (Control) protocol and the abbreviated protocol (3TF-SAM and 4EG-SAM). Cells were plated on undiluted Matrigel and induced with VEGF to assess their potential to form branched, capillary-like networks. Cells were stained with CMFDA and visualized with an Axiovert 200M microscope (Zeiss).

(d) Quantitative metrics analyzing the results from the angiogenesis assay. The bar graphs present the structural complexity of the formed networks, including mesh area, segment length, nodes and branches for iBECs differentiated under various protocols. Analysis was performed with Angiogenesis Analyzer plugin on Fiji Software, and the data are normalized to the 10X field of view in mm. Bar graphs (mean ± SEM) present percentage positive from 3 biological replicates (n=3).

3.6 Transcriptomic analysis of 3TF-SAM and 4EG-SAM iBECs

To investigate the global transcriptional changes associated with CRISPRa-mediated differentiation, RNA-sequencing analysis was performed on CRISPRa iBECs and iBECs derived from the standard protocol. When comparing the transcriptomic profile between 3TF-SAM and 4EG-SAM iBECs (**Figure S6**) we found that the majority of variation (62.5%) was attributable to differences between the cell lines, while approximately 20% variation was observed between samples. Next, we assessed CRISPRa-differentiated iBECs (abbreviated protocol) against control iBECs (standard differentiation protocol) revealing approximately 3,000 differentially expressed genes (**Figure 3.7a**). Among the most significantly upregulated genes in both 3TF-SAM and 4EG-SAM iBECs shared robust upregulation of *GJA1*[443–445] (Connexin 43) and *RGS5*[446] which are implicated in barrier communication and stabilization (Table S1 and S2). Meanwhile, we observed upregulation of *CTSV*[447,448] (Cathepsin V), which has been reported to be decreased in endothelial senescence, leading to *ALDH1A2* accumulation and p21 arrest. The upregulation of *CTSV* may reflect a mechanism that counteracts cell senescence and maintains blood vessel formation and repair during early differentiation. On the other hand, an epithelial marker *KRT7* (Keratin 7) was drastically suppressed in CRISPR-iBECs compared to the controls. Together this upregulation of barrier-specific markers and suppression of non-endothelial signatures show that CRISPR-based activation promotes unique brain endothelial lineage commitment and marker expression patterns compared to the standard protocol. While activation of these genes suggests reinforcement of brain endothelial-associated lineages, this interpretation remains preliminary and warrants further investigation into the involvement of the respective signaling pathways.

Comprehensive transcriptomic analysis examining endothelial and mesodermal gene networks is necessary to confirm whether CRISPR-based activation promotes lineage commitment rather than solely on upregulation of individual genes.

To benchmark and discriminate the expression profile to bona fide primary human BEC and epithelial cells, global expression heatmaps were generated with CRISPRa-transformed iBECs, standard protocol differentiated iBECs (Control), primary human brain endothelial cells (hBMECs) and epithelial cells (Caco2) (**Figure 3.7b)**. The hierarchal clustering suggests that CRISPRa-transformed iBECs retain a unique transcriptional profile that retain partial overlaps between control iBECs, epithelial cells and hBMECs. However, certain regions show some variability between control iBECs, for example, some reduced expression of epithelial genes that overlap with Caco2 populations. Furthermore, to gain insight into the biological pathways associated with CRISPR-mediated differentiation, gene set enrichment analysis (GSEA) was performed relative to control iBECs (**Figure 3.7c**). In both 3TF-SAM and 4EG-SAM iBECs, strong activation of RNA processing and spliceosome pathways relative to control iBECs is observed, consistent with transcriptional remodelling. On the other hand, pathways associated with cell-matrix adhesion signalling, integrin signalling and wound responses were suppressed. These findings suggest that CRISPRa promotes an active transcriptional remodelling of cell state while reducing adhesion and stress-associated processes compared to control iBECs. This transcriptional profile may indicate a less stress-activated or less senescent cell state than that of standard-protocol-derived iBECs.

Finally, to investigate the transcriptomic identity of CRISPRa iBECs, a meta-analysis was performed to compare global gene expression profiles across a panel of related cell types,

including previously published datasets: iBECs, hiPSCs, primary brain endothelial cells (hBMECs), generic primary endothelial cells (ECs) and epithelial cells (ECs) (**Table S3**)[251,283,449–453]. Principal component analysis (PCA) revealed that CRISPRa-iBECs cluster closely with other iBEC populations with clear separation from epithelial cells (PC2) and hiPSCs (PC1), which is confirmed by k-means clustering (**Figure 3.7d**). It is evident that CRISPRa iBECs continue to group closely with other iBEC models and are distinct from primary hBMEC cell lines and other generic primary endothelial cell lines. This pattern is also evident when focusing on expression heatmaps of the top 20 differentially expressed genes across all cell populations (**Figure S8**). A similar pattern was observed when focusing on marker genes for brain endothelial cells. **Figure S9** compares the expression of 74 genes enriched in brain capillary endothelial cells and 93 genes enriched in choroid plexus cells[454]. The figure confirms that hBMECs are enriched in endothelial cell genes and depleted of choroid plexus genes relative to the other cell types. Moreover, differentiation of iPSCs to iBECs resulted in an endothelial gene expression pattern that was increasingly distant from Caco2 cells in a manner that is consistent with **Figure 3.7d**. Whereas Caco2 cells express less than half the choroid plexus genes, the parental and 4EG iPSCs exhibited weak expression of most choroid plexus genes. Again, iBEC differentiation resulted in a noticeable decrease in the expression of a certain clusters of choroid plexus marker genes, including CLIC6[455]. While iBEC cells exhibit a BBB phenotype, it is clear that iPSC-based static BBB models require further refinement to achieve a gene expression pattern that is fully consistent with endogenous hBMECs.

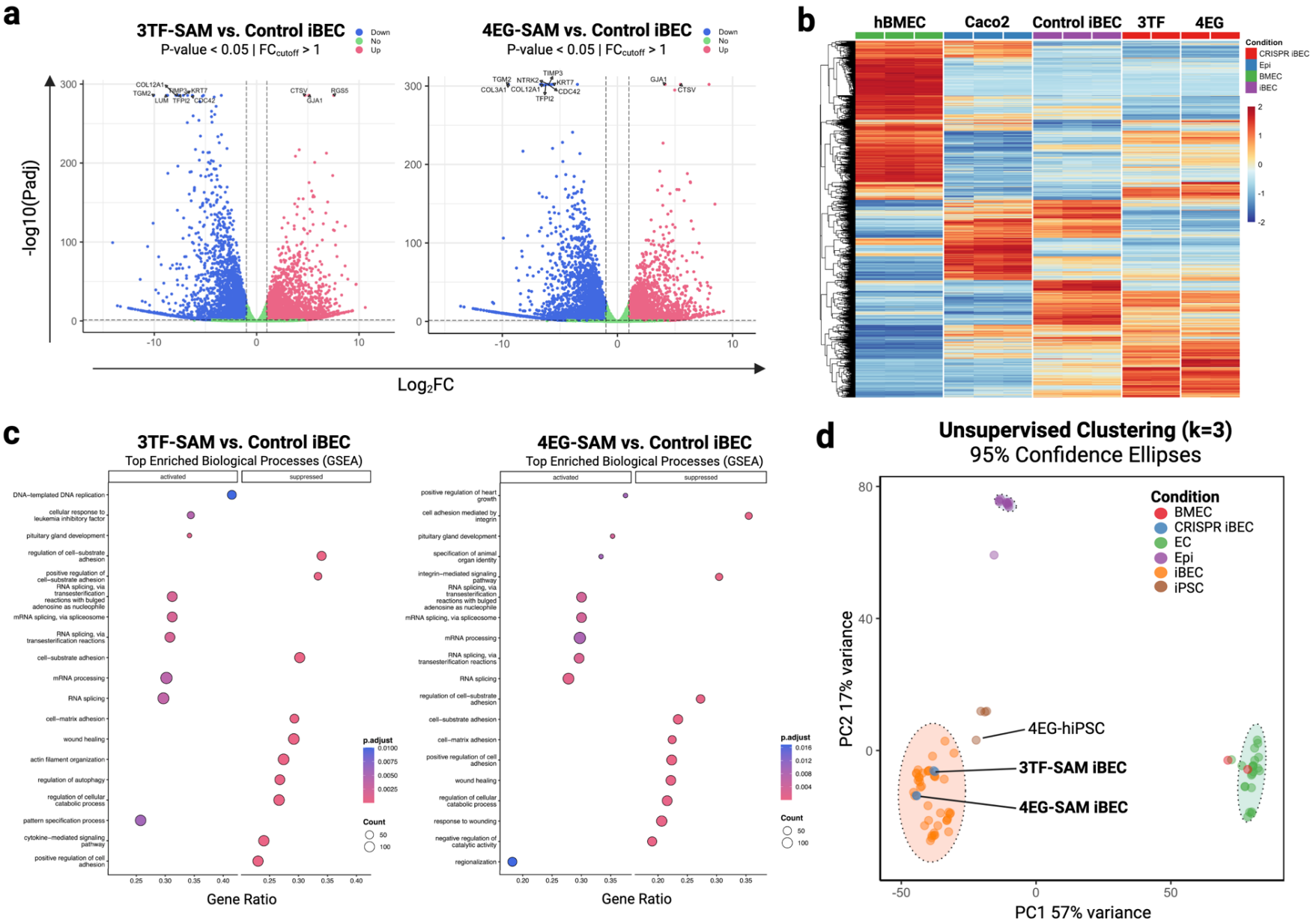

a
3TF-SAM vs. Control iBEC
P-value < 0.05 | FC$_{cutoff}$ > 1
4EG-SAM vs. Control iBEC
P-value < 0.05 | FC$_{cutoff}$ > 1
Down
No
Up
-log10(Padj)
Log$_2$FC
b
hBMEC
Caco2
Control iBEC
3TF
4EG
Condition
CRISPR iBEC
Epi
BMEC
iBEC
c
3TF-SAM vs. Control iBEC
Top Enriched Biological Processes (GSEA)
4EG-SAM vs. Control iBEC
Top Enriched Biological Processes (GSEA)
activated
suppressed
Gene Ratio
p.adjust
Count
d
Unsupervised Clustering (k=3)
95% Confidence Ellipses
Condition
BMEC
CRISPR iBEC
EC
Epi
iBEC
iPSC
4EG-hiPSC
3TF-SAM iBEC
4EG-SAM iBEC
PC2 17% variance
PC1 57% variance

Figure 3.7: Transcriptome Profiling of CRISPRa Differentiated iPSC-derived Brain Endothelial Cells.

(a) Volcano plot visualizing differentially expression analysis (DGE) between 3TF-SAM (left) and 4EG-SAM (right) iBECs compared against control iBECs. Significantly upregulated (red) and downregulated (blue) genes are defined by an adjusted p-value < 0.05 and an absolute log2 fold change > 1. Top differentially expressed genes are annotated.

(b) Hierarchical clustering heatmap of global gene expression profiles. The colour gradient represents row-scaled, log_2-transformed read counts (Z-scores) comparing primary human brain microvascular endothelial cells (hBMEC, positive control), Caco2 (epithelial control), standard control iBECs, and CRISPRa-engineered 3TF and 4EG iBECs.

(c) Gene set enrichment analysis (GSEA) dot plots illustrating the top activated and suppressed Gene (GO) biological processes in both 3TF-SAM and 4EG-SAM iBECs relative to control iBECs. Node size indicates the number of genes associated with the respective pathway, and node colour corresponds to the adjust p-value.

(d) Principal component analysis (PCA) comparing transcriptomic variance between CRISPRa iBECs against external meta-dataset of established primary endothelial cells (ECs), hBMECs, epithelial cells (Epi), standard control iBECs, and undifferentiated hiPSCs. The first two principal components (PC1 and PC2) account for 57% and 17% of the total variance respectively. Unsupervised k means clustering (k=3) was performed on the PCA coordinates with 95% confidence ellipses to identify transcriptional grouping of all analyzed cell populations. Data sources for the external meta-analysis are listed in Table S3

Chapter 4: Establishing a Site-specific and Modular Genomic Engineering Platform in hiPSCs

4.0 Findings Summary

This chapter describes the design, construction and validation of a two-system genomic engineering platform leveraging both PASTE and RMCE for stable, facile, site-specific insertion of large transgenes into hiPSCs. For this strategy, a synthetic reporter cassette and delivery protocols were iteratively optimized in HeLa cells, revealing that post-transcriptional regulatory elements (i.e., WPRE) and utilizing endotoxin-free minicircles were requirements to enable transgene expression and improved viability in sensitive hiPSC cultures. Following the proof-of-concept model, the PASTE-RMCE platform was deployed and tested in HeLa cells and hiPSCs. Furthermore, analysis of promoter-loci interactions in hiPSCs revealed a hierarchy of transcriptional activity using GFP as a reporter gene. The novel CBX3.1 promoter drove robust and consistent GFP expression across all loci, whereas the CAG promoter demonstrated high but locus-dependent expression. The EFS promoter demonstrated the weakest baseline expression. Subsequently, a single-cell sorting pipeline to isolate PASTE-edited clonal iPSC lines was generated with CAG-*AAVS1b* hiPSCs and validated with genomic DNA PCR amplification.

To assess the impact of PASTE-edited cell lines on gene silencing, pooled hiPSC cell lines were subjected to trilineage differentiation into individual germ layer cells (mesoderm, endoderm, ectoderm). Flow cytometric analysis of intracellular germ layer markers confirmed that all differentiated cells were successfully differentiated into their respective germ layers. Subsequently, live-cell tracking and endpoint flow cytometry revealed transgene expression during differentiation for both the CAG and CBX3.1 promoters,

which maintained moderate-high expression throughout the 5–7-day differentiation protocol. EFS-driven transgenes exhibited signal degradation, while endpoint flow cytometry confirmed gene silencing across most loci (particularly in endoderm lineage with less than 4% GFP+ cells). Subsequently, RMCE-mediated exchange of the GFP reporter for a therapeutically relevant CAR-BFP transgene was conducted. Cassette exchange eliminates GCV sensitivity, thereby permitting selection of edited cells.

Together, these findings establish a plug-and-play landing pad platform in hiPSCs which can be added to the regenerative medicine cell engineering toolkit, streamlining the future development of complex cell engineering applications.

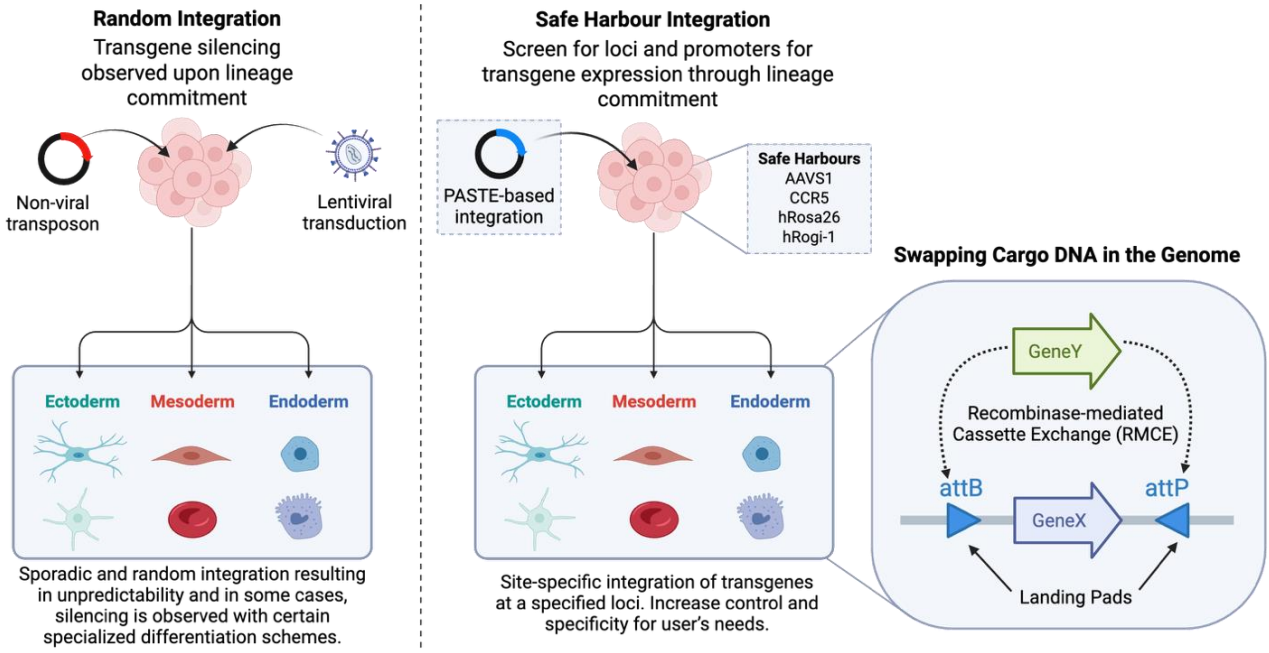


Figure 4.0: Graphical Summary of Chapter 4.

4.1 Design and Strategy for Site-specific Genomic Engineering

To establish a strategy for specific, modular and stable transgene integration into hiPSCs, we developed a hybrid platform merging two novel technologies, programmable addition via site-specific targeting elements (PASTE) (**Figure 4.1**) and recombinase-mediated cassette exchange (RMCE). While PASTE technology enables targeted integration of large transgenes without processing double-stranded DNA breaks, RMCE allows for downstream cargo swapping, enabling a plug-and-play pipeline for site-specific transgene insertion. By combining these technologies, I propose a platform for site-specific, stable and modular genomic engineering in hiPSC lines.

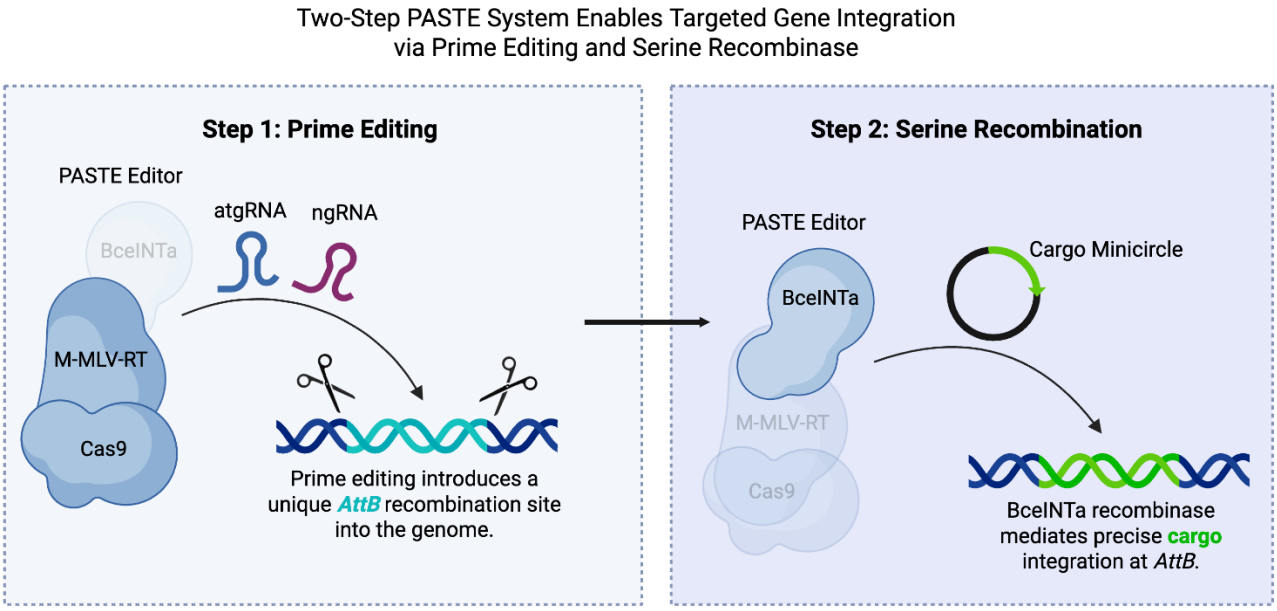


Figure 4.1: Mechanism of Programmable Addition via Site-specific Targeting Elements (PASTE). The PASTE editor is a fusion protein consisting of a Cas9 nickase, M-MLV reverse transcriptase, and a BceINTa serine recombinase. In this two-step procedure, the PASTE editor, together with the atgRNA and ngRNA, introduces a unique attB recombination site into the targeted locus via a prime editing step. Subsequently, the BceINTa paired with the donor cargo minicircle containing the complementary attP site mediates site-specific integration of a transgenes directly into the engineered attB site.

To design this platform, I first adapted, cloned and generated PASTE components optimized for site-specific delivery. The PASTE editing strategy is comprised of four major components, including: (1) a PASTE editor expression cassette (Cas9n-integrase fusion), (2) a nicking guide RNA (ngRNA), (3) an attachment guide RNA (pegRNA fused with an attachment site), and (4) a parental minicircle vector (pMini). Within this pipeline, I hypothesized that both the promoter elements and the integration site played a large role in the expression and stability of integrated transgenes. Thus, I first surveyed and identified candidate GSHs to investigate the influence of the genomic context on gene silencing and transgene expression (**Table 4.1**). As a result, I targeted three "gold-standard" GSHs (*AAVS1*, *CCR5*, and *hRosa26*) alongside the novel locus *hRogi1* (**Figure 4.2a**).

GSH gRNA spacer sequences were processed and identified through the sequence-to-function machine learning algorithm described by Yarnall et al.[331] that uses a multilayer perceptron (MLP) classifier to accurately predict atgRNA efficiency. To benchmark these predictions, I also adapted a previously validated pegRNA[456] to an atgRNA format. To minimize crosstalk between the two engineering systems, I substituted the standard Bxb1 attB site with the BceINTa attB site for the primary integration step in the atgRNA design. Moreover, to maximize integration efficiency, ngRNA were designed within a 90 bp window flanking the atgRNA-induced nick, which was previously reported to enhance prime-editing-assisted integration[331] . Accordingly, all nicking guide RNAs corresponding to each GSH site were subcloned into the pDY0209 vector backbone, while all atgRNA fragments were cloned into the pDY0663 vector.

Table 4.1: Landscape of Characterized Genomic Safe Harbour Loci.

Genomic Safe Harbour Site	Year Discovered	Gene	Genic Location	Chromosome	Citation
AAVS1	1992	PPP1R12C	Intragenic	19q	[457]
CCR5	1996/2008	CCR5	Intragenic	3p	[416,418]
COL1A	2000	COL1A	Intragenic	17q	[458]
hRosa26	2007	THUMPD3	Intragenic	3p	[414]
CLYBL	2015	CLYBL	Intragenic	13q	[428]
Chr1-eGSH	2019	-	Extragenic	1q	[431]
CNVR7	2019	-	Extragenic	3p	[432]
CNVR22	2019	-	Extragenic	8q	[432]
SHS231	2019	-	Extragenic	4p	[405]
hRogi1	2022	-	Extragenic	1q	[402]
hRogi2	2022	-	Extragenic	3p	[402]

To evaluate the efficiency of PASTE editing and downstream cargo swapping, a multifunctional reporter cassette was designed with unique promoter elements to assess expression levels driven by the transgene's regulatory elements. As a result, three promoters were selected based on their hypothesized expression in hiPSC studies, including: CAG (characterized as highly active), EFS (characterized as minimally active), and the synthetic CBX3.1 (characterized as stable and active) (**Figure 4.2b**). The CBX3.1 promoter is a proprietary synthetic promoter (by the NRC) composed of an EFS promoter sequence paired with an upstream CBX ubiquitous chromatin opening element (UCOE). The initial design of the cargo includes an EFS promoter flanked by BbsI sites for versatile Golden-gate compatible promoter swapping (for CAG and CBX3.1 promoters), a copGFP reporter, and an HSV-TK suicide gene linked to a P2A self-cleaving peptide, followed by

a terminator sequence (**Figure 4.2c**). The entire transgene is flanked by Bxb1 attachment sites (attP) containing orthogonal central dinucleotide modifications at the 5' end ("AC nucleotides) and at the 3' end ("GT" nucleotides) to ensure directionality and proper orientation for RMCE-mediated editing. This cassette utilizes both the copGFP for long-term tracking of stable cell lines and HSV-TK as a negative selection marker. In this system, when cells are treated with the prodrug Ganciclovir (GCV), apoptosis is triggered in parental PASTE-edited cells through DNA chain termination[459,460]. Therefore, this negative selection strategy can be used to isolate cells that have undergone RMCE by selectively eliminating parental PASTE-edited cells (**Figure 4.2d**).

To execute the RMCE phase, a compatible exchange donor cassette was designed to deliver a functional, therapeutically relevant transgene. This cassette comprises an EGFR-targeting chimeric antigen receptor (EGFR-CAR) linked to a blue fluorescent protein (BFP) by a P2A self-cleaving peptide, driven by an EFS promoter. To ensure targeted RMCE into the PASTE-edited landing pads, the donor transgene is flanked by Bxb1 attachment B (attB) sites. These sites incorporate the central dinucleotide modifications that correspond to the attP sites of the PASTE-integrated transgenes. This design assures that the transgene is integrated in the proper orientation while preventing inverse or improper integration events (**Figure 4.2d**).

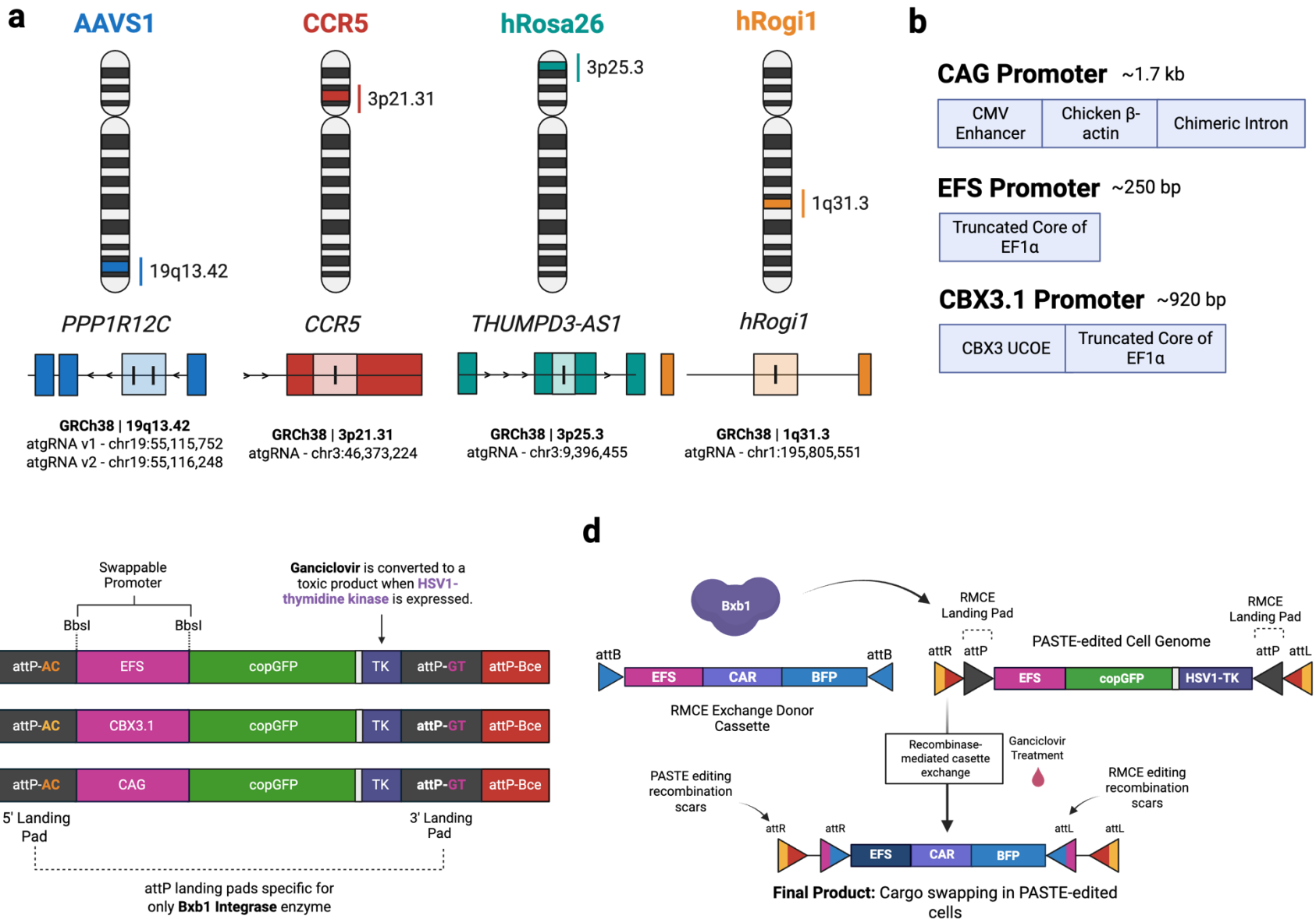

a
AAVS1
CCR5
hRosa26
hRogi1
19q13.42
3p21.31
3p25.3
1q31.3
PPP1R12C
CCR5
THUMPD3-AS1
hRogi1
GRCh38 | 19q13.42
atgRNA v1 - chr19:55,115,752
atgRNA v2 - chr19:55,116,248
GRCh38 | 3p21.31
atgRNA - chr3:46,373,224
GRCh38 | 3p25.3
atgRNA - chr3:9,396,455
GRCh38 | 1q31.3
atgRNA - chr1:195,805,551
b
CAG Promoter ~1.7 kb
CMV Enhancer
Chicken β-actin
Chimeric Intron
EFS Promoter ~250 bp
Truncated Core of EF1α
CBX3.1 Promoter ~920 bp
CBX3 UCOE
Truncated Core of EF1α
c
Swappable Promoter
BbsI
BbsI
Ganciclovir is converted to a toxic product when HSV1-thymidine kinase is expressed.
attP-AC
EFS
copGFP
TK
attP-GT
attP-Bce
attP-AC
CBX3.1
copGFP
TK
attP-GT
attP-Bce
attP-AC
CAG
copGFP
TK
attP-GT
attP-Bce
5' Landing Pad
3' Landing Pad
attP landing pads specific for only Bxb1 Integrase enzyme
d
Bxb1
RMCE Landing Pad
RMCE Landing Pad
PASTE-edited Cell Genome
attB
attB
attR
attP
attP
attL
EFS
CAR
BFP
EFS
copGFP
HSV1-TK
RMCE Exchange Donor Cassette
Recombinase-mediated casette exchange
Ganciclovir Treatment
PASTE editing recombination scars
RMCE editing recombination scars
attR
attR
attL
attL
EFS
CAR
BFP
Final Product: Cargo swapping in PASTE-edited cells

Figure 4.2 Schematic Overview of a Two-step PASTE and RMCE Platform for Modular and Site-specific Genomic Engineering in hiPSCs.

(a) Chromosomal idiograms revealing cytogenetic band and sites of integration at targeted Genomic Safe Harbour (GSH) sites (*AAVS1, CCR5, hRosa26,* and *hRogi1*). Increased resolution of the locus schematic reveals the specific integration windows for PASTE editing within the host genes (*PPP1R12C, CCR5, THUMPD3-AS1, hRogi1*) extracted and adapted from the USC genome browser. The exact GRCh38 genomic coordinates are provided for the atgRNAs utilized for the initial PASTE insertion step.

(b) Composition of transgene promoters CAG, EFS, and CBX3.1. The CAG promoter is composed of a CMV enhancer, chicken beta-actin promoter and a chimeric intron. The EFS promoter is derived from the proximal region of the EF1α promoter by taking the truncated core. The CBX3.1 promoter is a synthetic promoter composed of a CBX UCOE element upstream of an EFS promoter element.

(c) Schematic of the final transgene cargo to be integrated in the parental minicircle plasmid. The integration cargo contains a modular EFS promoter region flanked by *BbsI* restriction sites to enable downstream promoter swapping via Golden Gate assembly (for CAG and CBX3.1 promoters). Additionally, two genes (*copGFP* and an *HSV1-TK)* are used as a reporter and for downstream negative selection. To ensure directionality for Recombinase-Mediated Cassette Exchange (RMCE), the transgene is flanked by *Bxb1*-specific *attP* sites containing unique central dinucleotides (5'-AC and 3'-GT). The cassette also contains an *attP-Bce* site for the initial PASTE delivery step.

(d) The mechanism of *Bxb1*-mediated RMCE and isolation. An exchange donor cassette containing a CAR and BFP reporter is flanked by attB sites correlating to the dinucleotides from PASTE-edited cells. With the delivery of both donor cassette and Bxb1 serine recombinase, cassette exchange occurs at the genomic landing pad. Following the exchange, GCV treatment negatively selects parental PASTE-edited cells that express the HSV1-TK gene.

4.2 Iterative Optimizations to the Synthetic Reporter Transgene

To ensure effective PASTE-editing, I first had to test and empirically optimize the reporter cassette in the parental minicircle (pMini) vector to guarantee both fluorescent expression and HSV-TK activity. The initial transfections of the prototype pMini design into hiPSCs yielded negative reporter expression and insufficient TK-mediated cytotoxicity, which required iterative redesign of the transgene prior to minicircle generation.

To isolate variables limiting the expression in iPSCs, four distinct versions of the cargo cassette were systematically designed, cloned and evaluated for their expression in iPSCs (**Figure 4.3a**). The designs and alterations are detailed as follows:

- **Version 1 (V1):** The initial design of pMini utilizing a StayGold-TK fusion reporter.
- **Version 2 (V2):** To remedy inefficiencies related to translation initiation, the 5' untranslated region (UTR) was modified by removing pre-emptive ATG sequences, and an optimized Kozak consensus sequence was added to enhance ribosomal assembly.
- **Version 3 (V3):** Due to potential folding or maturation issues with the novel StayGold fluorescent protein, the well-characterized and standard copGFP gene was swapped into the cassette.
- **Version 4 (V4):** To mitigate potential transcript instability, a Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element (WPRE) and an optimized polyadenylation (poly-A) signal were subcloned at the 3' end of the cassette.

The WPRE element is known to enhance RNA stability while preventing premature transcript degradation, and GFP fluorescence with the V4 pMini vector was dramatically improved relative to previous versions (**Figure 4.3b**). Furthermore, the optimized V4 construct was functionally validated for negative selection with a cytotoxicity assay in hAF09 hiPSCs by treating cells with 20 μM GCV. IncuCyte imaging showed that GCV treatment induced cell death (reduced confluence) in pMini-V4 expressing iPSCs, whereas increased confluence was observed in GCV-treated control iPSCs and untreated pMini-V4 expressing iPSCs (**Figure 4.3c**). This confirmed that the V4 optimized transgene expressed both the copGFP reporter and HSV-TK proteins required for subsequent PASTE-editing experiments.

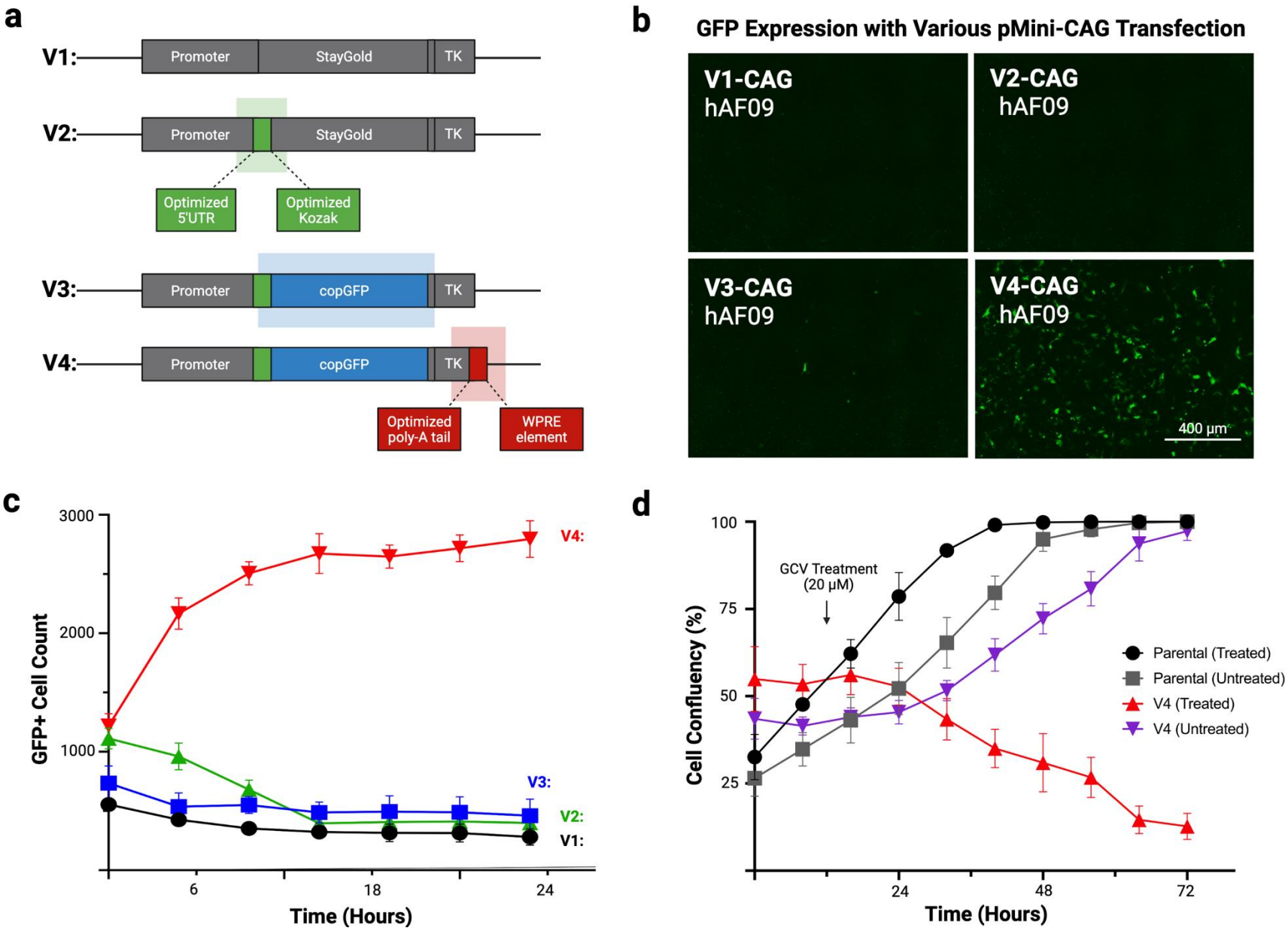

a
V1:
Promoter
StayGold
TK
V2:
Promoter
StayGold
TK
Optimized 5'UTR
Optimized Kozak
V3:
Promoter
copGFP
TK
V4:
Promoter
copGFP
TK
Optimized poly-A tail
WPRE element
b
GFP Expression with Various pMini-CAG Transfection
V1-CAG hAF09
V2-CAG hAF09
V3-CAG hAF09
V4-CAG hAF09
400 µm
c
GFP+ Cell Count
3000
2000
1000
V4:
V3:
V2:
V1:
6
18
24
Time (Hours)
d
Cell Confluency (%)
100
75
50
25
GCV Treatment (20 µM)
Parental (Treated)
Parental (Untreated)
V4 (Treated)
V4 (Untreated)
24
48
72
Time (Hours)

Figure 4.3: Iterative Optimization of Expression Cassettes to Enhance Transgene Expression.

(**a**) Schematic illustration of vector constructs from version 1 to version 4 (V1-V4). Sequential modifications of vectors include the integration of optimized 5' UTR and Kozak sequence (V2), substitution of StayGold reporter with copGFP (V3), and the addition of an optimized poly-A tail with a WPRE element (V4).

(**b**) Representative fluorescence microscopy images comparing relative GFP expression across the four vector designs with the CAG promoter construct electroporated in hAF09 hiPSCs. Scale bar: 400 μm.

(**c**) Quantification of GFP+ cell counts over 24 hours after electroporation in hAF09 and hAF12 hiPSCs. Error bars denote ± SEM, with n=2.

(**d**) Cytotoxicity assay in V4-transfected hiPSCs (hAF09, hAF12). Cells were seeded and cultured overnight, and confluency was measured and tracked for 72 hours. 20 μM of ganciclovir (GCV) was administered at the indicated time point (12 hours). V4-transfected cells demonstrate rapid cytotoxicity following GCV treatment, whereas V4 untreated and parental hiPSC control lines (treated and untreated) remain with normal proliferation.

4.3 Prototyping the Delivery and Integration of PASTE-editing

To evaluate the functionality of PASTE editing prior to implementation in hiPSCs, I opted to use Henrietta Lacks (HeLa) cells as a robust, cost-effective model to prototype and optimize the genomic modification pipeline. A recently published PASTE protocol[434] was adapted for a high throughput 96-well nucleofection format. A final optimized PASTE reaction components are detailed in **Table 4.2**.

Table 4.2: PASTE-editing Reaction Components for 96-well Nucleofection Format. Protocol adapted and modified from Fell et al. 2024[434].

PASTE Component	Concentration of DNA
PASTE V4 Editor (BceINTa)	300 ng
ngRNA plasmid	40 ng
atgRNA plasmid	40 ng
Cargo Minicircle	500 ng

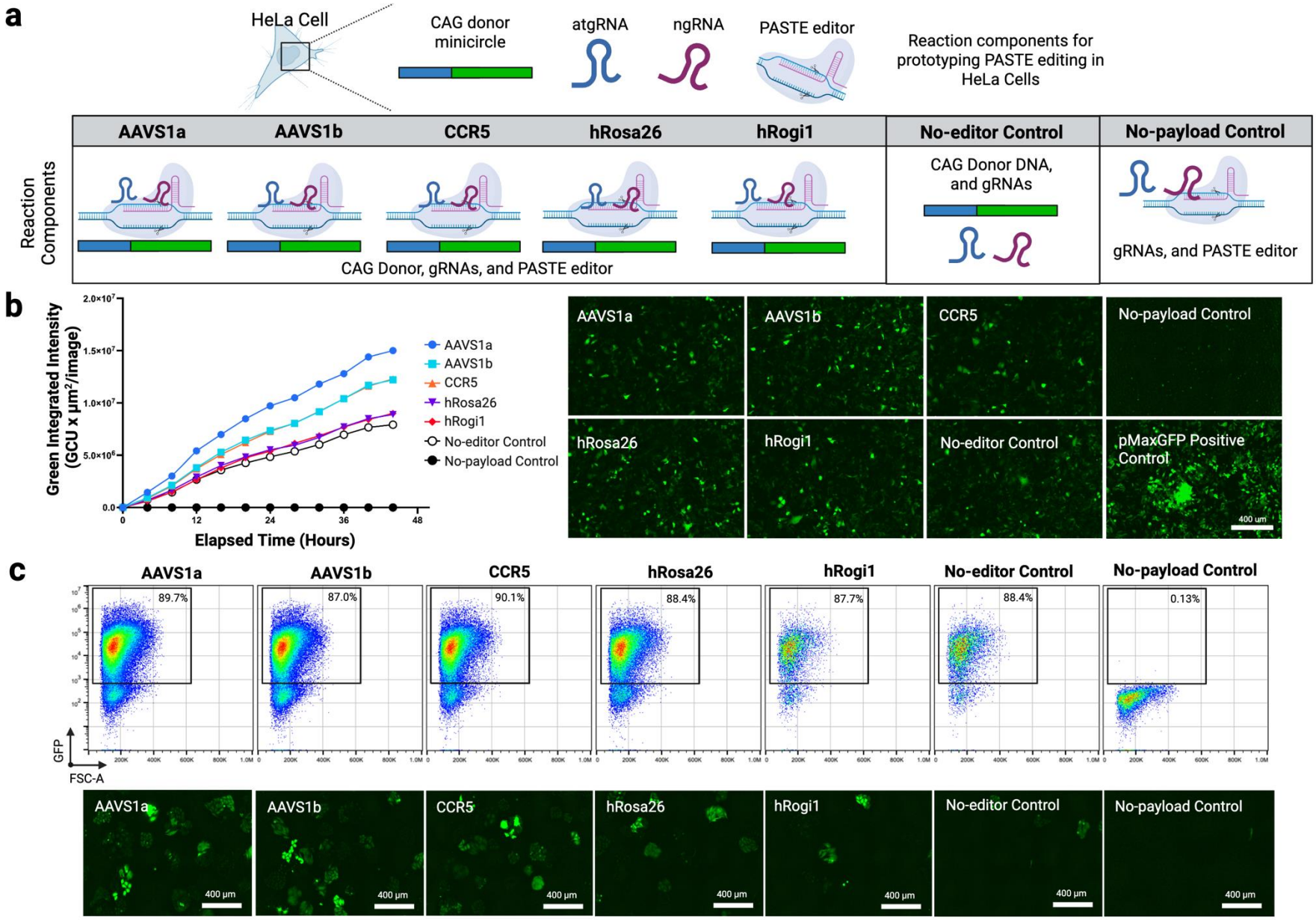
To establish a proof-of-principle of PASTE editing in HeLa cells, a single cargo minicircle (CMV) was electroporated to target five different genomic loci (*AAVS1a, AAVS1b, CCR5, hRosa26, hRogi1*). In parallel, a no-cargo control and a no-editor control were cultured (**Figure 4.4a**). Fluorescence was monitored for 48 hours post-electroporation to assess transfection efficiency and baseline GFP expression. GFP expression was observed within the first 12 hours, and continually increased over 48 hours, peaking at 1.5 x 10^7 green integrated intensity (GII: GCU x μm^2/image) across all transgene-transfected cell lines (**Figure 4.4b**). Upon reaching confluency at 48 hours, an initial round of FACS was performed to evaluate and sort GFP+ populations. As the differential signal intensity

between stably integrated PASTE-edited cells and transient episomal expression of the CAG donor minicircle was not fully characterized, a broad gating strategy was used. Approximately 60,000 cells were sorted per sample, with an average GFP positivity rate of ~89% (**Figure 4.4c**).

The sorted cells were subsequently monitored using an IncuCyte system to track fluorescence over time. As expected, with the dilution and degradation of unintegrated episomal minicircles, a dramatic reduction in GFP signal was observed after 25 hours, with sustained expression stabilizing at $2.5x10^4$ GII. Following the initial recovery phase, the HeLa cultures were maintained and expanded for approximately two weeks. During this period, it became evident that the cultures were highly heterogenous comprised of both GFP+ (edited) and GFP- (non-edited) cells. To isolate the PASTE-edited lines, a second FACS procedure was performed utilizing a stringent GFP+ gate (**Figure 4.4d**). While the no-editor control culture showed no detectable GFP+ cells for sorting, the experimental conditions revealed GFP+ populations ranging from approximately 0.5%-5.5%. These purified cell lines were then cultured for an additional month which conclusively demonstrated the generation of stable, PASTE-edited HeLa cell lines.

In parallel with PASTE editing optimizations, I aimed to develop a recombinase-mediated cassette exchange (RMCE) functionality that could be utilized following landing-pad integration. EFS and CAG PASTE-edited HeLa cells were co-transfected with the RMCE donor vector and the Bxb1 expression plasmid and cultured for 3 days. Successful cassette exchange was validated through flow cytometry analysis (**Figure 4.4e**), which showed a shift from GFP (landing pad) to BFP (cargo) expression. Finally, a TK cytotoxicity assay was performed to select for cells that have undergone cassette

exchanged (**Figure 4.4f**). Upon GCV treatment, non-exchanged PASTE-edited populations revealed decrease in GFP fluorescence and relative confluence, confirming the functionality of the integrated TK suicide gene.


a
HeLa Cell
CAG donor minicircle
atgRNA
ngRNA
PASTE editor
Reaction components for prototyping PASTE editing in HeLa Cells
AAVS1a
AAVS1b
CCR5
hRosa26
hRogi1
No-editor Control
No-payload Control
Reaction Components
CAG Donor, gRNAs, and PASTE editor
CAG Donor DNA, and gRNAs
gRNAs, and PASTE editor
b
Green Integrated Intensity (GCU x µm²/image)
Elapsed Time (Hours)
AAVS1a
AAVS1b
CCR5
hRosa26
hRogi1
No-editor Control
No-payload Control
pMaxGFP Positive Control
400 µm
c
89.7%
87.0%
90.1%
88.4%
87.7%
88.4%
0.13%
GFP
FSC-A
400 µm

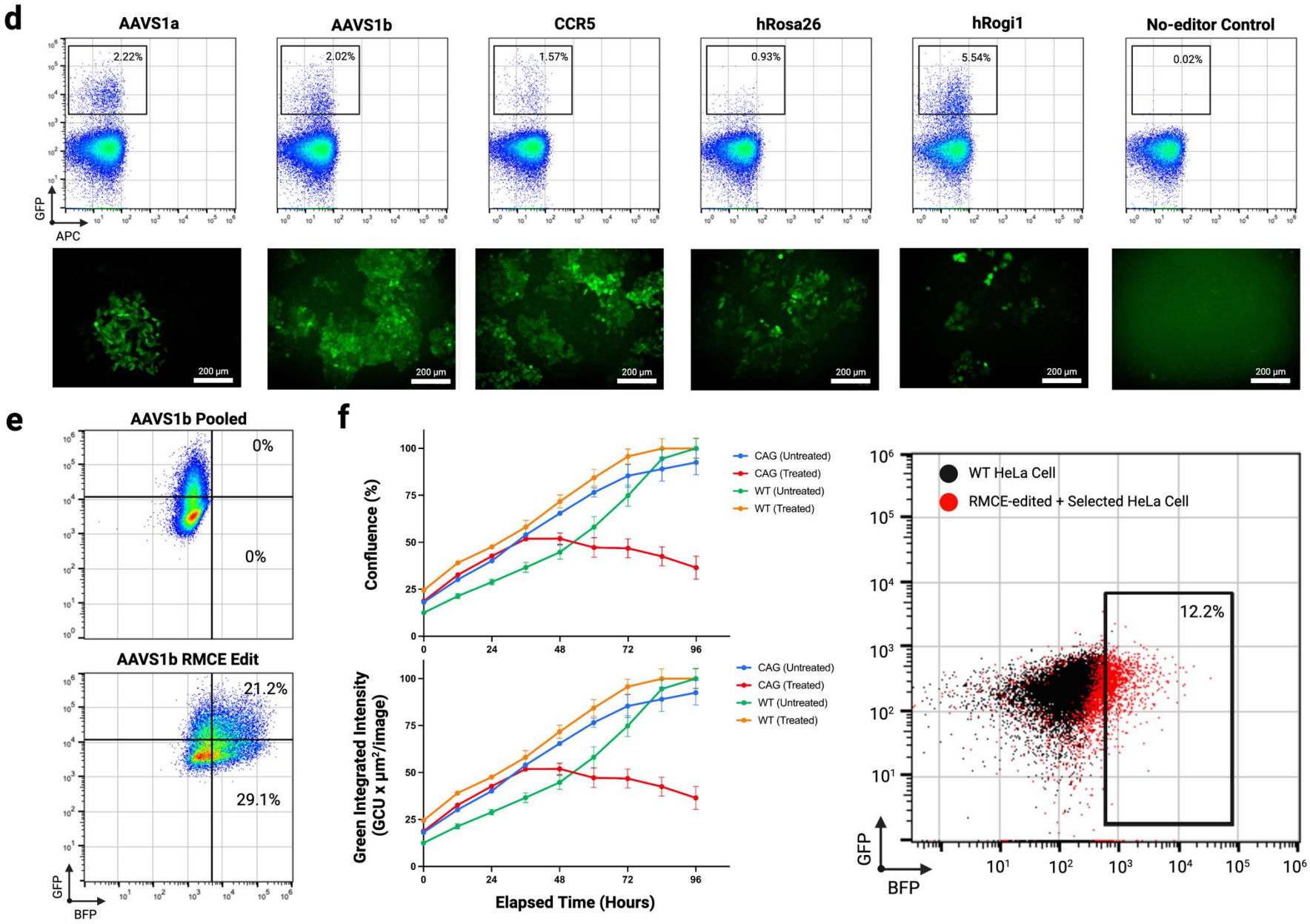

d
AAVS1a
AAVS1b
CCR5
hRosa26
hRogi1
No-editor Control
2.22%
2.02%
1.57%
0.93%
5.54%
0.02%
GFP
APC
200 µm
e
AAVS1b Pooled
0%
0%
AAVS1b RMCE Edit
21.2%
29.1%
GFP
BFP
f
Confluence (%)
CAG (Untreated)
CAG (Treated)
WT (Untreated)
WT (Treated)
Green Integrated Intensity (GCU x µm²/image)
Elapsed Time (Hours)
WT HeLa Cell
RMCE-edited + Selected HeLa Cell
12.2%
GFP
BFP

Figure 4.4: Developing a Prototype PASTE-editing and Recombinase-Mediated Cassette Exchange (RMCE) Platform in HeLa Cells.

(a) Schematic overview of reaction components and experimental design. PASTE components: CAG donor minicircle (plasmids containing no bacterial features), atgRNA, ngRNA, and PASTE editor are formatted to target individual safe harbour sites (*AAVS1a, AAVS1b, CCR5, hRosa26,* and *hRogi1*). A no-editor control (containing no PASTE editor) and a no-cargo control (containing no donor minicircle) were included in the experimental design.

(b) Fluorescent tracking of electroporated HeLa cells shows transfection efficiency (green integrated intensity; GII) over 48 hours (left). Fluorescent images (Incucyte) corresponding to 24 hr are shown on the right.

(c) First round of fluorescent-activated cell sorting (FACS) to identify and sort for pooled GFP+ populations of HeLa cells.

(d) Second round of FACS sorting to isolate for PASTE-edited GFP+ cell populations.

(e) RMCE editing in HeLa AAVS1b PASTE-edited cell. Flow cytometric analysis shows that 50.1% of cells were BFP+ (GFP+/BFP+ and GFP-/BFP+).

(f) Ganciclovir (GCV, 20 μM) selection of RMCE-edited CAG-AAVS1b and wildtype (WT) HeLa cells. GCV treatment reveals normal proliferation in wildtype and no GFP expression. Contrastingly, CAG-AAVS1b RMCE-edited cells exhibit diminished GFP expression and reduced confluence in the presence of GCV. After 2 weeks of culture and expansion, RMCE-edited cells were subjected to flow cytometric analysis, revealing a 12.2% population of BFP+ cells (red) compared to WT HeLa cells (black). All flow cytometry analyses employed a gating strategy to select for viable HeLa cells (FSC-A, SSC-A) followed by a singlet gating strategy (FSC-H vs. FSC-A) prior to marker analysis.

4.4 Optimization of the Cargo Expression Cassette in hiPSCs

Following the validation of PASTE and RMCE methodologies in HeLa cells (detailed in Section 4.2), this platform was transitioned to the hAF09 hiPSC line. Initial PASTE editing attempts utilized a CAG-driven minicircle to target GSH loci. However, preliminary results yielded undetectable GFP expression in all of the experimental conditions. hiPSCs are inherently sensitive to transfection protocols compared to robust immortalized cell lines. Therefore, it was first hypothesized that integration and transfection efficiency may require increased concentration of donor DNA (**Figure 4.5a**). Despite the increased cargo dose, GFP expression remained undetectable even at concentrations up to 1 µg. The lack of GFP signal suggested that the challenge remained in the transfection protocol or the expression environment and not the dosage of the PASTE components. Moreover, to rule out minicircle degradation and plasmid degradation, the same CAG minicircle preparation was re-validated in HeLa cells. This control experiment confirmed the structural integrity and functionality of the plasmid and yielded robust GFP expression consistent with the initial baseline validation outlined in Section 4.2.

The investigation shifted toward the minicircle generation process being incompatible for hiPSC transfection efficiency. First, the parental minicircle (pMini) vectors containing CAG promoters were electroporated to hiPSCs and transient GFP expression was observed under these conditions (**Figure 4.5b**). While the overall efficiency was expectedly lower compared to HeLa cells, the functionality of the pMini vectors in hiPSCs indicated that the minicircle generation procedure was not compatible with hiPSC transfection.

By comparing the protocols for minicircle generation, it was realized that PASTE-ready minicircles were harvested from large-scale bacterial cultures (>200 mL), whereas the

successfully expressed pMini vectors were isolated from smaller scale cultures (5 mL). Therefore, it was suspected that the maxiprep used for minicircle preparation yielded greater concentrations of endotoxin contamination, which compromised hiPSC viability and transfection efficiency. Thus, the PASTE-ready minicircles were re-cloned, purified using an endotoxin-free maxiprep system, and the structural integrity and purity were verified by Next-Generation Sequencing (NGS) (**Figure 4.5c**). Interestingly, a major peak corresponding to the correct minicircle length is observed (CAG: 5.1kb, EFS: 3.6kb, CBX3.1: 3.9kb). However, a minor enrichment peak is also observed at the 2X size of each minicircle. During the minicircle generation step, some vectors ligate with an additional minicircle generating a hybrid minicircle containing two cassettes. This prompted the use of an additional purification step with a restriction digest and re-ligation step for impure solutions. Transient electroporation of the endo-free CAG-minicircles into hAF09 hiPSCs (500 ng) successfully yielded GFP-expressing populations (**Figure 4.5d**), confirming that endotoxin contamination was problematic for minicircle expression.

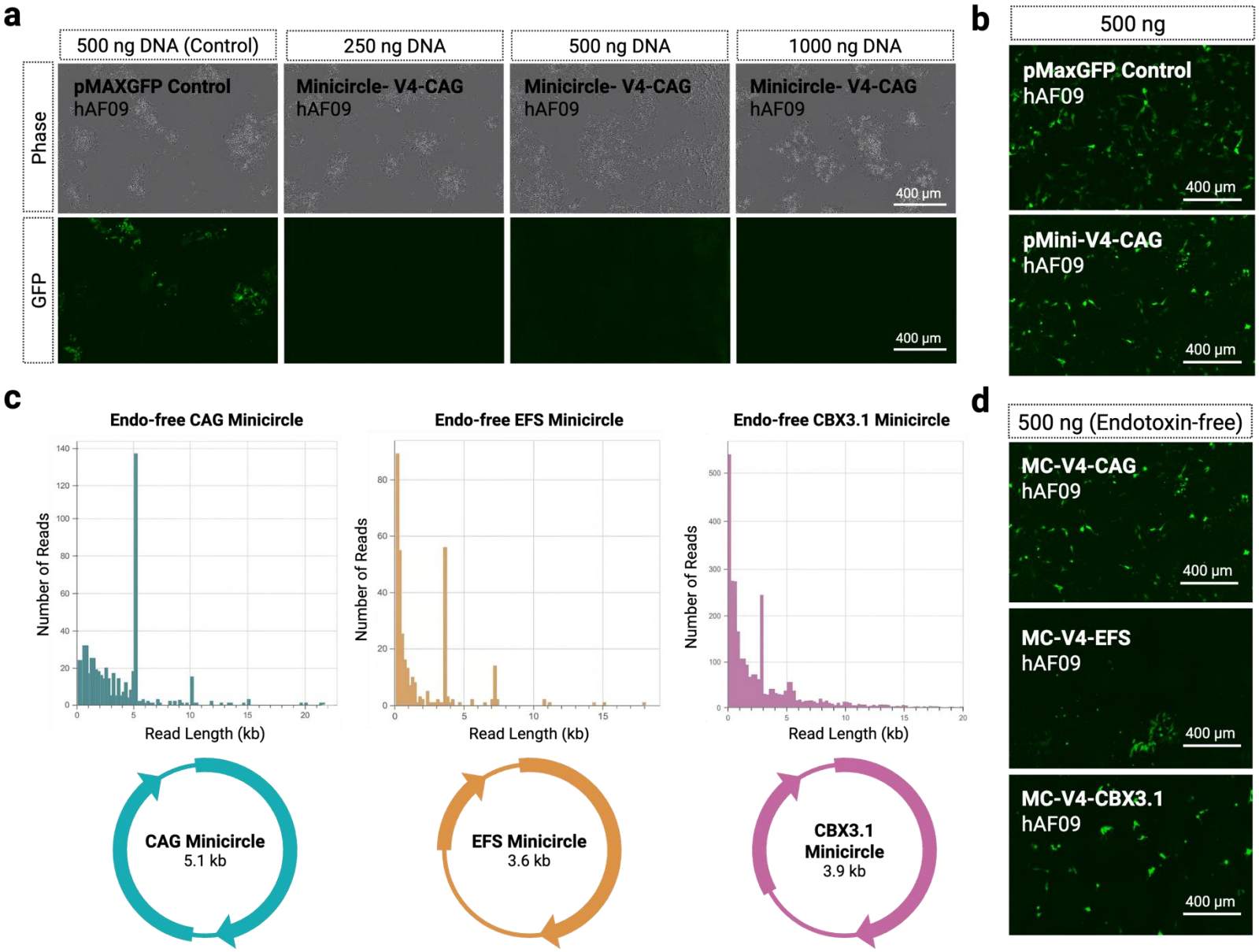

a
500 ng DNA (Control)
250 ng DNA
500 ng DNA
1000 ng DNA
Phase
GFP
pMAXGFP Control
hAF09
Minicircle- V4-CAG
hAF09
Minicircle- V4-CAG
hAF09
Minicircle- V4-CAG
hAF09
400 µm
400 µm
b
500 ng
pMaxGFP Control
hAF09
400 µm
pMini-V4-CAG
hAF09
400 µm
c
Endo-free CAG Minicircle
Endo-free EFS Minicircle
Endo-free CBX3.1 Minicircle
Number of Reads
Read Length (kb)
CAG Minicircle
5.1 kb
EFS Minicircle
3.6 kb
CBX3.1
Minicircle
3.9 kb
d
500 ng (Endotoxin-free)
MC-V4-CAG
hAF09
400 µm
MC-V4-EFS
hAF09
400 µm
MC-V4-CBX3.1
hAF09
400 µm

Figure 4.5: Optimizing Minicircle Transgene Expression in hAF09 hiPSC Lines.

(**a**) Minicircle transfection DNA dose analysis. hAF09 hiPSCs were subjected to a range of minicircle DNA concentrations to investigate whether DNA concentration impacts transgene expression. Phase-contrast (top row) and matching fluorescence (GFP, bottom row) microscopy images of hAF09 hiPSCs 24 hours after transfection are shown. Cells were transfected with an electroporation control (pMaxGFP plasmid, 500 ng) or increasing amounts of minicircle-V4-CAG (250 ng, 500 ng, 1000 ng). No fluorescence was observed for all minicircle dosages.

(**b**) Electroporation of parental minicircle (pMini) plasmid in hAF09 hiPSCs. Fluorescence microscopy images reveal GFP expression in transfected cells with control pMaxGFP (500 ng, top) and pMini-V4-CAG minicircle (500ng, bottom) plasmids.

(**c**) Minicircle vector purity assessment through next-generation sequencing. Histograms illustrating relative read-length distributions of newly synthesized endotoxin-free minicircles for CAG, EFS and CBX3.1 promoters. Corresponding circular vectors and expected size are presented below each histogram with matching colours.

(**d**) Fluorescence microscopy images of endotoxin-free minicircles transfected in hAF09 hiPSCs. GFP expression was observed using endotoxin-free minicircle vectors driven by CAG (top), EFS (middle), and CBX3.1 (bottom) promoters at 500 ng doses.

4.5 PASTE Editing in hiPSCs and Development of a PASTE-edited Clonal Cell Line Pipeline

The initial design of the project was to evaluate a broad panel of promoters for the generation, expansion and validation of PASTE-edited hiPSCs. However, time and resource constraints necessitated a narrowing of the project scope to edit cell lines and characterize promoter-locus effects on transgene expression. As a result, I selected the CAG promoter to serve as a benchmark of a strong promoter in hiPSCs (based on historical observation), EFS as a compact moderate-low expressing promoter, and CBX3.1 as a novel promoter with robust expression. Prior to genomic integration, transient minicircle screening was performed to establish baseline transcriptional strength of the GFP-TK cassette driven by a panel of promoters (CAG, CMV, CBX3.1, EFS, EF1α, and PGK) in hiPSCs. The cells were electroporated with equimolar minicircle concentrations, and a range of GFP expression levels was observed across the promoter panel (**Figure 4.6a**). In hiPSCs, the viral CMV promoter and novel CBX3.1 promoter produced robust GFP expression, CAG produced strong/moderate expression, and EFS, EF1α, and PGK promoters produced comparatively weak expression levels.

Subsequently, PASTE-editing was performed on hAF09 hiPSCs by electroporation with the reaction components described in **Table 4.2**. Here, the CAG, CBX3.1 and EFS cargo minicircles were used to target AAVS1b, CCR5, hRosa26 and hRogi1. Following electroporation, the cells were cultured for 3-4 days until confluency was reached. The PASTE-edited cells were then subjected to FACS, and pooled PASTE-edited populations were expanded into cultures for subsequent analysis.

While baseline transfection efficiencies were comparable across samples, analysis 3-5 days post-electroporation revealed distinct promoter- and locus-dependent differences in GFP expression levels (**Figure 4.6b**). Notably, the CBX3.1 promoter cassette exhibited robust and consistent expression across all loci. Distinct GFP+ sorted populations were observed for *AAVS1*b (38.8%), *CCR5* (44.1%), *hRosa26* (39.8%), and *hRogi1* (39.9%), with histogram analysis showing uniform GFP+ populations compared to wildtype parental hiPSCs (~67% - 68.5% between all conditions). On the other hand, the CAG promoter demonstrated robust GFP expression; however, it had decreased intensity compared to CBX3.1. For example, *AAVS1*b (31.5%), *hRosa26* (41.5%), and *hRogi1* (34.8%) yielded higher GFP-positive populations, while *CCR5* (13.7%) had drastically reduced editing efficiencies. This variability is also observed in the distribution of the peaks in the CAG histogram overlays. Lastly, the EFS promoter cassette resulted in lower GFP expression across all tested target sites. *AAVS1*b, *CCR5*, *hRosa26*, and *hRogi1* all hovered around 1.8% - 2.63% of GFP+ sorted cells, and approximately 10% of the cells were GFP+ compared to parental wildtype cells.

Additionally, the mean fluorescence intensity of GFP-positive populations was compared between the transgene promoters for the *AAVS1b* site. The analysis reveals a distinct hierarchy in promoter activity with PASTE-edited hiPSCs (**Figure 4.6c**). The CBX3.1 promoter seemingly drives the most robust level of transgene expression per cell, yielding the highest overall MFI (140,208 $\pm$ 17,333). Contrastingly, the CAG promoter generated moderate-high expression (25,969 $\pm$ 2,704) with a peak that was substantially narrower than CBX3.1. The EFS promoter exhibited the weakest overall transcriptional activity (13,077 $\pm$ 389).

Moreover, while pooled cell line generation and analysis were accomplished, the isolation of clonal hiPSC lines is essential to minimize variability, guarantee uniform genome integration and establish standardized models for downstream therapeutics development. Therefore, I established a framework and a proof-of-concept pipeline for single-cell isolation and clonal expansion of PASTE-edited hiPSCs with the CAG promoter lines as the model. Based on the optimized PASTE protocol, newly transfected PASTE-edited CAG hAF09 hiPSCs were electroporated and plated on 6-well plates and cultured until confluency. During this recovery window, GFP expression was monitored daily to evaluate the clearance of episomal expression and stabilization of GFP (**Figure 4.6d**). On day five, FACS sorting was used to plate individual GFP+ cells in a 96-well plate. To promote survival and attachment, sorted hiPSCs were cultured in eTeSR media containing 10 μM of ROCK inhibitor and 0.5X CloneR2 for 48 hours, with daily media changes thereafter. Approximately 3 weeks later, cells were assessed for GFP fluorescence and viability (**Figure 4.5e**). This revealed that approximately 11% of the sorted cells were viable GFP+ colonies. Additionally, 2% of the sorted cells were viable GFP- colonies. The remaining 87% of the wells failed to yield viable hAF09 colonies. The GFP+ colonies were subsequently expanded and cultured for another 2 weeks for characterization and banking.

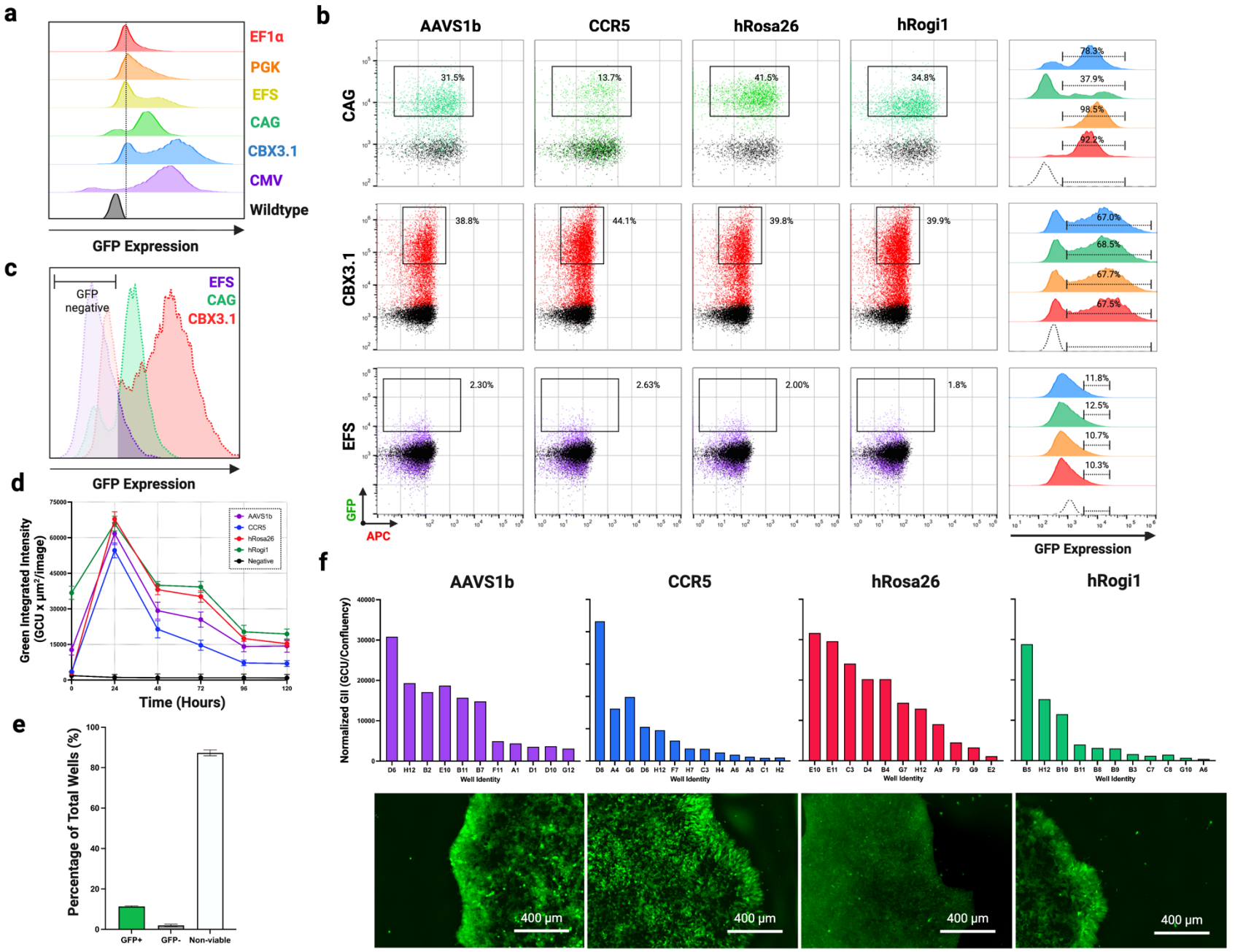

a
EF1α
PGK
EFS
CAG
CBX3.1
CMV
Wildtype
GFP Expression
b
AAVS1b
CCR5
hRosa26
hRogi1
CAG
CBX3.1
EFS
31.5%
13.7%
41.5%
34.8%
78.3%
37.9%
98.5%
92.2%
38.8%
44.1%
39.8%
39.9%
67.0%
68.5%
67.7%
67.5%
2.30%
2.63%
2.00%
1.8%
11.8%
12.5%
10.7%
10.3%
GFP
APC
GFP Expression
c
GFP negative
EFS
CAG
CBX3.1
GFP Expression
d
Green Integrated Intensity (GCU x µm²/image)
AAVS1b
CCR5
hRosa26
hRogi1
Negative
Time (Hours)
e
Percentage of Total Wells (%)
GFP+
GFP-
Non-viable
f
AAVS1b
CCR5
hRosa26
hRogi1
Normalized GII (GCU/Confluency)
Well Identity
400 µm

Figure 4.6: PASTE Editing for Pooled and Clonal hiPSCs Cell Line Generation.

(**a**) Promoter strength assessment in hAF09 hiPSCs. Cells were electroporated with equimolar amounts of minicircles with various promoters driving copGFP-TK expression. Ridgeline plots reveal variability in fluorescence intensity. All flow cytometry data were pre-gated for live hAF09 hiPSCs (FSC-A, SSC-A) and singlets (FSC-H, FSC-A) prior to analysis.

(**b**) Flow cytometric analysis of pooled hAF09 hiPSCs PASTE-edited with the CAG, CBX3.1 and EFS promoter constructs across four GSH sites (AAVS1b, CCR5, hRosa26, hRogi1). Dot plots indicate the cell sorting gating strategy for GFP+ populations, while ridgeline plots (right) show the total GFP-positive populations for each locus.

(**c**) Overlay histograms comparing fluorescence expression profiles of CAG, EFS, and CBX3.1 promoter constructs in pooled PASTE-edited hiPSCs.

(**d**) GFP expression levels after electroporation of CAG-cargoes targeting various GSHs. IncuCyte long-term imaging reveals the clearance of episomal expression over 120 hours prior to clonal cell line generation.

(**e**) Efficiency of single cell sorting and colony formation of hiPSCs. Cells were individually seeded in 96-well plates and cultured for 3 weeks until colony formation was observed.

(**f**) Clonal expansion and quantification of stable GFP expression with the CAG promoter. Bar charts show normalized green integrated intensity (GII/confluence) across each clonal well targeting each locus, revealing variance in GFP expression between clones after 1 month. Fluorescence imaging (bottom) confirms the morphology and stable GFP expression of the corresponding clonal lines at each GSH locus of the highest peak (AAVS1:D6, CCR5: D8, hRosa26: E10, hRogi1: B5).

4.6 Genomic Validation of GSH Targeted Edits

To validate PASTE-edited hiPSCs, an initial genomic validation experiment was performed by PCR analysis of genomic DNA (gDNA). Genomic DNA was extracted from PASTE-edited hiPSCs, and PCR was performed to confirm site-specific insertion. Primer pairs were designed to validate specific integration regions at the 5' and 3' GSH sites to indicate integrase-mediated recombination at the attB/attP sites (**Figure 4.7a**). The amplification strategy used three primer set configurations:

- Set A (Junction-spanning): Forward primer targeting the endogenous GSH integration region paired with a reverse primer targeting the 5' end of the integrated cargo. This amplicon can confirm the genomic location of the integration.
- Set B (Locus-flanking): Forward primer and reverse primer targeting the perimeter of the GSH region. This set analyzes the disruption of the endogenous safe harbour site. Thermocycling extension conditions were set to amplify shorter fragments (1-2 kb), which permits the amplification of undisturbed wild-type alleles only and prevents amplification of regions bearing a large transgene (4-6 kb), which serves as an indicator for GSH disruption.
- Set C (Internal cargo): Forward and reverse primers entirely within the copGFP sequence in the minicircle. This set verifies the presence of the integrated cargo.

First, control PCR amplifications were performed on wildtype parental hAF09 hiPSCs to verify the integrity of the GSH loci and to validate primer binding efficiency. All reactions were carried out with a high-fidelity polymerase (Platinum SuperFi II DNA Polymerase, Thermo Fisher) to ensure amplification and were visualized through gel electrophoresis.

As expected, Set B reactions generated 500-1000 bp in all wildtype cells, confirming undisturbed GSH sites, while Set A and Set C remain empty (**Figure 4.7b**). PASTE-edited pooled cell lines exhibited expected amplicon sizes with Set A (450-500 bp) with some signals observed in Set B and amplification observed in Set C (**Figure 4.7c**). Bands were observed in set A in PASTE-edited cell lines, indicating successful integration of the transgene at the correct *AAVS1b* locus. The expected amplicons were also observed with Set B, indicating that the pooled cell mixture contained varying ratios of edited and non-edited cells with potentially fluctuations in transgene copy. On the other hand, clonally generated PASTE-edited CAG-AAVS1b cell lines revealed no bands with Set B, indicating that all wildtype *AAVS1b* sites have been successfully edited (**Figure 4.7d**). The thermocycling extension time was insufficient to amplify the locus containing the large cassette, which further confirms the insertion of the transgene through PASTE.

Table 4.3: Expected Amplicon Length for gDNA PCR Validation of PASTE-edited Cell Lines.

Primer Set	Locus	Expected WT Amplicon Length (bp)	Expected PASTE-edited Amplicon Length (bp)
A (5' junction)	*AAVS1b*	-	515
	CCR5	-	420
	hRosa26	-	483
	hRogi1	-	495
B (WT allele)	*AAVS1b*	475	-
	CCR5	432	-
	hRosa26	958	-
	hRogi1	506	-
C (Insert)	*AAVS1b*	-	~350
	CCR5	-	~350
	hRosa26	-	~350
	hRogi1	-	~350

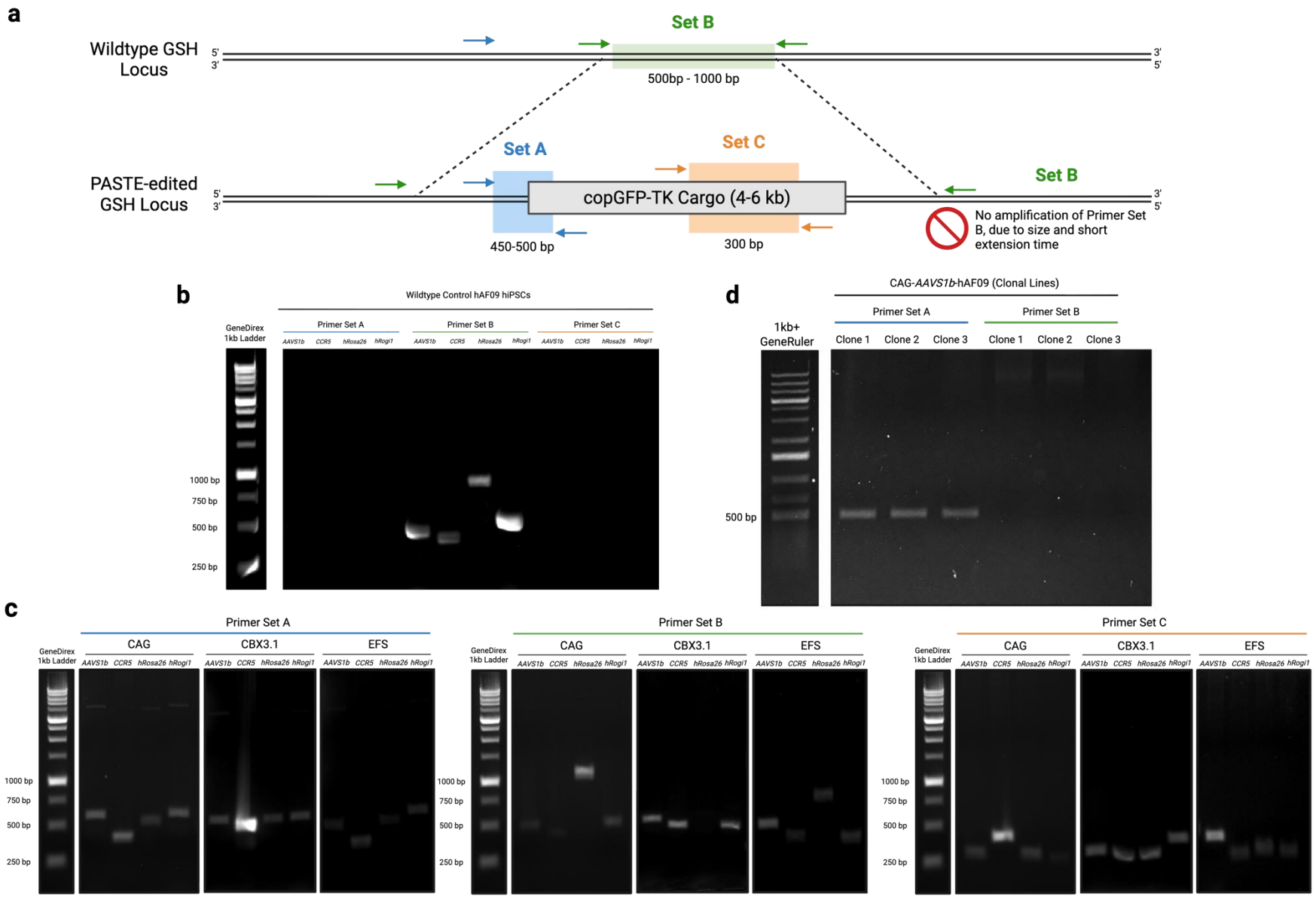

a
Wildtype GSH Locus
5'
3'
Set B
500bp - 1000 bp
3'
5'
PASTE-edited GSH Locus
Set A
Set C
Set B
copGFP-TK Cargo (4-6 kb)
450-500 bp
300 bp
No amplification of Primer Set B, due to size and short extension time
b
Wildtype Control hAF09 hiPSCs
GeneDirex 1kb Ladder
Primer Set A
Primer Set B
Primer Set C
AAVS1b CCR5 hRosa26 hRogi1
1000 bp
750 bp
500 bp
250 bp
d
CAG-AAVS1b-hAF09 (Clonal Lines)
1kb+ GeneRuler
Primer Set A
Primer Set B
Clone 1 Clone 2 Clone 3
500 bp
c
Primer Set A
Primer Set B
Primer Set C
GeneDirex 1kb Ladder
CAG
CBX3.1
EFS
AAVS1b CCR5 hRosa26 hRogi1
1000 bp
750 bp
500 bp
250 bp

Figure 4.7: Genomic DNA (gDNA) Validation of PASTE-edited hiPSCs.

(**a**) Schematic representation of genomic DNA PCR strategy. Set A primers were designed to target both the GSH locus and the integrated cargo cassette (450-500 bp). Set B spans the entire region of integration on the GSH locus, amplifying the wildtype locus (500-1000 bp), while the edited locus fails due to large size of the cargo (4-6 kb) combined with short PCR extension times (0bp). Set C specifically amplifies the insert sequence (300-400 bp).

(**b**) Genotyping unedited wildtype hAF09 hiPSCs revealing exclusive locus spanning Set B primers.

(**c**) PCR validation of PASTE-edited hAF09 hiPSCs (CAG, CBX3.1, and EFS) targeting unique GSH loci (*AAVS1b, CCR5, hRosa26,* and *hRogi1*). Amplification with Sets A and C confirms cargo integration, while the presence of Set B bands indicates unedited alleles in pooled populations.

(**d**) Genotyping clonally isolated CAG-*AAVS1b* hAF09 hiPSC cell lines. Presence of integration specific band (Set A) and absence of wildtype band (Set B) confirms a more uniform integration of the cargo in the clones.

4.7 Gene Silencing in PASTE-edited Trilineage Germ Layer Differentiations

To assess the dependence of gene expression on locus and promoter factors and their impact during the differentiation process, a panel of PASTE-edited hiPSCs was subjected to trilineage differentiation into the three germ layer lines (mesoderm, endoderm, and ectoderm). PASTE-edited hiPSCs cell lines harbouring CAG, CBX3.1, and EFS constructs inserted in the *AAVS1b, CCR5, hRosa26, and hRogi1* GSH loci were differentiated into mesoderm and endoderm (5-day protocol) and ectoderm (7-day protocol) (**Figure 4.8a**).

In order to dynamically assess the potential of gene silencing, transgene expression was monitored throughout the differentiation process using live-cell fluorescence imaging with the IncuCyte (**Figure 4.8b**). Across all lineages, expression profiles were variable and dependent on the promoter and the integration locus. Interestingly, the CAG promoter drove the most robust GFP expression across nearly all GSHs, with variability in expression based on lineage specification. For example, in the ectoderm lineage, CAG populations maintained higher expression over the 7-day protocol. However, in the mesoderm and endoderm lineages, GFP expression showed more dynamic silencing at the *CCR5, hRosa26*, and *hRogi1* loci, with peak expression observed on day 4. Interestingly, the *AAVS1b* locus maintained a steady increase in CAG-driven expression throughout the mesoderm and endoderm differentiation, showing resistance to gene silencing.

Additionally, the CBX3.1 integrated cell lines exhibited moderate GFP expression compared to CAG, with variability in silencing based on locus and lineage. In the ectoderm

lineage, GFP expression demonstrated remarkable stability, revealing a relatively consistent expression profile with little evidence of silencing. Similar to the CAG constructs, in the mesoderm and endoderm differentiations, CBX3.1-driven GFP expression was less stable, with a peak observed on day 3-4 followed by decreased expression by day 5, reflecting similar gene silencing trends observed with the CAG promoter integrated at the *CCR5*, *hRosa26*, and *hRogi1* locus.

Comparatively, the EFS promoter exhibited markedly lower baseline GFP expression across all loci and lineages and exhibited a silencing trend over time. Silencing was most prominent at the *CCR5* locus during both the endoderm and ectoderm differentiations, where EFS expression peaked (day 2-4) before silencing through the remainder of the differentiation (day 5 or 7). Wildtype controls maintained negligible baseline fluorescence throughout the differentiation process, which was as expected and used to compare the expression profiles with other cell lines.

Taken altogether, tracking the expression dynamics during the differentiation process provides insight into the shifting epigenetic landscape. Transgene expression across the *CCR5*, *hRosa26*, and *hRogi1* loci all exhibit coordinated, wave-like trajectories between days 2-4 across all three lineages before undergoing a decline. This shared expression pattern across diverse promoters implies that the expression dynamics during early lineage commitment are dictated by locus factors. On the other hand, the *AAVS1b* locus demonstrates resistance to that fluctuation and maintains a steadier increase in expression throughout differentiation. Ultimately, **Figure 4.8b** illustrates the dynamic relationship between early cell fate specification and locus-dependent accessibility for transgene expression.

At the terminal stage of the differentiation procedure, flow cytometry was performed to confirm germ layer specification using intracellular markers (**Table 4.3**) and to quantify end-stage GFP expression levels across the panel of locus and promoter combinations. Notably, all differentiated cell lines revealed pure populations of germ layer specification based on the markers used to identify lineage identity (**Figure 4.8c**). On the other hand, final GFP expression remained highly variable based on the promoter, GSH site and the differentiated lineage (**Figure 4.8d**).

Endpoint flow cytometric analysis was performed on all differentiated germ layer cells, and the analysis was grouped by their germ layer cell and the unique promoter cassette integrated. The CAG promoter produced the highest levels of expression, while its functionality was sensitive to the locus and the resulting germ layer. For example, the CAG-AAVS1b clone resulted in a more robust and stable population of GFP+ cells across the entire panel in all three lineages: mesoderm (~85%), endoderm (~90%), and ectoderm (98%) compared to WT differentiated controls. On the other hand, the transgene expression of the CAG-*CCR5*/h*Rosa26*/*hRogi1* revealed variable and low-premature silencing. For example, at these loci, the ectoderm lineage retained high GFP+ populations (ranging from 45%-80%), whereas the mesoderm and endoderm lineages resulted in silencing with detectable transgene expression dropping to 10-25% of the population.

Moreover, the CBX3.1 promoter exhibited transgene expression in all four safe harbours in the ectoderm lineage. Regardless of the site where it was integrated, the ectoderm lineage combined with the CBX3.1 promoter demonstrated resistance to gene silencing, achieving over 90% GFP+ cell population. However, in the mesoderm and endoderm

differentiated cells, GFP expression was drastically reduced to under 25% (mesoderm) and below 15% (endoderm).

Flow cytometric analysis on the EFS promoter in differentiated cells revealed populations of GFP+ cells ranging from near-background levels to up to 25%. In the mesoderm lineage, significant retention of GFP expression was observed at multiple loci when compared to WT control groups. The *hRosa26* locus revealed the highest percentage of GFP+ mesoderm cells (~22%), followed by *CCR5* (~8%) and *hRogi1* (~7%). Additionally, the *CCR5* locus maintained high GFP+ populations in the ectoderm lineage (~23%). Ectoderm differentiations at other loci resulted in populations that were near background levels (<5%). Finally, the endoderm lineage showed the most profound levels of silencing across all GSH sites. The percentage of GFP+ endoderm cells remained low in *AAVS1b, CCR5, hRosa26,* and *hRogi1,* which peaked under 4%, compared to the WT baseline. Altogether, EFS driven transgenes seem to be more susceptible to gene silencing and are more locus and cell fate dependent.

Table 4.4: Markers Assessed for Germ Layer Specification in Differentiated Cells.

Germ Layer Specification	Marker
Mesoderm	FLK1
Endoderm	CXCR4
Ectoderm	PAX6

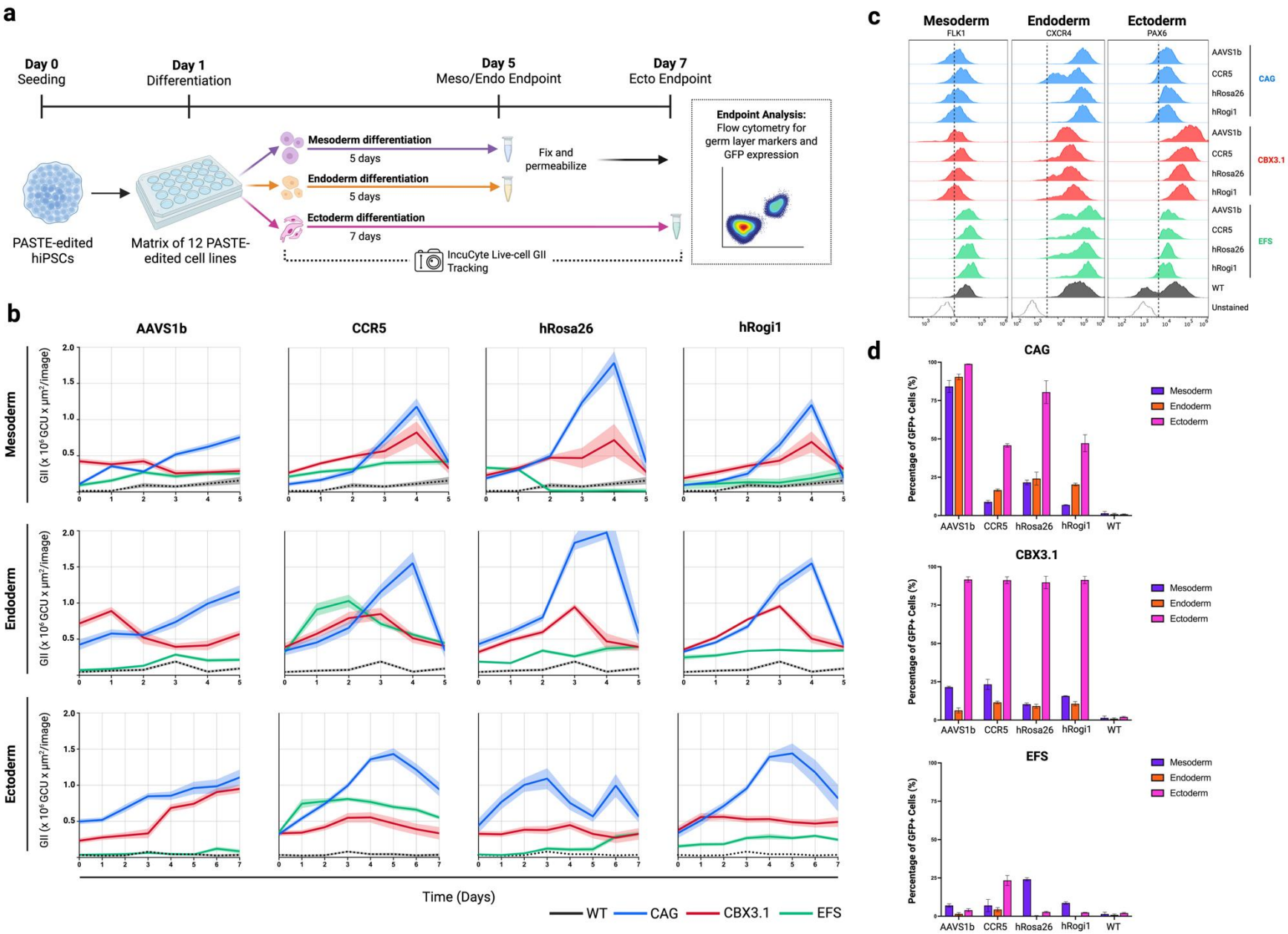

a
Day 0
Seeding
Day 1
Differentiation
Day 5
Meso/Endo Endpoint
Day 7
Ecto Endpoint
PASTE-edited hiPSCs
Matrix of 12 PASTE-edited cell lines
Mesoderm differentiation
5 days
Endoderm differentiation
5 days
Ectoderm differentiation
7 days
Fix and permeabilize
Endpoint Analysis:
Flow cytometry for germ layer markers and GFP expression
IncuCyte Live-cell GII Tracking
b
AAVS1b
CCR5
hRosa26
hRogi1
Mesoderm
Endoderm
Ectoderm
GII (x 10⁶ GCU x µm²/image)
Time (Days)
WT
CAG
CBX3.1
EFS
c
Mesoderm
Endoderm
Ectoderm
FLK1
CXCR4
PAX6
AAVS1b
CCR5
hRosa26
hRogi1
WT
Unstained
CAG
CBX3.1
EFS
d
CAG
CBX3.1
EFS
Percentage of GFP+ Cells (%)
Mesoderm
Endoderm
Ectoderm
AAVS1b
CCR5
hRosa26
hRogi1
WT

Figure 4.8: Evaluating locus and promoter dependencies of transgene silencing across a panel of PASTE-edited hiPSCs during trilineage differentiation.

(a) Experimental overview and timeline. 12 PASTE-edited hiPSC lines containing combinations of three promoters (CAG, CBX3.1 and EFS) integrated across four safe harbour loci (AAVS1b, CCR5, hRosa26, and hRogi1) were differentiated into mesoderm, endoderm, and ectoderm lineages. Transgene expression was monitored through live-cell imaging (IncuCyte), followed by a terminal flow cytometry analysis at day 5 (mesoderm and endoderm) or day 7 (ectoderm) to assess germ layer specification and transgene expression.

(b) Real-time tracking of transgene GFP expression during differentiation by live-cell quantification of green integrated intensity (GII). Data is organized by GSH locus (columns) and germ layer specification (rows). Line colours represent the promoter driving GFP expression (CAG, blue; CBX3.1; red, EFS, green), which is plotted against the wildtype unedited control differentiation (WT, dotted black line). Error bars represent SEM.

(c) Panel of flow cytometry histograms showing respective germ layer markers. Mesoderm cell lines were stained for FLK1 (column 1), endoderm cell lines were stained for CXCR4 (column 2), and ectoderm cell lines were stained for PAX6 (column 3). Wildtype hAF09 hiPSCs were differentiated and stained for each germ layer cell lines in parallel (WT). Last row shows unstained controls of each germ layer cell line (unstained).

(d) Endpoint transgene expression of differentiated PASTE-edited germ layer cells. Grouped bar chart quantifying the percentage of GFP+ population of cells in mesoderm (purple), endoderm (orange), and ectoderm (magenta) compared to unedited wildtype (WT) controls. Data are presented as mean ± SEM (n = 2 biological replicates).

4.8 Genomic Transgene Insertion by Recombinase-Mediated Cassette Exchange

PASTE editing permits the genomic insertion of large synthetic cassettes at specific safe harbour sites. Here, the landing pad was designed to further simplify the insertion of transgenes in the edited GSH sites, thereby expanding the utility of the engineered hiPSC cell lines. As implementation of PASTE in hiPSCs proved to be more challenging than in HeLa cells (Section 4.3), the ability to execute RMCE in hiPSCs was evaluated.

To demonstrate the landing pad RMCE functionality, a donor plasmid containing a CAR-BFP cassette flanked by matching attP/B sites was co-transfected with the eeBxb1 (Bxb1 serine recombinase) expression vector that was previously validated in HeLa cells in Section 4.3. RMCE edited cell lines were isolated following treatment with 20 μM GCV for 24 hours to eliminate cells retaining the landing pad following RMCE (**Figure 4.9a**). Live-cell imaging over 132 hours showed the GCV treatment efficacy as a dramatic reduction in cell confluence (from 85% to <25%), and green integrated intensity (GII) was observed between 48 and 96 hours post-treatment. Following GCV treatment, the surviving population was harvested for FACS analysis, which revealed populations of cells that had undergone RMCE, as indicated by the loss of GFP and acquisition of BFP (**Figure 4.9b**). The BFP+/GFP- population was gated and sorted for subsequent culture and expansion. Subsequently, flow cytometric analysis was performed on expanded cells to confirm the purity and stable expression of the CAR-BFP cassette (**Figure 4.9c**). These results demonstrated that the PASTE-integrated cassettes are accessible and functional for downstream genomic manipulations.

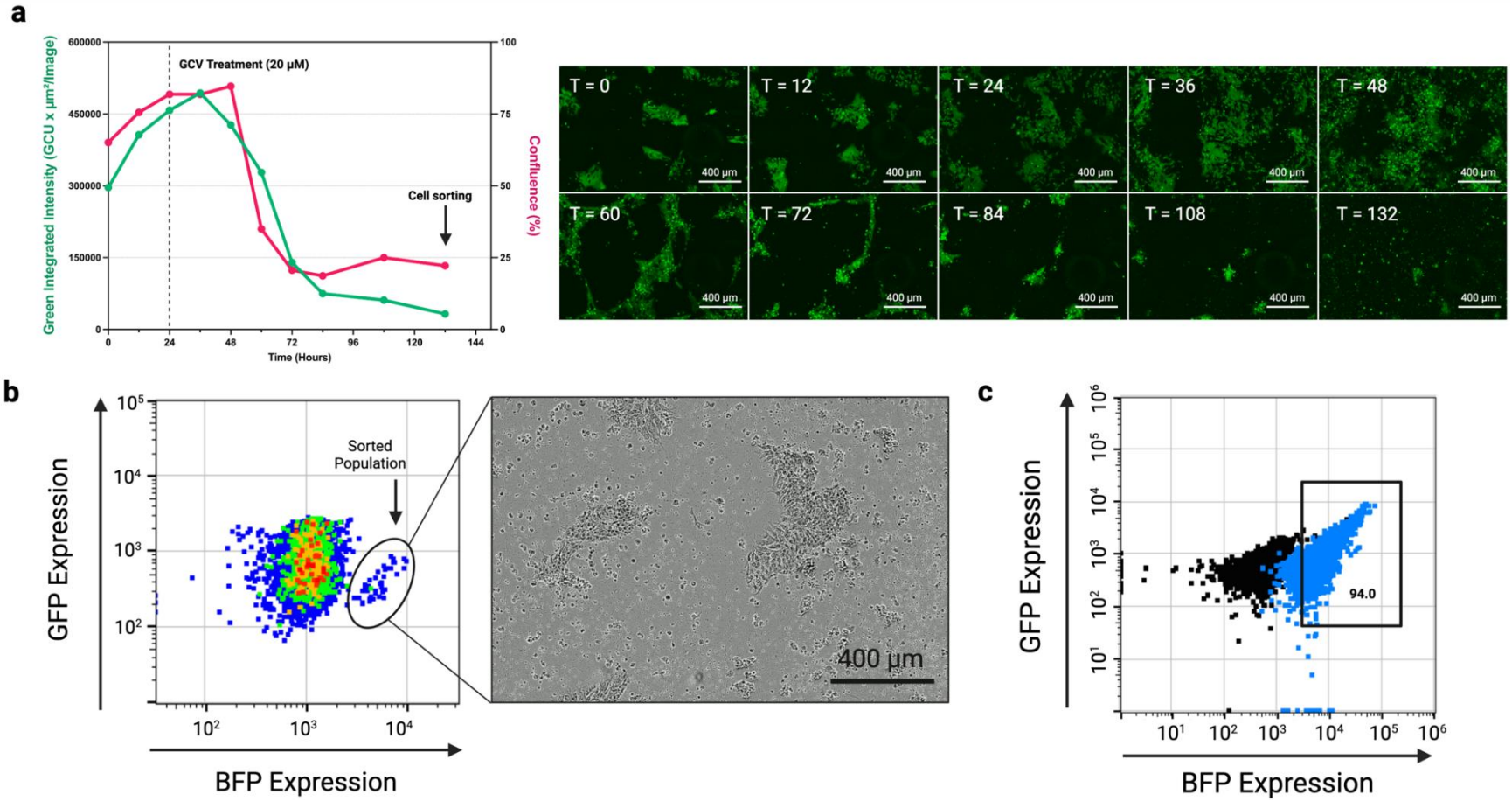

a
Green Integrated Intensity (GCU x µm²/Image)
Confluence (%)
Time (Hours)
GCV Treatment (20 µM)
Cell sorting
T = 0
T = 12
T = 24
T = 36
T = 48
T = 60
T = 72
T = 84
T = 108
T = 132
400 µm
b
GFP Expression
BFP Expression
Sorted Population
400 µm
c
GFP Expression
BFP Expression
94.0

Figure 4.9: Recombinase-mediated Cassette Exchange in CBX3.1 PASTE-edited hAF09 hiPSC (*AAVS1b*).

(**a**) Live-cell monitoring GCV treatment for negative selection. The graph (left) shows the decline in total cell confluence (pink, right axis) and green integrated intensity (green, left y-axis) over a 132-hour window. Treatment with 20 μM GCV induced apoptosis in PASTE-edited cells retaining the landing pad cassette. Corresponding time-lapse fluorescence images (right) visually confirm the clearance of GFP expressing cells over time.

(**b**) FACS analysis and sorting for the isolation and expansion of RMCE-edited cell lines. Flow cytometry scatter plot (left) shows a population of BFP+/GFP- cells that were sorted for expansion. A representative brightfield image (right) shows morphology of isolated cells after sorting and expansion.

(**c**) Flow cytometric analysis confirming stable BFP expression following expansion of cells.

Chapter 5: Discussion

5.0 Chapter Overview

This section comprises two main sections reflecting the research conducted in Chapters 3 and 4. First, we provide a comprehensive overview of the feasibility and promise of utilizing synthetic biology tools such as CRISPR technology to reprogram hiPSCs into brain endothelial-like cells. Next, we showcase a design, build, and test pipeline for integrating large synthetic transgenes into genomic safe harbour sites to assess gene silencing. Collectively, these findings highlight the potential of leveraging synthetic biology toolkits for forward and directed cellular reprogramming strategies, demonstrating how certain deficits in stem cell-based *in vitro* models (e.g., BBB) can be overcome by using this approach.

5.1 Accelerated Reprogramming of hiPSCs Into Functional Brain Endothelial-like Cells

The integrity of the BBB is essential for protecting the brain while enabling the selective transport of nutrients and molecules vital for CNS function. As such, the development of robust and predictive *in vitro* BBB models is critical for advancing CNS drug screening platforms and investigating vascular contributions to neurological disorders. Here, we have described a novel transcriptional reprogramming strategy that leverages CRISPR/dCas9 to generate hiPSC-derived brain endothelial-like cells. By inducing the gene expression of key BEC markers (*CLDN5, PECAM1, CDH5, KDR*) or endothelial transcription factors (*ETV2, ERG, FLI1*) we were able to establish a directed differentiation strategy. The targets were selected partly because these endothelial-

specific genes are known to be underexpressed or deficient in established BEC differentiation protocols.

5.1.1 Transcriptional Reprogramming Accelerates Differentiation and Induces Functional iBEC Phenotypes.

Leveraging CRISPRa technology for a targeted directed differentiation strategy, we were able to accelerate differentiation of hiPSCs into BECs that exhibit barrier properties and selective permeability, characteristic BEC morphology, and expression of requisite BBB tight junction and transporters/receptors. The CRISPRa hiPSCs appeared to be primed and responsive to differentiation cues, which correlate positively with higher gene activation. In contrast, hiPSCs expressing a mock CRISPRa vector exhibited culture behaviour indistinguishable from the parental cells, confirming that the differentiation-induced changes were being driven by targeted gene activation. Thus, our working hypothesis is that the CRISPRa hiPSC lines exist in a partially pre-differentiated state that enables efficient BEC differentiation upon exposure to BEC differentiation cues. This pre-induction of brain endothelial-associated transcriptional signals appeared to enhance the ability of the CRISPR-integrated hiPSCs to survive endothelial culture conditions during lineage commitment, providing a “priming” effect that facilitated the development of endothelial identity. We observed that a simplified five-day differentiation protocol successfully produced terminally differentiated and functional brain endothelial-like cells. The differentiated CRISPRa iBECs exhibited expression of key BBB tight junction, adherens junction and adhesion proteins, such as CLDN5, CDH5, PECAM1, ZO1, and Occludin, as well as BBB-enriched receptors and transporters such as TfR1, IGF1R, TMEM30A and GLUT1. Compared to conventional iBEC differentiation methods, our CRISPRa-based strategy generated cells with comparable BEC characteristics, including

the formation of a robust barrier, achieving TEER values up to 530 Ω·cm² and low NaFl paracellular diffusion rates between 0.01 and 0.12 x 10^{-3} cm/min. We found that 4EG-iBECs exhibited higher TEER values (up to 553 ± 39 Ω·cm²) than 3TF-iBECs (377 ± 49 Ω·cm²). We also expected that activating KDR would help iBECs develop the ability to undergo angiogenesis. Indeed, when treated with VEGF, both 3TF-iBECs and 4EG-iBECs showed cell sprouting and branching, confirming their angiogenic potential. Thus, CRISPRa-regulation of gene expression appears to be capable of inducing innate processes, such as angiogenesis, that are currently deficient in many iBEC models. Collectively, these findings highlight the potential in utilizing synthetic biology tools in enhancing BBB models.

5.1.2 Direct Activation of Structural Genes Yields Enhanced Marker Expression.

In line with previous studies that demonstrated transgene expression of the ETS transcription factors drives BEC differentiation and improves phenotype[266,461,462], we anticipated that the induction of BBB-specific transcription factors (3TF) could result in directed differentiation. However, the induction of BBB-specific tight junction genes was not expected to drive differentiation in a similar manner (**Figure 3.3**) suggesting that the 4EG genes may play an unrecognized role in supporting the BEC differentiation process. Thus, further investigation is warranted to explore the mechanisms by which the induced tight junction and BBB genes influence BEC differentiation, particularly to the potential involvement in the KDR and the Wnt signaling pathway[463,464]. Understanding these mechanisms could provide deeper insights to modelling, differentiation and enhance the robustness of future *in vitro* BBB models.

Subsequently, we characterized differentiated cells for BEC markers and function to assess their suitability as a BBB model. When compared to 3TF-iBECs, the expression of tight junction proteins (CLDN5, PECAM1) and an adherens junction protein (CDH5) was observed to be higher when directly targeted by CRISPRa in the 4EG-iBECs (**Figure 3.2c**). This suggests that directly targeting genes with CRISPRa yields greater induction than indirect targeting through transcription factors for downstream expression[465–467]. Moreover, it is not clear if the level of transcription factor activation in the 3TF-iBECs would be sufficient to activate all BEC-relevant downstream genes. Thus, consideration should be given to the activation strategy (direct or indirect) utilized along with sgRNA efficiency to select an appropriate outcome for individual genes in future multiplexed CRISPR reprogramming systems.

5.1.3 hiPSC-derived BECs Exhibit a Hybrid Heterogeneous Cell State and Late-stage Epigenetic Silencing.

A persistent challenge found in iPSC-derived BBB models is lineage heterogeneity, where gene expression in iBECs reveals both endothelial and epithelial characteristics and fails to fully leave the premature state. Despite the accelerated and optimized differentiation strategy leveraging CRISPR/Cas9 technology for iBEC programming, these models have yet to fully replicate the authentic BEC phenotype and lineage. This is supported by the transcriptomic analysis observed in the PCA plot in **Figure 3.7d** where CRISPRa-differentiated iBECs still cluster with other iBEC models. Moreover, the global expression heatmap (**Figure 3.7b)** demonstrates that CRISPRa iBECs also show partial overlapping of transcriptome patterns in the control iBEC populations despite it lacking some epithelial (Caco2) expression subsets (e.g., *KRT7* downregulation). This suggests that CRISPRa transformed iBEC expression profiles remain consistent with the standard protocol, which

is consistent with the PCA analysis in **Figure 3.7d** and the Pearson correlation coefficient similarity matrix (**Figure S7**). One possible explanation is that, despite improvements in marker expression, CRISPR-differentiated iBECs share similar limitations inherent to all *in vitro* hiPSC differentiation systems. These may include incomplete maturation, absence of key *in vivo* microenvironmental cues (such as interactions with pericytes, astrocytes, and neural cells), and limited exposure to physiological mechanical forces like shear stress. As a result, the transcriptomic profiles remain highly similar across iBEC models but still diverge from bona fide human BECs, which develop within a complex, dynamic vascular niche *in vivo*. Similarly, heatmaps based on canonical brain endothelial and choroid plexus genes[468] (**Figure S9**) further emphasize the distinct hybrid state of iBECs relative to gene signatures derived from endogenous cells. Moreover, it appears that differentiation toward iBECs does not produce a discernible trajectory toward a relevant epithelial cell lineage (choroid plexus). This indicates that, beyond targeted gene activation, incorporating additional biological and environmental cues, such as interactions with glia and neurons, extracellular matrix components, and physiological mechanical forces, may be essential to fully replicate the complex transcriptional programs and functional characteristics of bona fide BECs. Thus, while the directed reprogramming strategy described in this study achieved a BEC phenotype, there is a clear need for further improvements. To highlight this, further analysis of our RNA-seq datasets was performed on raw counts (normalized to transcripts per million, TPM) and key markers were extracted for comparison (**Figure S10**). At the differentiated iBEC state, both 3TF-SAM and 4EG-SAM populations revealed low to negligible expression of most induced targets, with the exception of *CLDN5* in 4EG-SAM and *ERG* in both 3TF-SAM

and 4EG-SAM iBECs. Notably, 4EG-SAM hiPSCs demonstrated upregulation of all four target transcripts at the pluripotent stage, indicating that the CRISPRa constructs were functional prior to differentiation. The BBB-associated transporters and surface markers (GLUT1, IGR1R, TfR1, ZO1 and TMEM30A) were readily detected in both CRISPRa iBEC populations, confirming that the differentiation protocol induced major BBB identity features. Notably, the epithelial marker *EPCAM* was robustly expressed at both the iPSC stage and in iBECs (**Figure S10**), indicating that iPSCs appear to exhibit a pre-existing epithelial characteristic that is retained following differentiation. This phenomenon is also observed in the heatmap of choroid plexus genes (**Figure S9B**). Despite the persistent dCas9/SAM expression at both the hiPSC and differentiated state, we observed a failure to sustain activation of endogenous targets at the terminal iBEC stage, indicating that transcriptional reprogramming barriers emerge as a result of differentiation. We propose that chromatin compaction and associated remodelling complexes may restrict accessibility of key loci in differentiating cells, limiting CRISPRa-driven activation during late-stage reprogramming and maturation. Therefore, future work should investigate epigenetic co-activation strategies, including chromatin-modifying CRISPR effectors such as CRISPRon/CRISPRof[354], or small-molecule epigenetic modulators, to overcome potential silencing mechanisms and improve the robustness of endogenous gene activation platforms for terminal iBEC differentiation.

5.1.4 Transcriptomic Upregulation of Key Endothelial Identity and Functional Genes.

Interestingly, when comparing the top differentially expressed genes between CRISPRa iBECs and control iBECs, we identified several key genes that were upregulated and relevant to BBB fidelity and commitment. For example, a recent study[446] demonstrated

that *RGS5* is a marker of BBB integrity and neuroprotection in focal cerebral ischemia. The loss of *RGS5* leads to endothelial dysfunction, tight junction protein relocalization and increased BBB permeability. Furthermore, *GJA1* (Connexin 43) was highly upregulated compared to control iBECs. Recent works by Krolak et al.[444] demonstrated that endothelial gap junctions are required for neurovascular coupling that coordinates blood flow with neuronal activity. Restoring *GJA1,* comparable to hBMECs, demonstrated that cells were hemodynamically competent and elucidated its role in the distribution of resources. Beyond its role in electrical signalling, another study[445] revealed a structural role in which it interacts with the tight junction protein ZO-1 to stabilize the barrier complex. Disruption of *GJA1* revealed dysfunction in CLDN5 and ZO-1 formation, resulting in hyperpermeability. Therefore, the robust expression in this marker reveals that CRISPR iBECs reveal sound electrical and structural factors for the BBB phenotype. In summary, 3TF and 4EG programming strategies do not solely address barrier tightness but may also have implicit roles in gene networks regulating brain endothelial-specific identity and function by silencing aberrant epithelial markers (e.g., *KRT7*) and enhancing machinery for neurovascular coupling and junction formation (e.g., *GJA1 and RGS5*).

5.1.5 Future Adaptations and Direction: Multiplexed Modulation of the Neurovascular Unit.

The persistence in cell lineage heterogeneity and hybrid cell states observed in our transcriptomic data suggests that future iterations of this platform should move away from bulk population analysis. Integrating pooled CRISPR screens with synthetic biology tools such as cellular barcoding[469] and lineage-resolved single-cell profiling[470] will be advantageous. These approaches would enable a systematic approach to identify cell brain endothelial molecular determinants or off-lineage neuroepithelial cell states where

we could map various clonal trajectories from hiPSCs to differentiated iBECs. Furthermore, downstream modifications to this multiplexed gene regulation system could implement CRISPR interference (CRISPRi) to target epithelial genes and other genes not expressed in endogenous BECs. Additionally, recent studies identifying transcription factors and genes not previously associated with BEC differentiation and function should be incorporated into future CRISPRa/i system designs[465]. Moreover, additional phenotypic deficiencies observed in transcriptomic analysis (**Figure 3.7C**), such as immune cell interactions and transient barrier properties, could be enabled by the expression of key functional genes and utilization of more representative multicellular environments with flow characteristics. Experiments exploring the effect of shear stress and more complex multicellular microphysiological systems on iBEC phenotype will continue to improve BEC differentiation protocols[471]. Ultimately, we expect that phenotype refinement of *in vitro* models will involve addressing the various limitations imposed by static 2D monoculture systems, with multiplexed gene regulation playing an important role in fully realizing the vast potential of iPSC-derived cell and tissue models.

5.1.6 Technical Limitations: Vector Integration Variability and Cell Line Generalizability.

A key limitation of this study is the variability and potential genotoxicity associated with piggyBac transposase-mediated genomic integration of CRISPRa constructs into hiPSCs[472]. The genomic integration site for PiggyBac transposase is the TTAA motif, which is widespread throughout the human genome, with reports indicating hundreds of genomic integration sites[473]. Consequently, PiggyBac integration does not permit precise control of the integration site or copy number, leading to heterogeneous transgene expression, variable gene induction and increased cell-cell variability. To address these

limitations, future iterations of this platform should incorporate site-specific integration strategies into characterized genomic safe harbours (e.g., *AAVS1, CLYBL, CCR5, hRosa26, hRogi1, hRogi2*) to achieve uniform expression, controlled copy number and the generation of clonally defined hiPSC lines. In parallel, non-integrating approaches such as mRNA-based CRISPR systems or episomal vectors should be explored to reduce genotoxic risk while maintaining tunable transgene expression. The development of inducible CRISPRa platforms would further provide temporal control over transcriptional activation, enhancing the flexibility and scalability of directed differentiation workflows. Lastly, differentiation consistency and phenotypic characterizations in this study were limited to hiPSCs derived from human amniotic fluid. Expanding this approach to other donor cell types, such as fibroblast and bone marrow-derived cells, will be critical for assessing robustness and donor-dependent variability in our CRISPR-mediated brain endothelial cell differentiation scheme.

5.2 Establishing a Site-specific and Modular Genomic Engineering Platform in hiPSCs

A significant challenge evident in Section 5.1 was the reliance on semi-random gene integration toolkits, such as PiggyBac and lentiviral vectors, due to their limited control over copy number and heterogeneity in clonal hiPSC development. As such, we have designed, built and tested a pipeline for genomically engineering hiPSCs through integrating transgenes into genomic safe harbour sites and characterizing the expression profiles and their susceptibility to gene silencing during germ layer differentiation. This work elaborates on the dependencies of promoter-gene locus interplay for transgene expression and establishes a versatile, modular pipeline for downstream engineering.

The final product is hiPSCs that are capable of a plug-and-play transgene insertion, leveraging recombinase-mediated cassette exchange (RMCE), which later provides for facile hiPSC engineering and the development of clonal engineered cell lines.

5.2.1 Design and Validation of Modular PASTE Integration Strategy

During the design and testing phase of this project, extensive optimizations were required before deployment into hiPSCs. The design process involved multiple iterations of the reporter construct (parental minicircle vectors), as expression of version 1 of the reporter plasmid yielded an extremely weak signal. For example, version 2 of the plasmid used an extended Kozak consensus sequence and an optimized 5' UTR sequence in an attempt to prevent alternative upstream initiation and to maximize translation efficiency at the correct start codon reading frame. In version 3, the novel StayGold GFP reporter was replaced with a well-characterized copGFP reporter, which had previously been used in the PASTE studies[331] to ensure more reliable readouts. Ultimately, integrating the Woodchuck Hepatitis Virus Post-Transcriptional Regulatory Element (WPRE) downstream of the poly (A) tail resulted in the pMini vectors that exhibited stable and enhanced fluorescent protein expression (later denoted as version 4). All subsequent pMini vectors with this modification were subcloned with unique promoters and verified for downstream genomic integration.

Similar to the prior challenge, a secondary challenge emerged when electroporating the minicircle DNA into hiPSCs. In hiPSCs, the transfections resulted in silenced or unviable cells that contained no GFP signals (**Figure 4.4a**). However, it was clear that the same minicircle preparations showed strong and functional expression in HeLa cells (**Figure 4.3**). First, we ruled out DNA concentration optimization as the limiting factor (**Figure 4.4a**).

We hypothesized that the procedure going from pMini to minicircle generation compromised the hiPSC viability. This was later validated by electroporating the parental pMini plasmids into hiPSCs, demonstrating that the lack of expression and hiPSC toxicity were caused by the minicircle preparation step. Upon investigation, minicircle preparations utilize large inoculates of bacterial culture, which, by nature, results in the accumulation and copurification of endotoxins (lipopolysaccharides, LPS)[474]. LPS have been shown to bind Toll-like Receptor 4 (TLR4), triggering a downstream signalling pathway that activates inflammatory transcription factors such as NF-kB[475,476]. The discrepancy between the cell lines is explained as a result of their innate immune profiles, where HeLa cells express non-functional TLR4 receptors due to the lack of MD2 co-receptors[477,478]. On the other hand, hiPSCs are primary cells that maintain tightly regulated innate immune receptors (including TLR4/MD2), which actively induce inflammatory stress and apoptosis in the presence of LPS[479,480]. Therefore, an endotoxin-free minicircle extraction protocol was implemented for all minicircles, effectively eliminating LPS contamination and consequent immune response, thereby restoring cell viability and transgene expression in electroporated hiPSCs (**Figure 4.4d**).

Subsequently, to validate the successful integration of the cassette, genotypic validation was performed to ensure that the PASTE machinery was integrated at the targeted loci rather than through random integration. A PCR strategy was used on genomic DNA (gDNA) to investigate the integration junctions and the insert, respectively. Primer Set A validated the 5' GSH site and the 5' cargo band to be amplified, while Primer Set C verified the internal insert cassette only. Interestingly, Primer Set B was designed to flank the entire integration region extending from the 5' end to the 3' end of the safe harbour. In

wildtype (WT) conditions, this amplification should yield a smaller band, ranging from ~400 to ~1000 bp. However, in all edited cell lines, these bands should amplify a much larger ~6-7kb band. Due to the restrictions in thermocycling conditions for gDNA amplifications, this extension would often lead to many off-target amplifications. Thus, the PCR reactions were deliberately shortened to account for this. Therefore, any PASTE-edited cell lines should yield no amplifications in a perfect clone. However, as expected, the PASTE-edited populations revealed amplifications of Primer Set B in pooled populations, indicating heterogeneity and or the presence of unedited alleles within the culture. To mitigate this, analysis of a clonal hiPSC line (CAG-AAVS1b) showed near-complete loss of all Set B signals, indicating effective bi-allelic disruption of the safe harbour.

5.2.2 Gene Expression in PASTE-edited hiPSCs

We first characterized and evaluated an extended panel of promoters (CBX3.1, CAG, EFS, PGK, CMV, and EF1α) (**Figure 4.5a**). The preliminary expression profile resulted in a broad and titratable spectrum of fluorescence intensity shifts, establishing a clear hierarchy of promoter strengths in hiPSCs (CMV > CBX3.1 > CAG > EFS > PGK > EF1α). These results emphasize the importance of careful consideration of design requirements when selecting and building synthetic cargoes for hiPSCs.

In order to evaluate promoter-locus dependencies on gene expression, we benchmarked top-performing promoters such as CAG and CBX3.1 with the weakest endogenous promoter (EFS) at all four GSH sites. The flow cytometric analysis revealed that the viral CAG-driven GFP expression was retained in ~30-40% of the cells, with the CCR5 locus dropping to a 13.7% GFP+ population. Comparatively, the CBX3.1 promoter drove robust

and consistent transgene expression (sorting efficiency ranging from 38.8% and 44.1%) across all safe harbour sites. The direct comparison between CBX3.1 and EFS in this project also illustrates the structural mechanisms involved for gene stability. The CBX3.1 promoter is a synthetic in-house promoter generated at the NRC by integrating a Chromobox 3 (CBX3) ubiquitous chromatin opening element (UCOE) upstream of the core EFS promoter. It is apparent that the EFS promoter was subjected to nearly complete silencing across all loci in hiPSCs (under 3% GFP-positive), whereas the CBX3.1 construct remained active. This is consistent with other iPSC-based research demonstrating that UCOEs are barriers to promoter activity due to DNA methylation and heterochromatin spread. The CBX3 UCOE effectively insulates the core EFS promoter from positional silencing at the pluripotent state and assures that transgenes are expressed reliably regardless of the four safe harbour sites chosen.

Another consideration in interpreting the flow cytometric analysis is distinguishing between promoter-locus interactions in the context of baseline PASTE integration efficiency. The differences in GFP+ populations are represented as the metric of successful PASTE integration and expression strength. However, comparisons across loci also indicate the promoter's effect on various GSHs. For example, the drastic reduction in the GFP+ population in the CAG-CCR5 combination cannot be attributed solely to poor PASTE integration, as the CBX3.1 promoter achieved 44.1% at the same locus using the same atgRNA and ngRNA targeting machinery. This internal control suggests that there is an inter-dependence between the promoter and GSH site and both factors are involved in the net outcome.

Consequently, the hierarchical expression profile established in this study provides an informative decision-making blueprint for modular cellular engineering. Not all synthetic biology applications would require maximal transcriptional output. Therefore, it is recommended for researchers to make purposeful and selective design strategies when titrating protein expression for developing cell models and therapeutic cell lines. While high levels of CMV-driven expression might be advantageous for transient expression, it is often unsuitable for directed differentiation, where extremely high expression can induce toxicity, apoptosis or aberrant off-target lineages. For example, in cellular immunotherapies, high expression of chimeric antigen receptors (CARs) was linked to T cell exhaustion and apoptosis[481,482]. For these applications, UCOE-stabilized promoters such as the CBX3.1 promoter produce consistent, high expression levels, whereas other promoters, such as EFS, offer a tighter, lower level of expression. Therefore, the characterized hierarchy of promoter strengths established in this research can serve as a preliminary blueprint to aid researchers to rationally selecting promoters with predictive expression strengths. Synthetic biologists can as a result prevent overexpression-associated toxicity while also achieving threshold to achieve a therapeutic or functional output.

5.2.3 Transgene Silencing in Lineage-specific Differentiations

A major consideration for cellular reprogramming strategies is to ensure that genomic engineering platforms enable expression during and throughout the terminal stage of differentiation. Pluripotent stem cells were previously characterized as transcriptionally permissive and highly dynamic, as they retain an open chromatin architecture, reflecting their uncommitted cell fate[483–485]. However, as these cells transition towards a somatic

cell fate, global epigenetic remodelling occurs to reinforce lineage specification[486,487]. Differentiating cells require epigenetic checkpoints to prevent dedifferentiation, which are accomplished by condensing euchromatin into heterochromatin with DNA methyltransferases (DNMT3A/3B) in pluripotent networks and off-target lineage pathways[488–490]. As a result, regenerative medicine researchers often struggle to obtain persistent expression of synthetic genes with high yield throughout differentiation [404,491]. Notably, epigenetic silencing is not uniform across all cell developmental trajectories[492,493]. For example, pluripotent cells are naturally biased towards the neuroectodermal cell fate, thus requiring a stringent epigenetic silencing process to reinforce mesodermal and endodermal lineages[492,494–496]. Therefore, we monitored transgene expression in PASTE-edited hiPSCs under diverse epigenetic conditions by monitoring GFP expression over time and at the terminal stage of differentiation in all three embryonic germ layers (**Figure 4.7**). When plotting relative expression strengths as a heat map (**Figure 5.1**), lineage bias for gene silencing patterns can be observed. For example, with strong acting promoters (e.g., CAG and CBX3.1), ectoderm-derived cells seem to resist silencing (3.5-5). Conversely, endoderm-derived cells were more susceptible to silencing as shown by reduced expression with strong promoters, and near complete silencing under the EFS promoter. This further reinforces the idea that there is no universal one-size-fits-all strategy for safe genome engineering when considering factors such as loci, promoter, and lineage interplay.

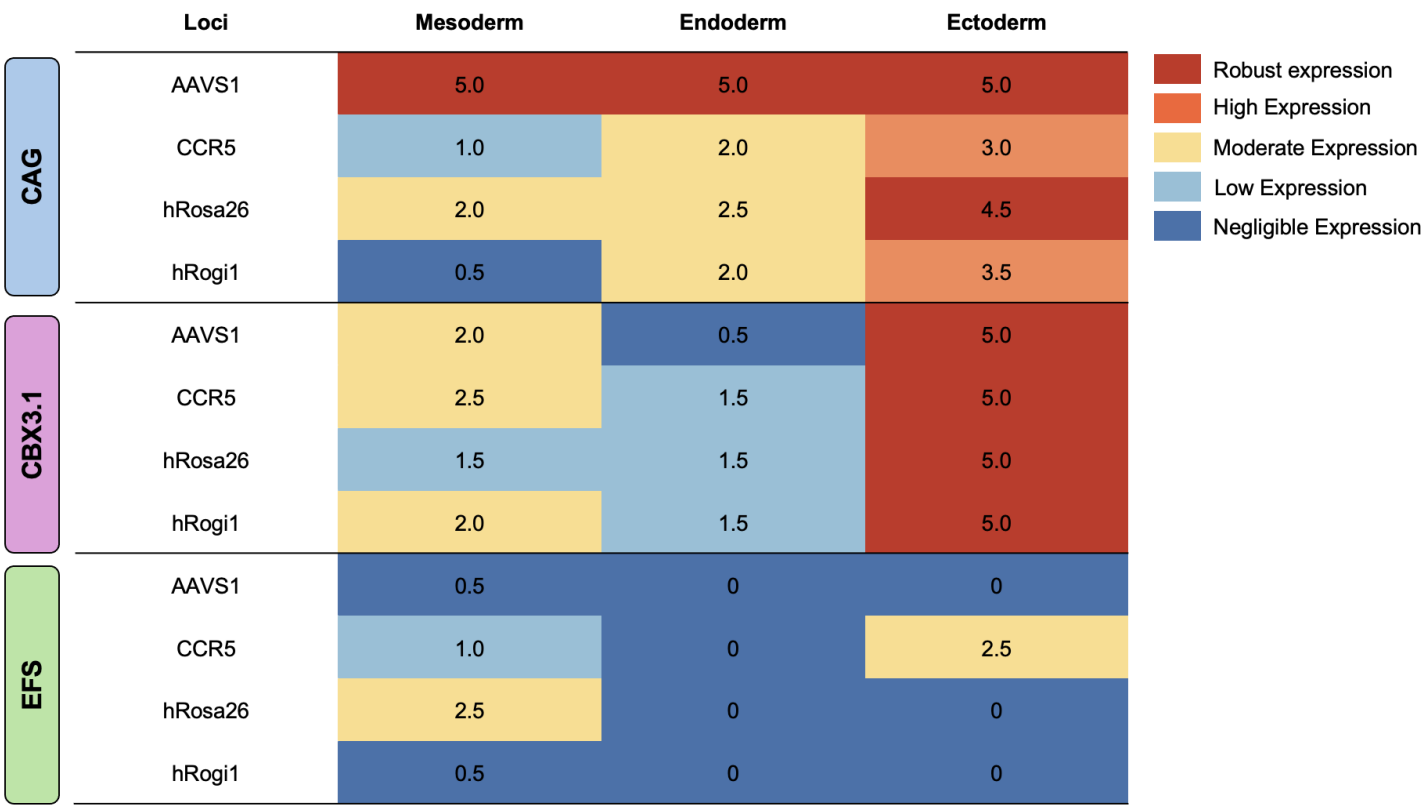


Figure 5.1 Heatmap of Transgene Expression Across Germ Layer Lineages. This heatmap matrix visualizes the stability of various promoter-locus-lineage configurations for stable GFP expression at terminal differentiation. To facilitate a macro-level summary, the raw flow cytometric data on the percentage of GFP+ cells detailed in Figure 4.7 were transformed into an expression score (0-5). This scoring criteria are as follows: 0 = Not Significant (NS) compared to control; 1 = <10% GFP+; 2 = 10-24% GFP+; 3 = 25-49% GFP+; 4 = 50-74% GFP+; 5 = ≥ 75% GFP+. The scoring was performed by calculating the mean expression score across replicate values (n=2).

5.2.4 The Oversimplification of GSH Safety and Stability Criteria

As validation of the trends and variability observed in the dataset, similar trends were found in reports that have targeted the GSH sites. As summarized in **Table 5.1,** the expression variability and silencing phenotypes observed in our platform align strongly with independent reports across diverse differentiation and cell lineages. Early studies established that *AAVS1* and *hRosa26* are permissive sites for transgene studies. However, newer reports are revealing that they cannot all universally resist silencing, especially during differentiation. Bhagwan et al.[412] demonstrated this clearly, revealing that CAG-driven constructs resulted in a some decline in transgene expression as early as the mesoderm stage, with expression completely silenced at the terminal cardiomyocyte stage. Similarly, Karbassi et al.[497] reported that CRISPRa constructs also experienced silencing when differentiating into cardiomyocyte and endothelial populations despite testing three loci. The authors importantly linked that the repression is likely caused by promoter methylation at these loci during lineage commitment, potentially overriding stability even in the described GSHs. Collectively, these studies and our dataset reinforce that GSH sites alone are not the sole factor in sustaining transgene expression in bioengineering applications. This demonstrates that genomic safe harbours are not truly “safe” or universally insulated environments, and that the assumption that integration into a designated GSH guarantees stable, indefinite transgene expression may be an oversimplification. Therefore, this criterion fails to account for the complexity and interplay among various factors and requires researchers to consider other factors, such as promoter designs, epigenetic state, differentiation trajectory, and transgene architecture, to optimize and mitigate variability and gene silencing.

Table 5.1: Reported Expression Variabilities Across Different Safe Harbours, Promoters and Targeted Cell Types.

Safe Harbours	Promoters	Targeted Cell Lineage	Observations	References
AAVS1, hRosa26, CLYBL	CAG, Tissue-specific promoters	hiPSCs, cardiomyocytes, endothelial cells	dCas9-VPR CRISPRa constructs underwent differentiation-associated silencing across various safe harbours. Promoter DNA methylation was observed	Karbassi et al. 2024
AAVS1	CAG	hiPSCs, mesoderm, cardiomyocytes	Clonal variability and transgene silencing during the differentiation process from 98% to 1.3% upon terminal cardiomyocytes	Bhagwan et al. 2020[412]
AAVS1 hRosa26	CAG, EF1α, Tet-On	hiPSCs, neurons, skeletal myocytes, oligodendrocytes	Systematically optimized a dual-locus system to perform directed differentiation by placing one cassette in hRosa26 and another cassette into AAVS1 for forward reprogramming of hiPSCs	Pawlowski et al. 2017[498]
AAVS1 CCR5	SFFV, EF1α, PGK	T-lymphocytes, lymphoblastoid cells, neural stem cells, hiPSCs, hematopoietic progenitors	Demonstrated that tailoring locus and cassette designs enabled certain configurations of optimal transgene expression	Lombardo et al. 2011[499]
AAVS1	CAG, Tet-On	hiPSC, definitive endoderm, pancreatic progenitor	Demonstrated using a Tet-On system at the AAVS1 locus yield sustained expression at each checkpoint of the differentiation step	González et al. 2014[500]
hRosa26	Used endogenous promoters in the locus	Embryonic stem cell, embryoid body, germ layer	Functionally validated hRosa26 as a potential site for transgene expression when differentiated into all three germ layers	Irion et al. 2007

5.2.5 Evaluating UCOE Insulators and Promoters Across Germ Layers

The EFS and CBX3.1 promoters are structurally similar, primarily differing in the presence of a CBX3-derived UCOE sequence. Comparing their performance throughout differentiation provides insight into the function of synthetic insulators in iPSC differentiation. Under all germ layer differentiations, the EFS promoter was silenced across the GSH sites, dropping to ~4% in expression. In contrast, the CBX3.1 promoter produced stable GFP expression of > 90%, suggesting that the UCOE sequence may prevent gene silencing. However, under the mesoderm and endoderm conditions, the stability of expression plummets to ~25% and 15%, respectively. Additionally, differences in the timing of peak expression were observed, days 3-4 (mesoderm, endoderm) and day 5 (ectoderm). This suggests that the timing in which cell lines commit to lineage specification is variable.

Moreover, these results highlighted that the CBX3.1 promoter configuration containing a UCOE element is not perfect. Although overall expression was increased in mesoderm and endoderm lineages, CBX3.1 alone cannot fully prevent silencing, as efficiency remains below a desirable range. These findings suggest that a single factor, such as a UCOE acting as a universal safeguard against silencing, is unrealistic and an oversimplification. Especially in the context of cellular reprogramming, where differentiation and chromatin remodelling are involved, solely relying on strategies such as this will inevitably result in insufficient and or variable transgene expression. Future synthetic biology efforts should focus on discovering novel insulators[501,502] and screening combinations of UCOE sequence elements[503,504] with various promoter combinations at GSH sites to characterize the panel in depth across all germ layers.

Alternatively, the CAG promoter demonstrated a unique pattern of lineage-specific resistance to silencing. For example, the CAG-AAVS1b clone exhibited highly persistent GFP expression across all mesoderm (~85%), endoderm (~90%) and ectoderm (~98%) germ layers. The *AAVS1b* locus is located within the *PPP1R12C* gene and is an actively transcribed housekeeping region with a well-defined open chromatin structure [406,505,506]. It is apparent that the combination of a strong viral promoter (CAG) and an open chromatin region synergize to enable robust transgene expression in both mesodermal and endodermal lineages.

5.2.6 Cargo Constraints and Transgene Structure

Although this study focused on promoter strength, integration site and lineage, it is also important to consider the effect of the cargo itself. In all the assays, the expression efficiency and strength are standardized to GFP. However, especially across various cell types, epigenetic silencing is not limited to the transcriptional regulatory elements (e.g., promoters, TSS, terminators, machinery) and can be affected by the sequence of the transgene as well. For example, sequences within the transgene, such as CpG dinucleotide[507], dormant splice sites[499] , and bacterial/viral elements,[508] can be recognized as foreign factors by the cell and degraded. Additionally, sequences specific to therapeutic transgenes, such as chimeric antigen receptor (CAR) constructs or multi-gene transcription factor arrays[509], can also increase the complexity in stem cell engineering, as they would require careful dosage titration. Our study provides a preliminary blueprint for the cellular reprogramming toolbox that identifies optimal lineage-promoter-loci configurations. However, future studies should also independently validate their own transgenes to ensure stability and efficacy catered to their needs.

5.2.7 Transgene Expression at Epigenetic Checkpoints to Predict Silencing

Pluripotent cells undergo global shifts in their epigenetic profile that can be used to predict long-term transgene behaviour[492]. Previous research has shown that early lineage commitment involves widespread histone modification of heterochromatin[510], which is consolidated by DNMT-mediated methylation[486]. Some of these marks (e.g., DNMT1) behave as molecular checkpoints that are passed mitotically throughout their cellular lineage[490]. Hence, the transgene silencing observed within the first week of germ layer differentiation can serve as a suitable indicator of gene silencing throughout the maturation of somatic cell lines. While we can use these predictions to assess early silencing patterns, they do not always guarantee long-term epigenetic stability in all cell types. For example, germ layer cells are proliferative progenitor lines and are not fully mature cell types. In conventional differentiation pipelines, progenitor cells undergo longer stages of refined differentiation lasting weeks to months (e.g., NK cells, cardiomyocytes, endothelial cells, neurons). Therefore, it is likely that the clones containing promoter-locus combinations with partial silencing or weak expression could later undergo stronger silencing as a result of the cellular maturation process. Ultimately, this provides a predictive baseline for reprogramming studies, enabling researchers to pre-select optimal sites and make informed decisions for downstream transgene integration applications. While our data can serve to focus transgene design, verifying transgene performance in fully differentiated cell types is still required.

5.2.8 Recombinase-mediated Cassette Exchange Function

Additionally, PASTE-edited cell lines allow for subsequent site-specific cassette exchange, enabling the initial transgene to be swapped. This enables downstream modifications to the edited cell lines to swap transgenes by introducing a secondary cargo flanked by

recognition sites matching the landing pads. The orientation and recognition are mediated by the unique central dinucleotide sequences (AC, GT), which were previously sequence optimized and validated *in vitro* in a panel study[331]. In this study, the RMCE process was facilitated by the eeBxb1 serine recombinase system utilizing these modified central dinucleotide attP and attB recombination sites, which were flanked in both the primary cargo of PASTE cell lines and the donor transgene (CAR-BFP). This integrase design allows for the GFP transgene to be effectively swapped without inducing a DNA double-stranded break and preserves the genomic integrity whilst also being site-specific at the specified loci.

It is important to note that the findings in this section serve as a proof-of-concept, as time constraints affected the thorough characterization of RMCE-edited cell lines and protocol. Our preliminary data demonstrates the exchange process. However, the efficiency of RMCE still requires stringent enrichment to yield a homologous population required for downstream differentiation. The HSV-TK gene was utilized to negative select populations of non-exchanged cell lines with the coupling of Ganciclovir (GCV), which showed significant reduction in confluence and green integrated intensity (GII) within 96 hours post-treatment. Additionally, FACS isolation of the surviving BFP+/GFP- populations followed by the expansion yielded a pure population stably expressing the CAR-BFP cassette.

A significant limitation in the preliminary validation is that RMCE was performed in pooled cell populations. The bulk populations demonstrated successful transgene exchange, thus, the pooled lines reveal heterogeneous integration profiles and background in the stability of the cargo. Future work should prioritize using clonally-derived lines for

characterization to establish a uniform and streamlined pipeline and dataset for the swapped cassettes.

Further optimization for this mechanism is required to optimize, streamline and validate this protocol, as these preliminary data suggest that the PASTE cassettes remain accessible for downstream genomic manipulations. This proof-of-principle pipeline, enables downstream bankable and easy-to-engineer hiPSC lines, permitting the engineering of various of genetic transgenes, ranging from complex synthetic circuits to transcription factors. The modularity of this platform not only accelerates and streamlines the process for stem cell engineering but also places more control into synthetic biologists in controlling where and how many copies should be integrated for their transgene cassettes.

5.2.9 Future Directions

The PASTE/RMCE platform formed in this study offers a versatile and effective way to characterize and advance the engineering of hiPSCs. However, to translate this strategy to therapeutic and synthetic biology platforms, future work should focus on extending the characterization and validation of this pipeline.

First, the results seen in this study can be used as a predictive indicator in germ layer differentiation when paired with various promoters and integrated at certain loci. This prediction criteria however can only be leveraged at the hiPSC and germ layer differentiated state. Future studies should next optimize the promoter-loci configurations, especially with the UCOE stabilized promoters to prolonged and specialized differentiation platforms. For example, generating mature neurons, NK cells and or 3D

organoids can last 30-60 days in culture. Observing for functional transgene expression is critical for the translation of this study to clinical and medical applications.

Second, the majority of the analysis was performed on pooled cell populations, which yield heterogeneous integration and expression profiles. To ensure consistency and reliability, the streamlined clonal isolation pipeline developed in Section 4.5 should be employed on all PASTE-edited cell lines. Future efforts should focus on performing next-generation sequencing on the clones to also validate off-target recombinase activity and genomic rearrangements to contrast with gold-standard integration tools (e.g., Cas9).

Additionally, the original goal of this study was to employ a broader panel of promoters. Our current characterization study focused on the contrast between the active CAG/CBX3.1 promoters and the silenced EFS promoter, while the preliminary data established a spectrum of expression strengths in hiPSCs (**Figure 4.5a**). Future directions of this study should continue to evaluate the entire panel of promoters at targeted genomic safe harbours throughout germ layer differentiation. Having an expanded panel of promoters is critical for being able to titrate the desired expression level of cargos (e.g., CAR-T cell activation requires lower expression to prevent genotoxicity[511,512]). By developing a comprehensive matrix assessing their promoter-locus-lineage interplay, this would provide synthetic biologists a blueprint that could allow for rational design for their desired application.

5.3 Moving Towards Precision Bioengineering Platforms in Regenerative Medicine

Although section 5.1 and 5.2 highlights distinct methodological challenges in iPSC cell platforms, they both address a grander challenge: Overcoming the unpredictability

observed in stem cell models to drive true clinical translation. The long-term goal of this research is to establish a programmable hiPSC platform that can be safely, uniformly and predictably engineered to recapitulate complex human microenvironments like the NVU. Without this level of precision, next-generation and further advancement in *in vitro* disease modelling, drug screening and regenerative cell therapies remain limited by current bottlenecks such as heterogeneous cell populations, variable transcriptome profiles and unstable phenotypes.

Forward reprogramming through transcriptional modulation (Chapter 3) successfully demonstrated the biological aspect of this solution by modulating gene networks involved in cell fate control (e.g., 3TF and 4EG arrays). Researchers can bypass developmental bottlenecks and rapidly induce differentiation of cells to functional brain endothelial-like phenotypes. However, applying these principles also exposed another bioengineering challenge which is delivering massive multiplexed CRISPRa cassettes through a semi-random integration (e.g., piggyBac) yielding heterogeneous populations and risking genotoxicity. The PASTE and RMCE platform (Seen in Chapter 4) provides the mechanical solution to this limitation. By establishing a site-specific modular landing pad with pre-described genomic safe harbours combined with an array of testable promoters we can now map and predict gene silencing and stability of transgenes by altering configurations of regulatory design, the landing site and the developmental pathway. Together these two projects establish a synergistic blueprint where genetic payload delivery such as PASTE, and transcriptional reprogramming circuits required to generate a more faithful BBB model could be done harmoniously which is a vital prerequisite for screening advanced therapeutics like lipid nanoparticles (LNP) drug delivery systems

Just a few years ago, the field heavily relied on semi-random integration methods (e.g., lentiviral and transposons) or risk averse DSB-dependent CRISPR/Cas9 HDR pathways to introduce transgenes into the genome. Today, genomic engineering is rapidly transforming by adopting precision bioengineering capabilities such as PASTE, prime editing and CRISPR-associated transposons (CASTs)[513], to maintain genomic integrity and safety. Concurrently, the expectations for *in vitro* stem cell models have also matured from the past 15 years in the BBB field. Researchers no longer simply accept heterogeneous hybrid cell states (typically observed in iBECs). There is now demand for clonally defined, and molecularly faithful cell model that can recapitulate the complex NVU more accurately, driven by understanding lineage commitment and developmental trajectories.

Looking forward, the technological developments in genome engineering further enhance the bioengineering platform established in this thesis. For example, with highly accurate AI-driven protein prediction tools (e.g., AlphaFold3)[514], miniaturized CRISPR effectors (e.g., CasMINI, hpCasMINI)[515,516], programmable CASTs[517], bridge RNAs[518] and machine learning pegRNA optimization tools (OptiPrime)[519] will allow for even more sophisticated designs of genetic payloads and delivery strategies. Furthermore, integrating novel synthetic biology tools such as single-cell multi-omics[470] and cellular barcoding[520,521] can enable researchers to precisely map out transcriptional trajectories as cells differentiate. Finally, researchers can anticipate non-viral delivery approaches to take precedence such as LNPs capable of delivering CRISPR molecules[522,523], and making them BBB-penetrable[524] for delivering neurological therapeutic treatments. Ultimately, by merging programmable genomic integration with transcriptional networks, we move closer to an

era where standardized precisely engineered hiPSC cell banks could be predictably reconfigured into any specialized cell type on demand.

5.4 Conclusion

This thesis presents two approaches to advancing hiPSC engineering: integrating a novel directed differentiation strategy and constructing a stable modular genome-engineering platform for hiPSCs. First, we showed that by targeting key genes and transcription factors with CRISPRa, transcription-based reprogramming can accelerate and rescue brain endothelial-like cells and their phenotype. Specifically, the CRISPR-differentiated iBECs generated in this study demonstrated functional tight junction formation, low paracellular permeability, angiogenic potential and receptor-mediated transport. However, the integration of these complex and large cassettes using semi-random transposon technology made directed differentiation uncontrollable and non-specific, yielding heterogeneous cells. Thus, the development of a PASTE/RMCE integration platform was done with a goal to evade gene silencing in stem cells and provide a blueprint to predict transgene silencing with various promoter-locus-lineage configurations. By mapping these interactions, we established that UCOE-stabilized promoters (e.g., CBX3.1) and specific safe harbour sites, such as AAVS1b, can prevent transgene silencing during differentiation compared to their EFS promoter counterparts. Furthermore, the RMCE platform enables a modular cargo swap function to enable downstream swapping of our transgenes

at the specified GSH sites with ease. In conclusion, this study, involving CRISPR-directed differentiation and the PASTE safe harbour integration platform, provides a comprehensive toolkit for synthetic biologists and regenerative medicine researchers and

promotes rational design and predictable engineering of hiPSCs for complex modelling and cell therapy applications.

Chapter 6: Bibliography

1. Till, J. E. & McCULLOCH, E. A. A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. *Radiat. Res.* **14**, 213–222 (1961).
2. Becker, A. J., McCULLOCH, E. A. & Till, J. E. Cytological demonstration of the clonal nature of spleen colonies derived from transplanted mouse marrow cells. *Nature* **197**, 452–454 (1963).
3. Siminovitch, L., Mcculloch, E. A. & Till, J. E. The distribution of colony-forming cells among spleen colonies. *J. Cell. Comp. Physiol.* **62**, 327–336 (1963).
4. Strauer, B. E. *et al.* Repair of infarcted myocardium by autologous intracoronary mononuclear bone marrow cell transplantation in humans. *Circulation* **106**, 1913–1918 (2002).
5. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676 (2006).
6. Campo, F. *et al.* Bioengineering of a human iPSC-derived vascularized endocrine pancreas for type 1 diabetes. *Cell Rep. Med.* **6**, 101938 (2025).
7. Li, G. *et al.* Bioprinting functional hepatocyte organoids derived from human chemically induced pluripotent stem cells to treat liver failure. *Gut* **74**, 1150–1164 (2025).
8. Fernandopulle, M. S. *et al.* Transcription factor-mediated differentiation of human iPSCs into neurons. *Curr. Protoc. Cell Biol.* **79**, e51 (2018).
9. Lancaster, M. A. *et al.* Cerebral organoids model human brain development and microcephaly. *Nature* **501**, 373–379 (2013).
10. Sharma, A. *et al.* High-throughput screening of tyrosine kinase inhibitor cardiotoxicity with human induced pluripotent stem cells. *Sci. Transl. Med.* **9**, eaaf2584 (2017).
11. Palasantzas, V. E. J. M. *et al.* iPSC-derived organ-on-a-chip models for personalized human genetics and pharmacogenomics studies. *Trends Genet.* **39**, 268–284 (2023).

12. Hayes, J. B. & Burnette, D. T. Pharmacological inhibition of β myosin II disrupts sarcomere assembly in human iPSC-derived cardiac myocytes. *Cytoskeleton (Hoboken)* **82**, 795–803 (2025).
13. Ebert, A. D. *et al.* Induced pluripotent stem cells from a spinal muscular atrophy patient. *Nature* **457**, 277–280 (2009).
14. Israel, M. A. *et al.* Probing sporadic and familial Alzheimer's disease using induced pluripotent stem cells. *Nature* **482**, 216–220 (2012).
15. Okano, H. & Morimoto, S. iPSC-based disease modeling and drug discovery in cardinal neurodegenerative disorders. *Cell Stem Cell* **29**, 189–208 (2022).
16. Nichols, J. & Smith, A. Naive and primed pluripotent states. *Cell Stem Cell* **4**, 487–492 (2009).
17. Mulas, C. *et al.* ERK signalling eliminates Nanog and maintains Oct4 to drive the formative pluripotency transition. *Development* **151**, (2024).
18. Niwa, H., Burdon, T., Chambers, I. & Smith, A. Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. *Genes Dev.* **12**, 2048–2060 (1998).
19. Mossahebi-Mohammadi, M., Quan, M., Zhang, J.-S. & Li, X. FGF signaling pathway: A key regulator of stem cell pluripotency. *Front. Cell Dev. Biol.* **8**, 79 (2020).
20. James, D., Levine, A. J., Besser, D. & Hemmati-Brivanlou, A. TGFbeta/activin/nodal signaling is necessary for the maintenance of pluripotency in human embryonic stem cells. *Development* **132**, 1273–1282 (2005).
21. Vallier, L., Alexander, M. & Pedersen, R. A. Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. *J. Cell Sci.* **118**, 4495–4509 (2005).
22. Boyer, L. A. *et al.* Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**, 947–956 (2005).
23. Loh, Y.-H. *et al.* The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat. Genet.* **38**, 431–440 (2006).
24. Young, R. A. Control of the embryonic stem cell state. *Cell* **144**, 940–954 (2011).
25. Zhang, X. *et al.* Pax6 is a human neuroectoderm cell fate determinant. *Cell Stem Cell* **7**, 90–100 (2010).
26. Lee, T. I. *et al.* Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* **125**, 301–313 (2006).
27. Suzuki, A. *et al.* Nanog binds to Smad1 and blocks bone morphogenetic protein-induced differentiation of embryonic stem cells. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 10294–10299 (2006).
28. Faial, T. *et al.* Brachyury and SMAD signalling collaboratively orchestrate distinct mesoderm and endoderm gene regulatory networks in differentiating human embryonic stem cells. *Development* **142**, 2121–2135 (2015).

29. Chen, M. *et al.* The roles of embryonic transcription factor BRACHYURY in tumorigenesis and progression. *Front. Oncol.* **10**, 961 (2020).
30. Niakan, K. K. *et al.* Sox17 promotes differentiation in mouse embryonic stem cells by directly regulating extraembryonic gene expression and indirectly antagonizing self-renewal. *Genes Dev.* **24**, 312–326 (2010).
31. Mukherjee, S. *et al.* Sox17 and β-catenin co-occupy Wnt-responsive enhancers to govern the endoderm gene regulatory network. *Elife* **9**, (2020).
32. Liang, J. *et al.* Nanog and Oct4 associate with unique transcriptional repression complexes in embryonic stem cells. *Nat. Cell Biol.* **10**, 731–739 (2008).
33. Bernstein, B. E. *et al.* A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315–326 (2006).
34. Werner, A., Manford, A. G. & Rape, M. Ubiquitin-dependent regulation of stem cell biology. *Trends Cell Biol.* **27**, 568–579 (2017).
35. Agger, K. *et al.* UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature* **449**, 731–734 (2007).
36. Wang, C. *et al.* UTX regulates mesoderm differentiation of embryonic stem cells independent of H3K27 demethylase activity. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 15324–15329 (2012).
37. Ding, Y. *et al.* JMJD3: a critical epigenetic regulator in stem cell fate. *Cell Commun. Signal.* **19**, 72 (2021).
38. LaVaute, T. M. *et al.* Regulation of neural specification from human embryonic stem cells by BMP and FGF. *Stem Cells* **27**, 1741–1749 (2009).
39. Reubinoff, B. E. *et al.* Neural progenitors from human embryonic stem cells. *Nat. Biotechnol.* **19**, 1134–1140 (2001).
40. Tropepe, V. *et al.* Direct neural fate specification from embryonic stem cells: a primitive mammalian neural stem cell stage acquired through a default mechanism. *Neuron* **30**, 65–78 (2001).
41. Simões-Costa, M. & Bronner, M. E. Establishing neural crest identity: a gene regulatory recipe. *Development* **142**, 242–257 (2015).
42. Nikolopoulou, E., Galea, G. L., Rolo, A., Greene, N. D. E. & Copp, A. J. Neural tube closure: cellular, molecular and biomechanical mechanisms. *Development* **144**, 552–566 (2017).
43. Smith, J. R. *et al.* Inhibition of Activin/Nodal signaling promotes specification of human embryonic stem cells into neuroectoderm. *Dev. Biol.* **313**, 107–117 (2008).
44. Chang, C. & Harland, R. M. Neural induction requires continued suppression of both Smad1 and Smad2 signals during gastrulation. *Development* **134**, 3861–3872 (2007).
45. Yang, J. & Jiang, W. The role of SMAD2/3 in human embryonic stem cells. *Front. Cell Dev. Biol.* **8**, 653 (2020).

46. Chambers, S. M. *et al.* Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nat. Biotechnol.* **27**, 275–280 (2009).
47. Liang, R. *et al.* Efficient definitive endoderm differentiation from human parthenogenetic embryonic stem cells induced by activin A and Wnt3a. *Ann. Clin. Lab. Sci.* **50**, 468–473 (2020).
48. Pera, M. F. *et al.* Regulation of human embryonic stem cell differentiation by BMP-2 and its antagonist noggin. *J. Cell Sci.* **117**, 1269–1280 (2004).
49. Stuhlmiller, T. J. & García-Castro, M. I. FGF/MAPK signaling is required in the gastrula epiblast for avian neural crest induction. *Development* **139**, 289–300 (2012).
50. Yoo, Y. D., Huang, C. T., Zhang, X., Lavaute, T. M. & Zhang, S.-C. Fibroblast growth factor regulates human neuroectoderm specification through ERK1/2-PARP-1 pathway. *Stem Cells* **29**, 1975–1982 (2011).
51. Patthey, C. & Gunhaga, L. Signaling pathways regulating ectodermal cell fate choices. *Exp. Cell Res.* **321**, 11–16 (2014).
52. Patthey, C., Gunhaga, L. & Edlund, T. Early development of the central and peripheral nervous systems is coordinated by Wnt and BMP signals. *PLoS One* **3**, e1625 (2008).
53. Wilson, S. I., Graziano, E., Harland, R., Jessell, T. M. & Edlund, T. An early requirement for FGF signalling in the acquisition of neural cell fate in the chick embryo. *Curr. Biol.* **10**, 421–429 (2000).
54. Marchal, L., Luxardi, G., Thomé, V. & Kodjabachian, L. BMP inhibition initiates neural induction via FGF signaling and Zic genes. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 17437–17442 (2009).
55. Manzari-Tavakoli, A. *et al.* The cross-talks among bone morphogenetic protein (BMP) signaling and other prominent pathways involved in neural differentiation. *Front. Mol. Neurosci.* **15**, 827275 (2022).
56. Wilson, S. I. *et al.* The status of Wnt signalling regulates neural and epidermal fates in the chick embryo. *Nature* **411**, 325–330 (2001).
57. Heeg-Truesdell, E. & LaBonne, C. Neural induction in Xenopus requires inhibition of Wnt-beta-catenin signaling. *Dev. Biol.* **298**, 71–86 (2006).
58. Kempf, H. *et al.* Bulk cell density and Wnt/TGFbeta signalling regulate mesendodermal patterning of human pluripotent stem cells. *Nat. Commun.* **7**, 13602 (2016).
59. Murry, C. E. & Keller, G. Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development. *Cell* **132**, 661–680 (2008).
60. Liu, J. *et al.* Wnt/β-catenin signalling: function, biological mechanisms, and therapeutic opportunities. *Signal Transduct. Target. Ther.* **7**, 3 (2022).
61. Zhao, M., Tang, Y., Zhou, Y. & Zhang, J. Deciphering Role of Wnt Signalling in Cardiac Mesoderm and Cardiomyocyte Differentiation from Human iPSCs: Four-

dimensional control of Wnt pathway for hiPSC-CMs differentiation. *Sci. Rep.* **9**, 19389 (2019).
62. Shang, S., Hua, F. & Hu, Z.-W. The regulation of β-catenin activity and function in cancer: therapeutic opportunities. *Oncotarget* **8**, 33972–33989 (2017).
63. Tan, J. Y., Sriram, G., Rufaihah, A. J., Neoh, K. G. & Cao, T. Efficient derivation of lateral plate and paraxial mesoderm subtypes from human embryonic stem cells through GSKi-mediated differentiation. *Stem Cells Dev.* **22**, 1893–1906 (2013).
64. Kennedy, M. W. *et al.* Sp5 and Sp8 recruit β-catenin and Tcf1-Lef1 to select enhancers to activate Wnt target gene transcription. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 3545–3550 (2016).
65. Yamaguchi, T. P. Genetics of Wnt signaling during early mammalian development. *Methods Mol. Biol.* **468**, 287–305 (2008).
66. Fei, T. *et al.* Smad2 mediates Activin/Nodal signaling in mesendoderm differentiation of mouse embryonic stem cells. *Cell Res.* **20**, 1306–1318 (2010).
67. Jia, S., Ren, Z., Li, X., Zheng, Y. & Meng, A. Smad2 and smad3 are required for mesendoderm induction by transforming growth factor-beta/nodal signals in zebrafish. *J. Biol. Chem.* **283**, 2418–2426 (2008).
68. Estarás, C., Benner, C. & Jones, K. A. SMADs and YAP compete to control elongation of β-catenin:LEF-1-recruited RNAPII during hESC differentiation. *Mol. Cell* **58**, 780–793 (2015).
69. Wang, Q. *et al.* The p53 family coordinates Wnt and nodal inputs in mesendodermal differentiation of embryonic stem cells. *Cell Stem Cell* **20**, 70–86 (2017).
70. Singh, A. M. *et al.* Signaling network crosstalk in human pluripotent cells: a Smad2/3-regulated switch that controls the balance between self-renewal and differentiation. *Cell Stem Cell* **10**, 312–326 (2012).
71. Yoney, A. *et al.* WNT signaling memory is required for ACTIVIN to function as a morphogen in human gastruloids. *Elife* **7**, (2018).
72. Teo, A. K. K. *et al.* Pluripotency factors regulate definitive endoderm specification through eomesodermin. *Genes Dev.* **25**, 238–250 (2011).
73. Loh, K. M. *et al.* Mapping the pairwise choices leading from pluripotency to human bone, heart, and other mesoderm cell types. *Cell* **166**, 451–467 (2016).
74. Schröder, C. M. *et al.* EOMES establishes mesoderm and endoderm differentiation potential through SWI/SNF-mediated global enhancer remodeling. *Dev. Cell* **60**, 735-748.e5 (2025).
75. Tosic, J. *et al.* Eomes and Brachyury control pluripotency exit and germ-layer segregation by changing the chromatin state. *Nat. Cell Biol.* **21**, 1518–1531 (2019).
76. Winnier, G., Blessing, M., Labosky, P. A. & Hogan, B. L. Bone morphogenetic protein-4 is required for mesoderm formation and patterning in the mouse. *Genes Dev.* **9**, 2105–2116 (1995).

77. Evseenko, D. *et al.* Mapping the first stages of mesoderm commitment during differentiation of human embryonic stem cells. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 13742–13747 (2010).
78. Ciruna, B. & Rossant, J. FGF signaling regulates mesoderm cell fate specification and morphogenetic movement at the primitive streak. *Dev. Cell* **1**, 37–49 (2001).
79. Thiery, J. P., Acloque, H., Huang, R. Y. J. & Nieto, M. A. Epithelial-mesenchymal transitions in development and disease. *Cell* **139**, 871–890 (2009).
80. D'Amour, K. A. *et al.* Efficient differentiation of human embryonic stem cells to definitive endoderm. *Nat. Biotechnol.* **23**, 1534–1541 (2005).
81. Zaret, K. S. & Carroll, J. S. Pioneer transcription factors: establishing competence for gene expression. *Genes Dev.* **25**, 2227–2241 (2011).
82. Séguin, C. A., Draper, J. S., Nagy, A. & Rossant, J. Establishment of endoderm progenitors by SOX transcription factor expression in human embryonic stem cells. *Cell Stem Cell* **3**, 182–195 (2008).
83. Sinner, D. *et al.* Sox17 and Sox4 differentially regulate beta-catenin/T-cell factor activity and proliferation of colon carcinoma cells. *Mol. Cell. Biol.* **27**, 7802–7815 (2007).
84. Chiao, E., Kmet, M., Behr, B. & Baker, J. Derivation of human embryonic stem cells in standard and chemically defined conditions. *Methods Cell Biol.* **86**, 1–14 (2008).
85. Kroon, E. *et al.* Pancreatic endoderm derived from human embryonic stem cells generates glucose-responsive insulin-secreting cells in vivo. *Nat. Biotechnol.* **26**, 443–452 (2008).
86. Ikonomou, L. & Kotton, D. N. Derivation of endodermal progenitors from pluripotent stem cells. *J. Cell. Physiol.* **230**, 246–258 (2015).
87. Kobayashi, Y., Hayashi, R., Shibata, S., Quantock, A. J. & Nishida, K. Ocular surface ectoderm instigated by WNT inhibition and BMP4. *Stem Cell Res.* **46**, 101868 (2020).
88. Leung, A. W., Kent Morest, D. & Li, J. Y. H. Differential BMP signaling controls formation and differentiation of multipotent preplacodal ectoderm progenitors from human embryonic stem cells. *Dev. Biol.* **379**, 208–220 (2013).
89. Tchieu, J. *et al.* A modular platform for differentiation of human PSCs into all major ectodermal lineages. *Cell Stem Cell* **21**, 399-410.e7 (2017).
90. Menezes, N. A. *et al.* Cell-context response to germ layer differentiation signals is predetermined by the epigenome in regionalized epiblast populations. *Nat. Commun.* **16**, 5000 (2025).
91. Shum, I. O., Akkermann, T., Krüger, R., Zeevaert, K. & Wagner, W. Primed-to-naive conversion of pluripotent stem cells can be tracked by specific DNA methylation changes for optimized culture conditions. *bioRxiv* (2026) doi:10.64898/2026.01.16.699914.

92. Cahan, P. *et al.* CellNet: network biology applied to stem cell engineering. *Cell* **158**, 903–915 (2014).
93. Bargaje, R. *et al.* Cell population structure prior to bifurcation predicts efficiency of directed differentiation in human induced pluripotent cells. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 2271–2276 (2017).
94. Hiyoshi, H. *et al.* Identification and removal of unexpected proliferative off-target cells emerging after iPSC-derived pancreatic islet cell implantation. *Proc. Natl. Acad. Sci. U. S. A.* **121**, e2320883121 (2024).
95. Morrison, G. *et al.* Evaluation of inter-batch differences in stem-cell derived neurons. *Stem Cell Res.* **16**, 140–148 (2016).
96. Palamà, M. E. F. *et al.* Batch variability and anti-inflammatory effects of iPSC-derived mesenchymal stromal cell extracellular vesicles in osteoarthritis in vitro model. *Front. Bioeng. Biotechnol.* **13**, 1536843 (2025).
97. Fowler, J. L., Ang, L. T. & Loh, K. M. A critical look: Challenges in differentiating human pluripotent stem cells into desired cell types and organoids. *Wiley Interdiscip. Rev. Dev. Biol.* **9**, e368 (2020).
98. Chavez, A. *et al.* Highly efficient Cas9-mediated transcriptional programming. *Nat. Methods* **12**, 326–328 (2015).
99. Black, J. B. & Gersbach, C. A. Synthetic transcription factors for cell fate reprogramming. *Curr. Opin. Genet. Dev.* **52**, 13–21 (2018).
100. Vogel, A. M., Persson, K. M., Seamons, T. R. & Deans, T. L. Synthetic biology for improving cell fate decisions and tissue engineering outcomes. *Emerg. Top. Life Sci.* **3**, 631–643 (2019).
101. Yi, S. A., Zhang, Y., Rathnam, C., Pongkulapa, T. & Lee, K.-B. Bioengineering approaches for the advanced organoid research. *Adv. Mater.* **33**, e2007949 (2021).
102. Clevers, H. Modeling development and disease with organoids. *Cell* **165**, 1586–1597 (2016).
103. Lancaster, M. A. & Knoblich, J. A. Organogenesis in a dish: modeling development and disease using organoid technologies. *Science* **345**, 1247125 (2014).
104. Place, E. S., Evans, N. D. & Stevens, M. M. Complexity in biomaterials for tissue engineering. *Nat. Mater.* **8**, 457–470 (2009).
105. Langer, R. & Vacanti, J. P. Tissue engineering. *Science* **260**, 920–926 (1993).
106. Irvine, S. A. & Venkatraman, S. S. Bioprinting and differentiation of stem cells. *Molecules* **21**, E1188 (2016).
107. Gornalusse, G. G. *et al.* HLA-E-expressing pluripotent stem cells escape allogeneic responses and lysis by NK cells. *Nat. Biotechnol.* **35**, 765–772 (2017).
108. Deuse, T. *et al.* Hypoimmunogenic derivatives of induced pluripotent stem cells evade immune rejection in fully immunocompetent allogeneic recipients. *Nat. Biotechnol.* **37**, 252–258 (2019).

109. Di Stasi, A. *et al.* Inducible apoptosis as a safety switch for adoptive cell therapy. *N. Engl. J. Med.* **365**, 1673–1683 (2011).
110. Folmes, C. D. L. *et al.* Somatic oxidative bioenergetics transitions into pluripotency-dependent glycolysis to facilitate nuclear reprogramming. *Cell Metab.* **14**, 264–271 (2011).
111. Endy, D. Foundations for engineering biology. *Nature* **438**, 449–453 (2005).
112. Cameron, D. E., Bashor, C. J. & Collins, J. J. A brief history of synthetic biology. *Nat. Rev. Microbiol.* **12**, 381–390 (2014).
113. Pulecio, J., Verma, N., Mejía-Ramírez, E., Huangfu, D. & Raya, A. CRISPR/Cas9-based engineering of the epigenome. *Cell Stem Cell* **21**, 431–447 (2017).
114. Dominguez, A. A., Lim, W. A. & Qi, L. S. Beyond editing: repurposing CRISPR-Cas9 for precision epigenome regulation and genome imaging. *Nature Reviews Molecular Cell Biology* **17**, 5–15 (2016).
115. Kipniss, N. H., Dingal, P. C. D. P., Abbott, T. R., Shi, Y. & Qi. Engineering cell fate and function by multiplexed transcriptional regulation. *Nature Communications* **8**, (2017).
116. Abbott, N. J., Patabendige, A. A. K., Dolman, D. E. M., Yusof, S. R. & Begley, D. J. Structure and function of the blood-brain barrier. *Neurobiol. Dis.* **37**, 13–25 (2010).
117. Montagne, A. *et al.* Blood-brain barrier breakdown in the aging human hippocampus. *Neuron* **85**, 296–302 (2015).
118. Matsuo, K. & Nshihara, H. Rebuilding insight into the pathophysiology of Alzheimer's disease through new blood-brain barrier models. *Neural Regen. Res.* **19**, 1954–1960 (2024).
119. Stolp, H. B. & Dziegielewska, K. M. Review: Role of developmental inflammation and blood-brain barrier dysfunction in neurodevelopmental and neurodegenerative diseases. *Neuropathol. Appl. Neurobiol.* **35**, 132–146 (2009).
120. Okada, T., Suzuki, H., Travis, Z. D. & Zhang, J. H. The stroke-induced blood-brain barrier disruption: Current progress of inspection technique, mechanism, and therapeutic target. *Curr. Neuropharmacol.* **18**, 1187–1212 (2020).
121. Schaeffer, S. & Iadecola, C. Revisiting the neurovascular unit. *Nat. Neurosci.* **24**, 1198–1209 (2021).
122. Alkayed, N. J. & Cipolla, M. J. The ever-evolving concept of the neurovascular unit. *Stroke* **54**, 2178–2180 (2023).
123. Xiang, Y., Gu, Q. & Liu, D. Brain endothelial cells in blood-brain barrier regulation and neurological therapy. *Int. J. Mol. Sci.* **26**, 5843 (2025).
124. Langen, U. H., Ayloo, S. & Gu, C. Development and cell biology of the blood-brain barrier. *Annu. Rev. Cell Dev. Biol.* **35**, 591–613 (2019).
125. Daneman, R. & Prat, A. The blood–brain barrier. *Cold Spring Harb. Perspect. Biol.* **7**, (2015).

126. Furuse, M. Molecular basis of the core structure of tight junctions. *Cold Spring Harb. Perspect. Biol.* **2**, a002907 (2010).
127. Sweeney, M. D., Zhao, Z., Montagne, A., Nelson, A. R. & Zlokovic, B. V. Blood-brain barrier: From physiology to disease and back. *Physiol. Rev.* **99**, 21–78 (2019).
128. Nałęcz, K. A. Solute carriers in the blood-brain barier: Safety in abundance. *Neurochem. Res.* **42**, 795–809 (2017).
129. Qosa, H., Miller, D. S., Pasinelli, P. & Trotti, D. Regulation of ABC efflux transporters at blood-brain barrier in health and neurological disorders. *Brain Res.* **1628**, 298–316 (2015).
130. Strazielle, N. & Ghersi-Egea, J.-F. Efflux transporters in blood-brain interfaces of the developing brain. *Front. Neurosci.* **9**, 21 (2015).
131. Hervé, F., Ghinea, N. & Scherrmann, J.-M. CNS delivery via adsorptive transcytosis. *AAPS J.* **10**, 455–472 (2008).
132. Kulkarni, M. *et al.* Nanomaterials as drug delivery agents for overcoming the blood-brain barrier: A comprehensive review. *ADMET DMPK* **12**, 63–105 (2024).
133. Pardridge, W. M. Drug transport across the blood-brain barrier. *J. Cereb. Blood Flow Metab.* **32**, 1959–1972 (2012).
134. Lajoie, J. M. & Shusta, E. V. Targeting receptor-mediated transport for delivery of biologics across the blood-brain barrier. *Annu. Rev. Pharmacol. Toxicol.* **55**, 613–631 (2015).
135. Sweeney, M. D., Ayyadurai, S. & Zlokovic, B. V. Pericytes of the neurovascular unit: key functions and signaling pathways. *Nat. Neurosci.* **19**, 771–783 (2016).
136. Daneman, R., Zhou, L., Kebede, A. A. & Barres, B. A. Pericytes are required for blood-brain barrier integrity during embryogenesis. *Nature* **468**, 562–566 (2010).
137. Armulik, A. *et al.* Pericytes regulate the blood-brain barrier. *Nature* **468**, 557–561 (2010).
138. Winkler, E. A., Bell, R. D. & Zlokovic, B. V. Central nervous system pericytes in health and disease. *Nat. Neurosci.* **14**, 1398–1405 (2011).
139. Bell, R. D. *et al.* Pericytes control key neurovascular functions and neuronal phenotype in the adult brain and during brain aging. *Neuron* **111**, 3131–3132 (2023).
140. Olson, L. E. & Soriano, P. PDGFRβ signaling regulates mural cell plasticity and inhibits fat development. *Dev. Cell* **20**, 815–826 (2011).
141. Jansson, D. *et al.* A role for human brain pericytes in neuroinflammation. *J. Neuroinflammation* **11**, 104 (2014).
142. Iadecola, C. & Nedergaard, M. Glial regulation of the cerebral microvasculature. *Nat. Neurosci.* **10**, 1369–1376 (2007).
143. Benarroch, E. What are the roles of pericytes in the neurovascular unit and its disorders? *Neurology* **100**, 970–977 (2023).

144. Farahani, R. M., Rezaei-Lotfi, S., Simonian, M., Xaymardan, M. & Hunter, N. Neural microvascular pericytes contribute to human adult neurogenesis. *J. Comp. Neurol.* **527**, 780–796 (2019).
145. Yao, Y., Chen, Z.-L., Norris, E. H. & Strickland, S. Astrocytic laminin regulates pericyte differentiation and maintains blood brain barrier integrity. *Nat. Commun.* **5**, 3413 (2014).
146. Vezzani, B., Pierantozzi, E. & Sorrentino, V. Not all pericytes are born equal: Pericytes from human adult tissues present different differentiation properties. *Stem Cells Dev.* **25**, 1549–1558 (2016).
147. Dore-Duffy, P., Katychev, A., Wang, X. & Van Buren, E. CNS microvascular pericytes exhibit multipotential stem cell activity. *J. Cereb. Blood Flow Metab.* **26**, 613–624 (2006).
148. Guimarães-Camboa, N. *et al.* Pericytes of multiple organs do not behave as mesenchymal stem cells in vivo. *Cell Stem Cell* **20**, 345-359.e5 (2017).
149. Göritz, C. *et al.* A pericyte origin of spinal cord scar tissue. *Science* **333**, 238–242 (2011).
150. Bélanger, M., Allaman, I. & Magistretti, P. J. Brain energy metabolism: focus on astrocyte-neuron metabolic cooperation. *Cell Metab.* **14**, 724–738 (2011).
151. Bell, A. H., Miller, S. L., Castillo-Melendez, M. & Malhotra, A. The neurovascular unit: Effects of brain insults during the perinatal period. *Front. Neurosci.* **13**, 1452 (2019).
152. Nuriya, M., Shinotsuka, T. & Yasui, M. Diffusion properties of molecules at the blood-brain interface: potential contributions of astrocyte endfeet to diffusion barrier functions. *Cereb. Cortex* **23**, 2118–2126 (2013).
153. Haj-Yasein, N. N. *et al.* Glial-conditional deletion of aquaporin-4 (Aqp4) reduces blood-brain water uptake and confers barrier function on perivascular astrocyte endfeet. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 17815–17820 (2011).
154. Rouach, N., Koulakoff, A., Abudara, V., Willecke, K. & Giaume, C. Astroglial metabolic networks sustain hippocampal synaptic transmission. *Science* **322**, 1551–1555 (2008).
155. Sofroniew, M. V. & Vinters, H. V. Astrocytes: biology and pathology. *Acta Neuropathol.* **119**, 7–35 (2010).
156. Hayakawa, K., Pham, L.-D. D., Katusic, Z. S., Arai, K. & Lo, E. H. Astrocytic high-mobility group box 1 promotes endothelial progenitor cell-mediated neurovascular remodeling during stroke recovery. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 7505–7510 (2012).
157. Helms, H. C. *et al.* In vitro models of the blood-brain barrier: An overview of commonly used brain endothelial cell culture models and guidelines for their use. *J. Cereb. Blood Flow Metab.* **36**, 862–890 (2016).

158. Roux, F. & Couraud, P.-O. Rat brain endothelial cell lines for the study of blood-brain barrier permeability and transport functions. *Cell. Mol. Neurobiol.* **25**, 41–58 (2005).
159. Hawkins, B. T. & Egleton, R. D. Pathophysiology of the blood-brain barrier: animal models and methods. *Curr. Top. Dev. Biol.* **80**, 277–309 (2008).
160. O'Brown, N. M., Pfau, S. J. & Gu, C. Bridging barriers: a comparative look at the blood-brain barrier across organisms. *Genes Dev.* **32**, 466–478 (2018).
161. Song, H. W. *et al.* Transcriptomic comparison of human and mouse brain microvessels. *Sci. Rep.* **10**, 12358 (2020).
162. Weksler, B., Romero, I. A. & Couraud, P.-O. The hCMEC/D3 cell line as a model of the human blood brain barrier. *Fluids Barriers CNS* **10**, 16 (2013).
163. Weksler, B. B. *et al.* Blood-brain barrier-specific properties of a human adult brain endothelial cell line. *FASEB J.* **19**, 1872–1874 (2005).
164. Sivandzade, F. & Cucullo, L. In-vitro blood-brain barrier modeling: A review of modern and fast-advancing technologies. *J. Cereb. Blood Flow Metab.* **38**, 1667–1681 (2018).
165. Ohtsuki, S. *et al.* Quantitative targeted absolute proteomic analysis of transporters, receptors and junction proteins for validation of human cerebral microvascular endothelial cell line hCMEC/D3 as a human blood-brain barrier model. *Mol. Pharm.* **10**, 289–296 (2013).
166. Eigenmann, D. E. *et al.* Comparative study of four immortalized human brain capillary endothelial cell lines, hCMEC/D3, hBMEC, TY10, and BB19, and optimization of culture conditions, for an in vitro blood-brain barrier model for drug permeability studies. *Fluids Barriers CNS* **10**, 33 (2013).
167. Lippmann, E. S. *et al.* Derivation of blood-brain barrier endothelial cells from human pluripotent stem cells. *Nat. Biotechnol.* **30**, 783–791 (2012).
168. Lippmann, E. S., Al-Ahmad, A., Azarin, S. M., Palecek, S. P. & Shusta, E. V. A retinoic acid-enhanced, multicellular human blood-brain barrier model derived from stem cell sources. *Sci. Rep.* **4**, 4160 (2014).
169. Ribecco-Lutkiewicz, M. *et al.* A novel human induced pluripotent stem cell blood-brain barrier model: Applicability to study antibody-triggered receptor-mediated transcytosis. *Sci. Rep.* **8**, 1873 (2018).
170. Appelt-Menzel, A. *et al.* Human iPSC-derived blood-brain barrier models: Valuable tools for preclinical drug discovery and development? *Curr. Protoc. Stem Cell Biol.* **55**, e122 (2020).
171. Choi, J.-H., Santhosh, M. & Choi, J.-W. In vitro blood-brain barrier-integrated neurological disorder models using a microfluidic device. *Micromachines (Basel)* **11**, 21 (2019).
172. Niazi, S. K. Non-invasive drug delivery across the blood-brain barrier: A prospective analysis. *Pharmaceutics* **15**, 2599 (2023).

173. Sethi, B., Kumar, V., Mahato, K., Coulter, D. W. & Mahato, R. I. Recent advances in drug delivery and targeting to the brain. *J. Control. Release* **350**, 668–687 (2022).
174. Burgio, F. *et al.* A perfused in vitro human iPSC-derived blood-brain barrier faithfully mimics transferrin receptor-mediated transcytosis of therapeutic antibodies. *Cell. Mol. Neurobiol.* **43**, 4173–4187 (2023).
175. Simonneau, C. *et al.* Investigating receptor-mediated antibody transcytosis using blood-brain barrier organoid arrays. *Fluids Barriers CNS* **18**, 43 (2021).
176. Jezierski, A. *et al.* Mouse embryonic stem cell-derived blood-brain barrier model: applicability to studying antibody triggered receptor mediated transcytosis. *Fluids Barriers CNS* **20**, 36 (2023).
177. Chen, W. *et al.* BBB-crossing adeno-associated virus vector: An excellent gene delivery tool for CNS disease treatment. *J. Control. Release* **333**, 129–138 (2021).
178. Drouyer, M. *et al.* Enhanced AAV transduction across preclinical CNS models: A comparative study in human brain organoids with cross-species evaluations. *Mol. Ther. Nucleic Acids* **35**, 102264 (2024).
179. Ding, Y., Shusta, E. V. & Palecek, S. P. Integrating in vitro disease models of the neurovascular unit into discovery and development of neurotherapeutics. *Curr. Opin. Biomed. Eng.* **20**, 100341 (2021).
180. González-Gallego, J. *et al.* A fully iPS-cell-derived 3D model of the human blood–brain barrier for exploring neurovascular disease mechanisms and therapeutic interventions. *Nat. Neurosci.* 1–14 (2025).
181. Sela, M. *et al.* AI-validated brain targeted mRNA lipid nanoparticles with neuronal tropism. *ACS Nano* **19**, 36106–36128 (2025).
182. Ferrari, I. *et al.* Microfluidic-based technologies for crossing the blood-brain barrier against Alzheimer's disease: Novel strategies and challenges. *Int. J. Mol. Sci.* **26**, 9478 (2025).
183. Stanimirovic, D. B. & Jezierski, A. Blood–brain barrier models in pre-clinical assessment of CNS-targeting cell and gene therapies. *Cell Gene Ther. Insights* **5**, 937–947 (2019).
184. Badawi, A. H. *et al.* In vitro blood-brain barrier models for neuroinfectious diseases: A narrative review. *Curr. Neuropharmacol.* **22**, 1344–1373 (2024).
185. Richards, A. *et al.* A human blood-brain barrier model reveals pericytes as critical regulators of viral neuroinvasion. *iScience* **29**, 114443 (2026).
186. Cheng, Y. *et al.* Intrinsic antiviral immunity of barrier cells revealed by an iPSC-derived blood-brain barrier cellular model. *Cell Rep.* **39**, 110885 (2022).
187. Alimonti, J. B. *et al.* Zika virus crosses an in vitro human blood brain barrier model. *Fluids Barriers CNS* **15**, 15 (2018).
188. Yamada, S. *et al.* SARS-CoV-2 causes dysfunction in human iPSC-derived brain microvascular endothelial cells potentially by modulating the Wnt signaling pathway. *Fluids Barriers CNS* **21**, 32 (2024).

189. Jacob, F. *et al.* Human pluripotent stem cell-derived neural cells and brain organoids reveal SARS-CoV-2 neurotropism predominates in choroid plexus epithelium. *Cell Stem Cell* **27**, 937-950.e9 (2020).
190. Krasemann, S. *et al.* The blood-brain barrier is dysregulated in COVID-19 and serves as a CNS entry route for SARS-CoV-2. *Stem Cell Reports* **17**, 307–320 (2022).
191. Gopinadhan, A. *et al.* A human pluripotent stem cell-derived in vitro model of the blood-brain barrier in cerebral malaria. *Fluids Barriers CNS* **21**, 38 (2024).
192. Huang, J. *et al.* Application of blood brain barrier models in pre-clinical assessment of glioblastoma-targeting CAR-T based immunotherapies. *Fluids Barriers CNS* **19**, 38 (2022).
193. Pang, Z., Wang, Z., Li, F., Feng, C. & Mu, X. Current progress of CAR-NK therapy in cancer treatment. *Cancers (Basel)* **14**, 4318 (2022).
194. Santomasso, B. D. *et al.* Data from clinical and biological correlates of neurotoxicity associated with CAR T-cell therapy in patients with B-cell acute lymphoblastic leukemia. *American Association for Cancer Research* (2023) doi:10.1158/2159-8290.c.6546760.v1.
195. Gust, J. *et al.* Endothelial activation and blood-brain barrier disruption in neurotoxicity after adoptive immunotherapy with CD19 CAR-T cells. *Cancer Discov.* **7**, 1404–1419 (2017).
196. Gust, J., Taraseviciute, A. & Turtle, C. J. Neurotoxicity associated with CD19-targeted CAR-T cell therapies. *CNS Drugs* **32**, 1091–1101 (2018).
197. Wälchli, T. *et al.* Single-cell atlas of the human brain vasculature across development, adulthood and disease. *Nature* **632**, 603–613 (2024).
198. Shahriar, S. *et al.* VEGF-A-mediated venous endothelial cell proliferation results in neoangiogenesis during neuroinflammation. *Nat. Neurosci.* **27**, 1904–1917 (2024).
199. Anderson, R. H. & Francis, K. R. Modeling rare diseases with induced pluripotent stem cell technology. *Mol. Cell. Probes* **40**, 52–59 (2018).
200. Dao, L. *et al.* Modeling blood-brain barrier formation and cerebral cavernous malformations in human PSC-derived organoids. *Cell Stem Cell* **31**, 818-833.e11 (2024).
201. Bosworth, A. M., Faley, S. L., Bellan, L. M. & Lippmann, E. S. Modeling neurovascular disorders and therapeutic outcomes with human-induced pluripotent stem cells. *Front. Bioeng. Biotechnol.* **5**, 87 (2017).
202. McTague, A., Rossignoli, G., Ferrini, A., Barral, S. & Kurian, M. A. Genome editing in iPSC-based neural systems: From disease models to future therapeutic strategies. *Front. Genome Ed.* **3**, 630600 (2021).
203. Williams, L. M. *et al.* Prolonged culturing of iPSC-derived brain endothelial-like cells is associated with quiescence, downregulation of glycolysis, and resistance to disruption by an Alzheimer's brain milieu. *Fluids Barriers CNS* **19**, 10 (2022).

204. Wu, Y.-C., Sonninen, T.-M., Peltonen, S., Koistinaho, J. & Lehtonen, Š. Blood-brain barrier and neurodegenerative diseases-modeling with iPSC-derived brain cells. *Int. J. Mol. Sci.* **22**, 7710 (2021).
205. Linville, R. M. *et al.* Brain microvascular endothelial cell dysfunction in an isogenic juvenile iPSC model of Huntington's disease. *Fluids Barriers CNS* **19**, 54 (2022).
206. Penney, J., Ralvenius, W. T. & Tsai, L. H. Modeling Alzheimer's disease with iPSC-derived brain cells. *Molecular Psychiatry* vol. 25 148–167 Preprint at https://doi.org/10.1038/s41380-019-0468-3 (2020).
207. Oikari, L. E. *et al.* Altered brain endothelial cell phenotype from a familial Alzheimer mutation and its potential implications for amyloid clearance and drug delivery. *Stem Cell Reports* **14**, 924–939 (2020).
208. Sonninen, T.-M. *et al.* LRRK2 G2019S mutated iPSC-derived endothelial cells exhibit increased α-synuclein, mitochondrial impairment, and altered inflammatory responses. *Int. J. Mol. Sci.* **25**, 12874 (2024).
209. Pediaditakis, I. *et al.* Modeling alpha-synuclein pathology in a human brain-chip to assess blood-brain barrier disruption. *Nat. Commun.* **12**, 5907 (2021).
210. Lall, D. *et al.* An organ-chip model of sporadic ALS using iPSC-derived spinal cord motor neurons and an integrated blood-brain-like barrier. *Cell Stem Cell* **32**, 1139-1153.e7 (2025).
211. Sances, S. *et al.* Human iPSC-derived endothelial cells and microengineered organ-chip enhance neuronal development. *Stem Cell Reports* **10**, 1222–1236 (2018).
212. Alexander, N. G. *et al.* Using a brain-like endothelial cell differentiation to characterize the CS79iBRCA-n2 BRCA1 mutated patient derived stem cell line. *Front. Cell Dev. Biol.* **13**, 1516669 (2025).
213. Moseki, M. C., Makhzoum, A., Masisi, K. & Gaobotse, G. Recent progress in the use of induced pluripotent stem cells in novel drug development over the past 5 years (2020–2024). *Discov. Biotechnol.* **2**, (2025).
214. Lu, T. M. *et al.* Pluripotent stem cell-derived epithelium misidentified as brain microvascular endothelium requires ETS factors to acquire vascular fate. *Proc. Natl. Acad. Sci. U. S. A.* **118**, (2021).
215. Gastfriend, B. D., Palecek, S. P. & Shusta, E. V. Modeling the blood-brain barrier: Beyond the endothelial cells. *Curr. Opin. Biomed. Eng.* **5**, 6–12 (2018).
216. Campisi, M. *et al.* 3D self-organized microvascular model of the human blood-brain barrier with endothelial cells, pericytes and astrocytes. *Biomaterials* **180**, 117–129 (2018).
217. Corada, M. *et al.* Fine-tuning of Sox17 and canonical Wnt coordinates the permeability properties of the blood-brain barrier. *Circ. Res.* **124**, 511–525 (2019).
218. Birdsey, G. M. *et al.* The endothelial transcription factor ERG promotes vascular stability and growth through Wnt/β-catenin signaling. *Dev. Cell* **32**, 82–96 (2015).

219. Sadanandan, J. *et al.* Key epigenetic and signaling factors in the formation and maintenance of the blood-brain barrier. *Elife* **12**, (2024).
220. Bonney, S., Dennison, B. J. C., Wendlandt, M. & Siegenthaler, J. A. Retinoic acid regulates endothelial β-catenin expression and pericyte numbers in the developing brain vasculature. *Front. Cell. Neurosci.* **12**, 476 (2018).
221. Canfield, S. G. *et al.* An isogenic blood-brain barrier model comprising brain endothelial cells, astrocytes, and neurons derived from human induced pluripotent stem cells. *J. Neurochem.* **140**, 874–888 (2017).
222. Noh, K. M., Park, S.-J., Moon, S.-H. & Jung, S. Y. Extracellular matrix cues regulate the differentiation of pluripotent stem cell-derived endothelial cells. *Front. Cardiovasc. Med.* **10**, 1169331 (2023).
223. Choi, J.-W., Youn, J., Kim, D. S. & Park, T.-E. Human iPS-derived blood-brain barrier model exhibiting enhanced barrier properties empowered by engineered basement membrane. *Biomaterials* **293**, 121983 (2023).
224. McConnell, H. L. & Mishra, A. Cells of the blood-brain barrier: An overview of the neurovascular unit in health and disease. *Methods Mol. Biol.* **2492**, 3–24 (2022).
225. Obermeier, B., Daneman, R. & Ransohoff, R. M. Development, maintenance and disruption of the blood-brain barrier. *Nat. Med.* **19**, 1584–1596 (2013).
226. Kim, H. *et al.* Reactive astrocytes transduce inflammation in a blood-brain barrier model through a TNF-STAT3 signaling axis and secretion of alpha 1-antichymotrypsin. *Nat. Commun.* **13**, 6581 (2022).
227. Ulbrich, H., Eriksson, E. E. & Lindbom, L. Leukocyte and endothelial cell adhesion molecules as targets for therapeutic interventions in inflammatory disease. *Trends Pharmacol. Sci.* **24**, 640–647 (2003).
228. Lu, T. M. *et al.* Human induced pluripotent stem cell-derived brain endothelial cells: Current controversies. *Front. Physiol.* **12**, 642812 (2021).
229. Yan, L., Moriarty, R. A. & Stroka, K. M. Recent progress and new challenges in modeling of human pluripotent stem cell-derived blood-brain barrier. *Theranostics* **11**, 10148–10170 (2021).
230. Girard, S. D. *et al.* High and low permeability of human pluripotent stem cell-derived blood-brain barrier models depend on epithelial or endothelial features. *FASEB J.* **37**, e22770 (2023).
231. Nishihara, H. *et al.* Advancing human induced pluripotent stem cell-derived blood-brain barrier models for studying immune cell interactions. *FASEB J.* **34**, 16693–16715 (2020).
232. Carcamo-Orive, I. *et al.* Analysis of Transcriptional Variability in a Large Human iPSC Library Reveals Genetic and Non-genetic Determinants of Heterogeneity. *Cell Stem Cell* **20**, 518-532.e9 (2017).

233. Galiakberova, A. A. *et al.* Different iPSC-derived neural stem cells shows various spectrums of spontaneous differentiation during long term cultivation. *Front. Mol. Neurosci.* **16**, 1037902 (2023).
234. Cui, A. *et al.* Generation of hiPSC-derived brain microvascular endothelial cells using directed differentiation and transcriptional reprogramming. *Arterioscler. Thromb. Vasc. Biol.* https://doi.org/10.1161/ATVBAHA.125.323397 (2025) doi:10.1161/ATVBAHA.125.323397.
235. Gastfriend, B. D. & Daneman, R. Development of the blood-brain barrier. *Development* **153**, (2026).
236. Gastfriend, B. D. *et al.* Wnt signaling mediates acquisition of blood–brain barrier properties in naïve endothelium derived from human pluripotent stem cells. *Elife* **10**, (2021).
237. Laksitorini, M. D., Yathindranath, V., Xiong, W., Hombach-Klonisch, S. & Miller, D. W. Modulation of Wnt/β-catenin signaling promotes blood-brain barrier phenotype in cultured brain endothelial cells. *Sci. Rep.* **9**, 19718 (2019).
238. Mizee, M. R. *et al.* Retinoic acid induces blood-brain barrier development. *J. Neurosci.* **33**, 1660–1671 (2013).
239. Est, C. B. & Murphy, R. M. An in vitro model for vitamin A transport across the human blood-brain barrier. *bioRxivorg* (2023) doi:10.1101/2023.04.11.536348.
240. Sabbagh, M. F. *et al.* Transcriptional and epigenomic landscapes of CNS and non-CNS vascular endothelial cells. *Elife* **7**, (2018).
241. Lippmann, E. S., Azarin, S. M., Palecek, S. P. & Shusta, E. V. Commentary on human pluripotent stem cell-based blood-brain barrier models. *Fluids Barriers CNS* **17**, 64 (2020).
242. Wilson, H. K., Canfield, S. G., Hjortness, M. K., Palecek, S. P. & Shusta, E. V. Exploring the effects of cell seeding density on the differentiation of human pluripotent stem cells to brain microvascular endothelial cells. *Fluids Barriers CNS* **12**, 13 (2015).
243. Hollmann, E. K. *et al.* Accelerated differentiation of human induced pluripotent stem cells to blood-brain barrier endothelial cells. *Fluids Barriers CNS* **14**, 9 (2017).
244. Neal, E. H. *et al.* A simplified, fully defined differentiation scheme for producing blood-brain barrier endothelial cells from human iPSCs. *Stem Cell Reports* **12**, 1380–1388 (2019).
245. Park, T.-E. *et al.* Hypoxia-enhanced Blood-Brain Barrier Chip recapitulates human barrier function and shuttling of drugs and antibodies. *Nat. Commun.* **10**, 2621 (2019).
246. Praça, C. *et al.* Derivation of brain capillary-like endothelial cells from human pluripotent stem cell-derived endothelial progenitor cells. *Stem Cell Reports* **13**, 599–611 (2019).

247. Linville, R. M. *et al.* Human iPSC-derived blood-brain barrier microvessels: validation of barrier function and endothelial cell behavior. *Biomaterials* **190–191**, 24–37 (2019).
248. Faley, S. L. *et al.* IPSC-derived brain endothelium exhibits stable, long-term barrier function in perfused hydrogel scaffolds. *Stem Cell Reports* **12**, 474–487 (2019).
249. Selvakumaran, J. *et al.* An induced pluripotent stem cell-derived human blood-brain barrier (BBB) model to test the Crossing by adeno-associated virus (AAV) vectors and antisense oligonucleotides. *Biomedicines* **11**, 2700 (2023).
250. Ozgür, B. *et al.* Characterization of an iPSC-based barrier model for blood-brain barrier investigations using the SBAD0201 stem cell line. *Fluids Barriers CNS* **20**, 96 (2023).
251. Workman, M. J. & Svendsen, C. N. Recent advances in human iPSC-derived models of the blood-brain barrier. *Fluids Barriers CNS* **17**, 30 (2020).
252. Warren, L. *et al.* Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. *Cell Stem Cell* **7**, 618–630 (2010).
253. Lyra-Leite, D. M. *et al.* A review of protocols for human iPSC culture, cardiac differentiation, subtype-specification, maturation, and direct reprogramming. *STAR Protoc.* **3**, 101560 (2022).
254. Ginsberg, M. *et al.* Efficient direct reprogramming of mature amniotic cells into endothelial cells by ETS factors and TGFβ suppression. *Cell* **151**, 559–575 (2012).
255. Morita, R. *et al.* ETS transcription factor ETV2 directly converts human fibroblasts into functional endothelial cells. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 160–165 (2015).
256. Lee, S. *et al.* Direct reprogramming of human dermal fibroblasts into endothelial cells using ER71/ETV2. *Circ. Res.* **120**, 848–861 (2017).
257. Cheng, F. *et al.* Conversion of human adipose-derived stem cells into functional and expandable endothelial-like cells for cell-based therapies. *Stem Cell Res. Ther.* **9**, 350 (2018).
258. Ding, Y. *et al.* ETV2 Overexpression promotes efficient differentiation of pluripotent stem cells to endothelial cells. *Biotechnol. Bioeng.* https://doi.org/10.1002/bit.28979 (2025) doi:10.1002/bit.28979.
259. Patel, R. *et al.* Generation of hiPSC-derived brain microvascular endothelial cells using a combination of directed differentiation and transcriptional reprogramming strategies. *bioRxivorg* https://doi.org/10.1101/2024.04.03.588012 (2024) doi:10.1101/2024.04.03.588012.
260. Roudnicky, F. *et al.* Identification of a combination of transcription factors that synergistically increases endothelial cell barrier resistance. *Sci. Rep.* **10**, (2020).
261. Fontijn, R. D. *et al.* SOX-18 controls endothelial-specific claudin-5 gene expression and barrier function. *Am. J. Physiol. Heart Circ. Physiol.* **294**, H891-900 (2008).

262. Darby, I. A. *et al.* Sox18 is transiently expressed during angiogenesis in granulation tissue of skin wounds with an identical expression pattern to Flk-1 mRNA. *Lab. Invest.* **81**, 937–943 (2001).
263. Hupe, M. *et al.* Gene expression profiles of brain endothelial cells during embryonic development at bulk and single-cell levels. *Sci. Signal.* **10**, (2017).
264. Reyahi, A. *et al.* Foxf2 is required for brain pericyte differentiation and development and maintenance of the blood-brain barrier. *Dev. Cell* **34**, 19–32 (2015).
265. Cottarelli, A. *et al.* Fgfbp1 promotes blood-brain barrier development by regulating collagen IV deposition and maintaining Wnt/β-catenin signaling. *Development* **147**, dev185140 (2020).
266. Zhang, H., Yamaguchi, T. & Kawabata, K. The maturation of iPS cell-derived brain microvascular endothelial cells by inducible-SOX18 expression. *Fluids Barriers CNS* **20**, 10 (2023).
267. Zhou, Y., Williams, J., Smallwood, P. M. & Nathans, J. Sox7, Sox17, and Sox18 cooperatively regulate vascular development in the mouse retina. *PLoS One* **10**, e0143650 (2015).
268. Downes, M. & Koopman, P. SOX18 and the transcriptional regulation of blood vessel development. *Trends Cardiovasc. Med.* **11**, 318–324 (2001).
269. Wang, Y. *et al.* Norrin/Frizzled4 signaling in retinal vascular development and blood brain barrier plasticity. *Cell* **151**, 1332–1344 (2012).
270. Sabbagh, M. F. & Nathans, J. A genome-wide view of the de-differentiation of central nervous system endothelial cells in culture. *Elife* **9**, (2020).
271. Patsch, C. *et al.* Generation of vascular endothelial and smooth muscle cells from human pluripotent stem cells. *Nat. Cell Biol.* **17**, 994–1003 (2015).
272. Porkoláb, G. *et al.* Synergistic induction of blood-brain barrier properties. *Proc. Natl. Acad. Sci. U. S. A.* **121**, e2316006121 (2024).
273. Ramezankhani, R. *et al.* Identification of cell fate determining transcription factors for generating brain endothelial cells. *Stem Cell Rev Rep* **21**, 744–766 (2025).
274. Tamhankar, S. *et al.* Forward programming identifies inducers of blood-brain barrier properties in human pluripotent stem cell-derived endothelial cells. *bioRxivorg* (2026) doi:10.64898/2026.02.02.699492.
275. Cantu Gutierrez, M. E. *et al.* Mapping the transcriptional and epigenetic landscape of organotypic endothelial diversity in the developing and adult mouse. *Nat. Cardiovasc. Res.* **4**, 473–495 (2025).
276. Iwasaki, H. *et al.* The order of expression of transcription factors directs hierarchical specification of hematopoietic lineages. *Genes Dev.* **20**, 3010–3021 (2006).
277. Zhang, H., Yamaguchi, T., Kokubu, Y. & Kawabata, K. Transient ETV2 expression promotes the generation of mature endothelial cells from human pluripotent stem cells. *Biol. Pharm. Bull.* **45**, 483–490 (2022).

278. Roney, I. J., Rudner, A. D., Couture, J. F. & Kærn, M. Improvement of the reverse tetracycline transactivator by single amino acid substitutions that reduce leaky target gene expression to undetectable levels. *Sci. Rep.* **6**, (2016).
279. Kazemian, P. *et al.* Lipid-nanoparticle-based delivery of CRISPR/Cas9 genome-editing components. *Mol. Pharm.* **19**, 1669–1686 (2022).
280. Chen, T. *et al.* DKK3 (dikkopf-3) transdifferentiates fibroblasts into functional endothelial cells-brief report. *Arterioscler. Thromb. Vasc. Biol.* **39**, 765–773 (2019).
281. Han, J.-K. *et al.* Direct conversion of adult human fibroblasts into functional endothelial cells using defined factors. *Biomaterials* **272**, 120781 (2021).
282. Wang, K. *et al.* Robust differentiation of human pluripotent stem cells into endothelial cells via temporal modulation of ETV2 with modified mRNA. *Sci. Adv.* **6**, eaba7606 (2020).
283. Qian, T. *et al.* Directed differentiation of human pluripotent stem cells to blood-brain barrier endothelial cells. *Sci. Adv.* **3**, e1701679 (2017).
284. Luo, A. C. *et al.* A streamlined method to generate endothelial cells from human pluripotent stem cells via transient doxycycline-inducible ETV2 activation. *Angiogenesis* **27**, 779–795 (2024).
285. Erzar, I. *et al.* Protocol for generation of endothelial cells from human-derived iPSCs and integration into mouse brain explants. *STAR Protoc.* **6**, 104239 (2025).
286. Zhang, S. *et al.* ETV2 mediated differentiation of human pluripotent stem cells results in functional endothelial cells for engineering advanced vascularized microphysiological models. *Adv. Healthc. Mater.* e04849 (2026).
287. Hwang, R. W. *et al.* Accelerated reprogramming of hiPSCs into functional brain endothelial-like cells using multiplexed CRISPR activation. *Sci. Rep.* 1–17 (2026).
288. Kelly, T. J., Jr & Smith, H. O. A restriction enzyme from Hemophilus influenzae. II. *J. Mol. Biol.* **51**, 393–409 (1970).
289. Smith, H. O. & Wilcox, K. W. A restriction enzyme from Hemophilus influenzae. I. Purification and general properties. *J. Mol. Biol.* **51**, 379–391 (1970).
290. Smithies, O., Gregg, R. G., Boggs, S. S., Koralewski, M. A. & Kucherlapati, R. S. Insertion of DNA sequences into the human chromosomal beta-globin locus by homologous recombination. *Nature* **317**, 230–234 (1985).
291. Thomas, K. R. & Capecchi, M. R. Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell* **51**, 503–512 (1987).
292. Capecchi, M. R. Altering the genome by homologous recombination. *Science* **244**, 1288–1292 (1989).
293. Rouet, P., Smih, F. & Jasin, M. Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Mol. Cell. Biol.* **14**, 8096–8106 (1994).

294. Choulika, A., Perrin, A., Dujon, B. & Nicolas, J. F. Induction of homologous recombination in mammalian chromosomes by using the I-SceI system of Saccharomyces cerevisiae. *Mol. Cell. Biol.* **15**, 1968–1973 (1995).
295. Kim, Y. G., Cha, J. & Chandrasegaran, S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 1156–1160 (1996).
296. Bibikova, M. *et al.* Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. *Mol. Cell. Biol.* **21**, 289–297 (2001).
297. Porteus, M. H. & Baltimore, D. Chimeric nucleases stimulate gene targeting in human cells. *Science* **300**, 763 (2003).
298. Christian, M. *et al.* Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* **186**, 757–761 (2010).
299. Boch, J. *et al.* Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326**, 1509–1512 (2009).
300. Miller, J. C. *et al.* A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.* **29**, 143–148 (2011).
301. Gaj, T., Gersbach, C. A. & Barbas, C. F., 3rd. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol.* **31**, 397–405 (2013).
302. Holkers, M. *et al.* Differential integrity of TALE nuclease genes following adenoviral and lentiviral vector gene transfer into human cells. *Nucleic Acids Res.* **41**, e63 (2013).
303. Joung, J. K. & Sander, J. D. TALENs: a widely applicable technology for targeted genome editing. *Nat. Rev. Mol. Cell Biol.* **14**, 49–55 (2013).
304. Cermak, T. *et al.* Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res.* **39**, e82 (2011).
305. Barrangou, R. *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709–1712 (2007).
306. Garneau, J. E. *et al.* The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **468**, 67–71 (2010).
307. Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826 (2013).
308. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).
309. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
310. Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J. & Almendros, C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* **155**, 733–740 (2009).
311. Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C. & Doudna, J. A. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* **507**, 62–67 (2014).

312. Nishimasu, H. *et al.* Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell* **156**, 935–949 (2014).
313. Brouns, S. J. J. *et al.* Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321**, 960–964 (2008).
314. Marraffini, L. A. & Sontheimer, E. J. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* **322**, 1843–1845 (2008).
315. Hale, C. R. *et al.* RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* **139**, 945–956 (2009).
316. Özcan, A. *et al.* Programmable RNA targeting with the single-protein CRISPR effector Cas7-11. *Nature* **597**, 720–725 (2021).
317. Zetsche, B. *et al.* Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* **163**, 759–771 (2015).
318. Pausch, P. *et al.* CRISPR-CasΦ from huge phages is a hypercompact genome editor. *Science* **369**, 333–337 (2020).
319. Harrington, L. B. *et al.* Programmed DNA destruction by miniature CRISPR-Cas14 enzymes. *Science* **362**, 839–842 (2018).
320. Abudayyeh, O. O. *et al.* C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science* **353**, aaf5573 (2016).
321. Shmakov, S. *et al.* Discovery and functional characterization of diverse Class 2 CRISPR-Cas systems. *Mol. Cell* **60**, 385–397 (2015).
322. Anzalone, A. V. *et al.* Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* **576**, 149–157 (2019).
323. Nelson, J. W. *et al.* Engineered pegRNAs improve prime editing efficiency. *Nat. Biotechnol.* **40**, 402–410 (2022).
324. Suzuki, K. *et al.* In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. *Nature* **540**, 144–149 (2016).
325. Ihry, R. J. *et al.* p53 inhibits CRISPR–Cas9 engineering in human pluripotent stem cells. *Nat. Med.* **24**, 939–946 (2018).
326. Kosicki, M., Tomberg, K. & Bradley, A. Repair of double-strand breaks induced by CRISPR-Cas9 leads to large deletions and complex rearrangements. *Nat. Biotechnol.* **36**, 765–771 (2018).
327. Alberts, B. *et al. Site-Specific Recombination*. (Garland Science, London, England, 2002).
328. Sauer, B. Site-specific recombination: developments and applications. *Curr. Opin. Biotechnol.* **5**, 521–527 (1994).
329. Smith, M. C. M., Brown, W. R. A., McEwan, A. R. & Rowley, P. A. Site-specific recombination by phiC31 integrase and other large serine recombinases. *Biochem. Soc. Trans.* **38**, 388–394 (2010).

330. Russell, J. P., Chang, D. W., Tretiakova, A. & Padidam, M. Phage Bxb1 integrase mediates highly efficient site-specific recombination in mammalian cells. *Biotechniques* **40**, 460, 462, 464 (2006).
331. Yarnall, M. T. N. *et al.* Drag-and-drop genome insertion of large sequences without double-strand DNA cleavage using CRISPR-directed integrases. *Nat. Biotechnol.* 1–13 (2022).
332. Roelle, S. M., Kamath, N. D. & Matreyek, K. A. Mammalian genomic manipulation with orthogonal Bxb1 DNA recombinase sites for the functional characterization of protein variants. *ACS Synth. Biol.* **12**, 3352–3365 (2023).
333. Ghosh, P., Kim, A. I. & Hatfull, G. F. The orientation of mycobacteriophage Bxb1 integration is solely dependent on the central dinucleotide of attP and attB. *Mol. Cell* **12**, 1101–1111 (2003).
334. Jusiak, B. *et al.* Comparison of integrases identifies Bxb1-GA mutant as the most efficient site-specific integrase system in mammalian cells. *ACS Synth. Biol.* **8**, 16–24 (2019).
335. Pandey, S. *et al.* Efficient site-specific integration of large genes in mammalian cells via continuously evolved recombinases and prime editing. *Nat. Biomed. Eng.* **9**, 22–39 (2025).
336. Blanch-Asensio, A. *et al.* STRAIGHT-IN Dual: a platform for dual, single-copy integrations of DNA payloads and gene circuits into human induced pluripotent stem cell. *bioRxivorg* (2025) doi:10.1101/2024.10.17.616637.
337. Gervais, N. C. *et al.* Development and applications of a CRISPR activation system for facile genetic overexpression in Candida albicans. *G3 (Bethesda)* **13**, jkac301 (2023).
338. Larson, M. H. *et al.* CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nat. Protoc.* **8**, 2180–2196 (2013).
339. Qi, L. S. *et al.* Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173–1183 (2013).
340. Campa, C. C., Weisbach, N. R., Santinha, A. J., Incarnato, D. & Platt, R. J. Multiplexed genome engineering by Cas12a and CRISPR arrays encoded on single transcripts. *Nat. Methods* **16**, 887–893 (2019).
341. Bendixen, L., Jensen, T. I. & Bak, R. O. CRISPR-Cas-mediated transcriptional modulation: The therapeutic promises of CRISPRa and CRISPRi. *Mol. Ther.* **31**, 1920–1937 (2023).
342. Maeder, M. L. *et al.* CRISPR RNA-guided activation of endogenous human genes. *Nat. Methods* **10**, 977–979 (2013).
343. Perez-Pinera, P. *et al.* RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat. Methods* **10**, 973–976 (2013).
344. Konermann, S. *et al.* Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517**, 583–588 (2015).

345. Tanenbaum, M. E., Gilbert, L. A., Qi, L. S., Weissman, J. S. & Vale, R. D. A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell* **159**, 635–646 (2014).
346. Chavez, A. *et al.* Comparison of Cas9 activators in multiple species. *Nat. Methods* **13**, 563–567 (2016).
347. Lino, C. A., Harper, J. C., Carney, J. P. & Timlin, J. A. Delivering CRISPR: a review of the challenges and approaches. *Drug Deliv.* **25**, 1234–1257 (2018).
348. Matharu, N. *et al.* CRISPR-mediated activation of a promoter or enhancer rescues obesity caused by haploinsufficiency. *Science* **363**, eaau0629 (2019).
349. Gilbert, L. A. *et al.* CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**, 442–451 (2013).
350. Schultz, D. C., Ayyanathan, K., Negorev, D., Maul, G. G. & Rauscher, F. J., 3rd. SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes Dev.* **16**, 919–932 (2002).
351. Thakore, P. I. *et al.* Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nat. Methods* **12**, 1143–1149 (2015).
352. Gilbert, L. A. *et al.* Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* **159**, 647–661 (2014).
353. Yeo, N. C. *et al.* An enhanced CRISPR repressor for targeted mammalian gene regulation. *Nat. Methods* **15**, 611–616 (2018).
354. Nuñez, J. K. *et al.* Genome-wide programmable transcriptional memory by CRISPR-based epigenome editing. *Cell* **184**, 2503-2519.e17 (2021).
355. Anthon, C., Corsi, G. I. & Gorodkin, J. CRISPRon/off: CRISPR/Cas9 on- and off-target gRNA design. *Bioinformatics* **38**, 5437–5439 (2022).
356. Liu, X. S. *et al.* Editing DNA methylation in the mammalian genome. *Cell* **167**, 233-247.e17 (2016).
357. Tahiliani, M. *et al.* Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**, 930–935 (2009).
358. He, Y.-F. *et al.* Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* **333**, 1303–1307 (2011).
359. Shi, P. & Wu, X. Programmable RNA targeting with CRISPR-Cas13. *RNA Biol.* **21**, 1–9 (2024).
360. Konermann, S. *et al.* Transcriptome engineering with RNA-targeting type VI-D CRISPR effectors. *Cell* **173**, 665-676.e14 (2018).
361. Abudayyeh, O. O. *et al.* RNA targeting with CRISPR-Cas13. *Nature* **550**, 280–284 (2017).
362. Grimm, D. The dose can make the poison: lessons learned from adverse in vivo toxicities caused by RNAi overexpression. *Silence* **2**, 8 (2011).

363. Cox, D. B. T. *et al.* RNA editing with CRISPR-Cas13. *Science* **358**, 1019–1027 (2017).
364. Naldini, L. *et al.* In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science* **272**, 263–267 (1996).
365. Durand, S. & Cimarelli, A. The inside out of lentiviral vectors. *Viruses* **3**, 132–159 (2011).
366. Maetzig, T., Galla, M., Baum, C. & Schambach, A. Gammaretroviral vectors: biology, technology and application. *Viruses* **3**, 677–713 (2011).
367. Wu, X., Li, Y., Crise, B. & Burgess, S. M. Transcription start regions in the human genome are favored targets for MLV integration. *Science* **300**, 1749–1751 (2003).
368. Mitchell, R. S. *et al.* Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. *PLoS Biol.* **2**, E234 (2004).
369. Wang, J.-H., Gessler, D. J., Zhan, W., Gallagher, T. L. & Gao, G. Adeno-associated virus as a delivery vector for gene therapy of human diseases. *Signal Transduct. Target. Ther.* **9**, 78 (2024).
370. Kotterman, M. A. & Schaffer, D. V. Engineering adeno-associated viruses for clinical gene therapy. *Nat. Rev. Genet.* **15**, 445–451 (2014).
371. Ottaviano, G. & Qasim, W. Current landscape of vector safety and genotoxicity after hematopoietic stem or immune cell gene therapy. *Leukemia* **39**, 1325–1333 (2025).
372. Chandler, R. J., Sands, M. S. & Venditti, C. P. Recombinant adeno-associated viral integration and genotoxicity: Insights from animal models. *Hum. Gene Ther.* **28**, 314–322 (2017).
373. Wu, Z., Asokan, A. & Samulski, R. J. Adeno-associated virus serotypes: vector toolkit for human gene therapy. *Mol. Ther.* **14**, 316–327 (2006).
374. Kumar, M., Keller, B., Makalou, N. & Sutton, R. E. Systematic determination of the packaging limit of lentiviral vectors. *Hum. Gene Ther.* **12**, 1893–1905 (2001).
375. Thomas, C. E., Ehrhardt, A. & Kay, M. A. Progress and problems with the use of viral vectors for gene therapy. *Nat. Rev. Genet.* **4**, 346–358 (2003).
376. Nayak, S. & Herzog, R. W. Progress and prospects: immune responses to viral vectors. *Gene Ther.* **17**, 295–304 (2010).
377. Geng, G. *et al.* Viral and non-viral vectors in gene therapy: current state and clinical perspectives. *EBioMedicine* **118**, 105834 (2025).
378. Rehman, U. U., Ghafoor, D., Ullah, A., Ahmad, R. & Hanif, S. Epigenetics regulation during virus-host interaction and their effects on the virus and host cell. *Microb. Pathog.* **182**, 106271 (2023).
379. Ellis, J. Silencing and variegation of gammaretrovirus and lentivirus vectors. *Hum. Gene Ther.* **16**, 1241–1246 (2005).
380. Allshire, R. C. & Madhani, H. D. Ten principles of heterochromatin formation and function. *Nat. Rev. Mol. Cell Biol.* **19**, 229–244 (2018).

381. Sandoval-Villegas, N., Nurieva, W., Amberger, M. & Ivics, Z. Contemporary transposon tools: A review and guide through mechanisms and applications of Sleeping Beauty, piggyBac and Tol2 for genome engineering. *Int. J. Mol. Sci.* **22**, 5084 (2021).
382. Burnight, E. R. *et al.* A hyperactive transposase promotes persistent gene transfer of a piggyBac DNA transposon. *Mol. Ther. Nucleic Acids* **1**, e50 (2012).
383. Moriarity, B. S. & Largaespada, D. A. Sleeping Beauty transposon insertional mutagenesis based mouse models for cancer gene discovery. *Curr. Opin. Genet. Dev.* **30**, 66–72 (2015).
384. Izsvák, Z., Hackett, P. B., Cooper, L. J. N. & Ivics, Z. Translating Sleeping Beauty transposition into cellular therapies: victories and challenges. *Bioessays* **32**, 756–767 (2010).
385. Li, M. A. *et al.* Mobilization of giant piggyBac transposons in the mouse genome. *Nucleic Acids Res.* **39**, e148 (2011).
386. Tschorn, N., van Heuvel, Y. & Stitz, J. Transgene expression and transposition efficiency of two-component Sleeping Beauty transposon vector systems utilizing Plasmid or mRNA encoding the transposase. *Mol. Biotechnol.* **65**, 1327–1335 (2023).
387. Reznik, B. *et al.* Heterogeneity of transposon expression and activation of the repressive network in human fetal germ cells. *Development* **146**, dev171157 (2019).
388. He, J. *et al.* Identifying transposable element expression dynamics and heterogeneity during development at the single-cell level with a processing pipeline scTE. *Nat. Commun.* **12**, 1456 (2021).
389. Durrant, M. G. *et al.* Systematic discovery of recombinases for efficient integration of large DNA sequences into the human genome. *Nat. Biotechnol.* https://doi.org/10.1038/s41587-022-01494-w (2022) doi:10.1038/s41587-022-01494-w.
390. Cerneckis, J., Cai, H. & Shi, Y. Induced pluripotent stem cells (iPSCs): molecular mechanisms of induction and applications. *Signal Transduct. Target. Ther.* **9**, 112 (2024).
391. Toofan, P., Singh, M., Brooks, A. & McLuckie, K. Non-clinical safety considerations on genome editing using the CRISPR/Cas system. *Genes Dis.* **13**, 101785 (2026).
392. Singh, A. *et al.* A high efficiency precision genome editing method with CRISPR in iPSCs. *Sci. Rep.* **14**, 9933 (2024).
393. Chaterji, S., Ahn, E. H. & Kim, D.-H. CRISPR genome engineering for human pluripotent stem cell research. *Theranostics* **7**, 4445–4469 (2017).
394. Müller-Kuller, U. *et al.* A minimal ubiquitous chromatin opening element (UCOE) effectively prevents silencing of juxtaposed heterologous promoters by epigenetic remodeling in multipotent and pluripotent stem cells. *Nucleic Acids Res.* **43**, 1577–1592 (2015).

395. Pfaff, N. *et al.* A ubiquitous chromatin opening element prevents transgene silencing in pluripotent stem cells and their differentiated progeny. *Stem Cells* **31**, 488–499 (2013).
396. Narsinh, K. H. *et al.* Single cell transcriptional profiling reveals heterogeneity of human induced pluripotent stem cells. *J. Clin. Invest.* **121**, 1217–1221 (2011).
397. Messmer, T. *et al.* Transcriptional heterogeneity in naive and primed human pluripotent stem cells at single-cell resolution. *Cell Rep.* **26**, 815-824.e4 (2019).
398. Reed, X. *et al.* Transcriptional signatures in iPSC-derived neurons are reproducible across labs when differentiation protocols are closely matched. *Stem Cell Res.* **56**, 102558 (2021).
399. Pellegrini, S. *et al.* Transcriptional dynamics of induced pluripotent stem cell differentiation into β cells reveals full endodermal commitment and homology with human islets. *Cytotherapy* **23**, 311–319 (2021).
400. Ahmed, A. *et al.* Human genome Safe Harbor sites: A comprehensive review of criteria, discovery, features, and applications. *Cells* **15**, 81 (2026).
401. Shrestha, D. *et al.* Genomics and epigenetics guided identification of tissue-specific genomic safe harbors. *Genome Biol.* **23**, 199 (2022).
402. Aznauryan, E. *et al.* Discovery and validation of human genomic safe harbor sites for gene and cell therapies. *Cell Reports Methods* vol. 2 100154 Preprint at https://doi.org/10.1016/j.crmeth.2021.100154 (2022).
403. Sadelain, M., Papapetrou, E. P. & Bushman, F. D. Safe harbours for the integration of new DNA in the human genome. *Nat. Rev. Cancer* **12**, 51–58 (2011).
404. Papapetrou, E. P. *et al.* Genomic safe harbors permit high β-globin transgene expression in thalassemia induced pluripotent stem cells. *Nat. Biotechnol.* **29**, 73–78 (2011).
405. Pellenz, S. *et al.* New human chromosomal sites with “Safe Harbor” potential for targeted transgene insertion. *Hum. Gene Ther.* **30**, 814–828 (2019).
406. DeKelver, R. C. *et al.* Functional genomics, proteomics, and regulatory DNA analysis in isogenic settings using zinc finger nuclease-driven transgenesis into a safe harbor locus in the human genome. *Genome Res.* **20**, 1133–1142 (2010).
407. Ordovás, L. Efficient Recombinase-Mediated Cassette Exchange in hPSCs to Study the Hepatocyte Lineage Reveals AAVS1 Locus-Driven Transcriptional Splicing. *Stem Cell Reports* **5**, 918–931 (2015).
408. Qian, K. *et al.* A simple and efficient system for regulating gene expression in human pluripotent stem cells and derivatives. *Stem Cells* **32**, 1230–1238 (2014).
409. Guye, P., Li, Y., Wroblewska, L., Duportet, X. & Weiss, R. Rapid, modular and reliable construction of complex mammalian gene circuits. *Nucleic Acids Res.* **41**, e156 (2013).
410. Mizutani, T., Li, R., Haga, H. & Kawabata, K. Transgene integration into the human AAVS1 locus enhances myosin II-dependent contractile force by reducing

expression of myosin binding subunit 85. *Biochem. Biophys. Res. Commun.* **465**, 270–274 (2015).
411. Li, C. *et al.* Targeted integration and High-Level transgene expression in AAVS1 transgenic mice after in vivo HSC transduction with HDAd5/35++ vectors. *Mol. Ther.* **27**, 2195–2212 (2019).
412. Bhagwan, J. R. *et al.* Variable expression and silencing of CRISPR-Cas9 targeted transgenes identifies the AAVS1 locus as not an entirely safe harbour. *F1000Res.* **8**, 1911 (2019).
413. Klatt, D. *et al.* Differential transgene silencing of myeloid-specific promoters in the AAVS1 Safe Harbor locus of induced pluripotent stem cell-derived myeloid cells. *Hum. Gene Ther.* **31**, 199–210 (2020).
414. Irion, S. *et al.* Identification and targeting of the ROSA26 locus in human embryonic stem cells. *Nat. Biotechnol.* **25**, 1477–1482 (2007).
415. Statello, L., Guo, C.-J., Chen, L.-L. & Huarte, M. Gene regulation by long non-coding RNAs and its biological functions. *Nat. Rev. Mol. Cell Biol.* **22**, 96–118 (2021).
416. Perez, E. E. *et al.* Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nat. Biotechnol.* **26**, 808–816 (2008).
417. Dragic, T. *et al.* HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5. *Nature* **381**, 667–673 (1996).
418. Deng, H. *et al.* Identification of a major co-receptor for primary isolates of HIV-1. *Nature* **381**, 661–666 (1996).
419. MacLeod, D. T. *et al.* Integration of a CD19 CAR into the TCR alpha chain locus streamlines production of allogeneic gene-edited CAR T cells. *Mol. Ther.* **25**, 949–961 (2017).
420. Charo, I. F. & Ransohoff, R. M. The many roles of chemokines and chemokine receptors in inflammation. *N. Engl. J. Med.* **354**, 610–621 (2006).
421. Glass, W. G. *et al.* CCR5 deficiency increases risk of symptomatic West Nile virus infection. *J. Exp. Med.* **203**, 35–40 (2006).
422. Lim, J. K. *et al.* CCR5 deficiency is a risk factor for early clinical manifestations of West Nile virus infection but not for viral transmission. *J. Infect. Dis.* **201**, 178–185 (2010).
423. Leitão, M. de C. *et al.* SHIP identifies genomic safe harbors in eukaryotic organisms using genomic general feature annotation. *Sci. Rep.* **15**, 7193 (2025).
424. Papapetrou, E. P. & Schambach, A. Gene Insertion Into Genomic Safe Harbors for Human Gene Therapy. *Mol. Ther.* **24**, 678–684 (2016).
425. Kotin, R. M. *et al.* Site-specific integration by adeno-associated virus. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 2211–2215 (1990).
426. Samson, M. *et al.* Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature* **382**, 722–725 (1996).

427. Beard, C., Hochedlinger, K., Plath, K., Wutz, A. & Jaenisch, R. Efficient method to generate single-copy transgenic mice by site-specific integration in embryonic stem cells. *Genesis* **44**, 23–28 (2006).
428. Cerbini, T. *et al.* Transcription activator-like effector nuclease (TALEN)-mediated CLYBL targeting enables enhanced transgene expression and one-step generation of dual reporter human induced pluripotent stem cell (iPSC) and neural stem cell (NSC) lines. *PLoS One* **10**, e0116032 (2015).
429. Lalonde, R. L. *et al.* pIGLET: Safe harbor landing sites for reproducible and efficient transgenesis in zebrafish. *Sci. Adv.* **10**, eadn6603 (2024).
430. Odak, A. *et al.* Novel extragenic genomic safe harbors for precise therapeutic T-cell engineering. *Blood* **141**, 2698–2712 (2023).
431. Kimura, Y. *et al.* Human Genomic Safe Harbors and the Suicide Gene-Based Safeguard System for iPSC-Based Cell Therapy. *Stem Cells Transl. Med.* **8**, 627–638 (2019).
432. Lee, E.-S. *et al.* Programmable nuclease-based integration into novel extragenic genomic Safe Harbor identified from Korean population-based CNV analysis. *Mol. Ther. Oncolytics* **14**, 253–265 (2019).
433. Li, S., Zhang, A., Xue, H., Li, D. & Liu, Y. One-Step piggyBac Transposon-Based CRISPR/Cas9 Activation of Multiple Genes. *Molecular Therapy - Nucleic Acids* **8**, 64–76 (2017).
434. Fell, C. W., Schmitt-Ulms, C., Tagliaferri, D. V., Gootenberg, J. S. & Abudayyeh, O. O. Precise kilobase-scale genomic insertions in mammalian cells using PASTE. *Nat. Protoc.* 1–38 (2024).
435. Garberg, P. *et al.* In vitro models for the blood-brain barrier. *Toxicol. In Vitro* **19**, 299–334 (2005).
436. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
437. Hervé Pagès, Marc Carlson, Seth Falcon, Nianhua Li. *AnnotationDbi*. (Bioconductor, 2017). doi:10.18129/B9.BIOC.ANNOTATIONDBI.
438. Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287 (2012).
439. Yu, G. *Enrichplot: Visualization of Functional Enrichment Result*. (2023).
440. Delsing, L. *et al.* Barrier properties and transcriptome expression in human iPSC-derived models of the blood-brain barrier. *Stem Cells* **36**, 1816–1827 (2018).
441. Auerbach, R., Lewis, R., Shinners, B., Kubai, L. & Akhtar, N. Angiogenesis assays: a critical overview. *Clin. Chem.* **49**, 32–40 (2003).
442. Goyal, D. & Goyal, R. Angiogenic transformation in human brain micro endothelial cells: Whole genome DNA methylation and transcriptomic analysis. *Front. Physiol.* **10**, 1502 (2019).

443. Zhang, J. *et al.* GJA1 gene polymorphisms and topographic distribution of cranial MRI lesions in cerebral small vessel disease. *Front. Neurol.* **11**, 583974 (2020).
444. Krolak, T. *et al.* Brain endothelial gap junction coupling enables rapid vasodilation propagation during neurovascular coupling. *Cell* **188**, 5003-5019.e22 (2025).
445. Phillips, C. M., Johnson, A. M., Stamatovic, S. M., Keep, R. F. & Andjelkovic, A. V. 20 kDa isoform of connexin-43 augments spatial reorganization of the brain endothelial junctional complex and lesion leakage in cerebral cavernous malformation type-3. *Neurobiol. Dis.* **186**, 106277 (2023).
446. Sladojevic, N., Yu, B. & Liao, J. K. Regulator of G-protein signaling 5 maintains brain endothelial cell function in focal cerebral ischemia. *J. Am. Heart Assoc.* **9**, e017533 (2020).
447. Li, C. *et al.* Critical role of cathepsin L/V in regulating endothelial cell senescence. *Biology (Basel)* **12**, 42 (2022).
448. Jia, G., Aroor, A. R., Jia, C. & Sowers, J. R. Endothelial cell senescence in aging-related vascular dysfunction. *Biochim. Biophys. Acta Mol. Basis Dis.* **1865**, 1802–1809 (2019).
449. Palikuqi, B. *et al.* Adaptable haemodynamic endothelial cells for organogenesis and tumorigenesis. *Nature* **585**, 426–432 (2020).
450. Vatine, G. D. *et al.* Modeling psychomotor retardation using iPSCs from MCT8-deficient patients indicates a prominent role for the blood-brain barrier. *Cell Stem Cell* **20**, 831-843.e5 (2017).
451. Linville, R. M. *et al.* Long-term cryopreservation preserves blood-brain barrier phenotype of iPSC-derived brain microvascular endothelial cells and three-dimensional microvessels. *Mol. Pharm.* **17**, 3425–3434 (2020).
452. Lee, C. A. A. *et al.* Modeling and rescue of defective blood–brain barrier function of induced brain microvascular endothelial cells from childhood cerebral adrenoleukodystrophy patients. *Fluids Barriers CNS* **15**, (2018).
453. Sandler, V. M. *et al.* Reprogramming human endothelial cells to haematopoietic cells requires vascular induction. *Nature* **511**, 312–318 (2014).
454. CZI Cell Science Program *et al.* CZ CELLxGENE Discover: a single-cell data platform for scalable exploration, analysis and modeling of aggregated data. *Nucleic Acids Res.* **53**, D886–D900 (2025).
455. Griffon, N., Jeanneteau, F., Prieur, F., Diaz, J. & Sokoloff, P. CLIC6, a member of the intracellular chloride channel family, interacts with dopamine D(2)-like receptors. *Brain Res. Mol. Brain Res.* **117**, 47–57 (2003).
456. Anzalone, A. V. *et al.* Programmable deletion, replacement, integration and inversion of large DNA sequences with twin prime editing. *Nat. Biotechnol.* **40**, 731–740 (2022).

457. Kotin, R. M., Linden, R. M. & Berns, K. I. Characterization of a preferred site on human chromosome 19q for integration of adeno-associated virus DNA by non-homologous recombination. *EMBO J.* **11**, 5071–5078 (1992).
458. McCreath, K. J. *et al.* Production of gene-targeted sheep by nuclear transfer from cultured somatic cells. *Nature* **405**, 1066–1069 (2000).
459. Moolten, F. L. Tumor chemosensitivity conferred by inserted herpes thymidine kinase genes: paradigm for a prospective cancer control strategy. *Cancer Res.* **46**, 5276–5281 (1986).
460. Fillat, C., Carrió, M., Cascante, A. & Sangro, B. Suicide gene therapy mediated by the Herpes Simplex virus thymidine kinase gene/Ganciclovir system: fifteen years of application. *Curr. Gene Ther.* **3**, 13–26 (2003).
461. Rieck, S. *et al.* Forward programming of human induced pluripotent stem cells via the ETS variant transcription factor 2: rapid, reproducible, and cost-effective generation of highly enriched, functional endothelial cells. *Cardiovasc. Res.* **120**, 1472–1484 (2024).
462. Grath, A. & Dai, G. SOX17/ETV2 improves the direct reprogramming of adult fibroblasts to endothelial cells. *Cell Rep. Methods* **4**, 100732 (2024).
463. Fetsko, A. R., Sebo, D. J. & Taylor, M. R. Brain endothelial cells acquire blood-brain barrier properties in the absence of Vegf-dependent CNS angiogenesis. *Dev. Biol.* **494**, 46–59 (2023).
464. Wang, Q. *et al.* Activation of Wnt/β-catenin pathway mitigates blood-brain barrier dysfunction in Alzheimer's disease. *Brain* **145**, 4474–4488 (2022).
465. Gumina, R. J., Kirschbaum, N. E., Piotrowski, K. & Newman, P. J. Characterization of the human platelet/endothelial cell adhesion molecule-1 promoter: Identification of a GATA-2 binding element required for optimal transcriptional activity. *Blood* **89**, 1260–1269 (1997).
466. Yuan, L. *et al.* ETS-related gene (ERG) controls endothelial cell permeability via transcriptional regulation of the claudin 5 (CLDN5) gene. *J. Biol. Chem.* **287**, 6582–6591 (2012).
467. Garry, D. J. Etv2 IS A MASTER REGULATOR OF HEMATOENDOTHELIAL LINEAGES. *Trans. Am. Clin. Climatol. Assoc.* **127**, 212–223 (2016).
468. CZI Single-Cell Biology Program *et al.* CZ CELL×GENE Discover: A single-cell data platform for scalable exploration, analysis and modeling of aggregated data. *bioRxiv* (2023) doi:10.1101/2023.10.30.563174.
469. Kebschull, J. M. & Zador, A. M. Cellular barcoding: lineage tracing, screening and beyond. *Nat. Methods* **15**, 871–879 (2018).
470. Biddy, B. A. *et al.* Single-cell mapping of lineage and identity in direct reprogramming. *Nature* **564**, 219–224 (2018).
471. Li, Y. B. *et al.* Angiogenesis driven extracellular matrix remodeling of 3D bioprinted vascular networks. *Bioprinting* **30**, e00258 (2023).

472. Zhao, S. *et al.* PiggyBac transposon vectors: the tools of the human gene encoding. *Transl. Lung Cancer Res.* **5**, 120–125 (2016).
473. Galvan, D. L. *et al.* Genome-wide mapping of PiggyBac transposon integrations in primary human T cells. *J. Immunother.* **32**, 837–844 (2009).
474. Maeshima, N. & Fernandez, R. C. Recognition of lipid A variants by the TLR4-MD-2 receptor complex. *Front. Cell. Infect. Microbiol.* **3**, 3 (2013).
475. Medzhitov, R., Preston-Hurlburt, P. & Janeway, C. A., Jr. A human homologue of the Drosophila Toll protein signals activation of adaptive immunity. *Nature* **388**, 394–397 (1997).
476. Poltorak, A. *et al.* Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. *Science* **282**, 2085–2088 (1998).
477. Moore, L. J., Pridmore, A. C., Dower, S. K. & Read, R. C. Penicillin enhances the toll-like receptor 2-mediated proinflammatory activity of Streptococcus pneumoniae. *J. Infect. Dis.* **188**, 1040–1048 (2003).
478. da Silva Correia, J. & Ulevitch, R. J. MD-2 and TLR4 N-linked glycosylations are important for a functional lipopolysaccharide receptor. *J. Biol. Chem.* **277**, 1845–1854 (2002).
479. Yücel, G. *et al.* Lipopolysaccharides induced inflammatory responses and electrophysiological dysfunctions in human-induced pluripotent stem cell derived cardiomyocytes. *Sci. Rep.* **7**, 2935 (2017).
480. Grasselli, C. *et al.* Toll-like receptor 4 modulation influences human neural stem cell proliferation and differentiation. *Cell Death Dis.* **9**, 280 (2018).
481. Gomes-Silva, D. *et al.* Tonic 4-1BB costimulation in chimeric antigen receptors impedes T cell survival and is vector-dependent. *Cell Rep.* **21**, 17–26 (2017).
482. Chen, Q., Yao, H., Xu, X., Lu, J. & Yang, H. TCF-1 overexpression enhances CD19.CAR-T cell potency by attenuating activation-induced cell death. *Blood* **146**, 4126–4126 (2025).
483. Meshorer, E. *et al.* Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Dev. Cell* **10**, 105–116 (2006).
484. Gaspar-Maia, A., Alajem, A., Meshorer, E. & Ramalho-Santos, M. Open chromatin in pluripotency and reprogramming. *Nat. Rev. Mol. Cell Biol.* **12**, 36–47 (2011).
485. Bulut-Karslioglu, A. *et al.* The transcriptionally permissive chromatin state of embryonic stem cells is acutely tuned to translational output. *Cell Stem Cell* **22**, 369-383.e8 (2018).
486. Atlasi, Y. & Stunnenberg, H. G. The interplay of epigenetic marks during stem cell differentiation and development. *Nat. Rev. Genet.* **18**, 643–658 (2017).
487. Meissner, A. *et al.* Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* **454**, 766–770 (2008).

488. Okano, M., Bell, D. W., Haber, D. A. & Li, E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**, 247–257 (1999).
489. Epsztejn-Litman, S. *et al.* De novo DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes. *Nat. Struct. Mol. Biol.* **15**, 1176–1183 (2008).
490. Smith, Z. D. & Meissner, A. DNA methylation: roles in mammalian development. *Nat. Rev. Genet.* **14**, 204–220 (2013).
491. Hotta, A. & Ellis, J. Retroviral vector silencing during iPS cell induction: an epigenetic beacon that signals distinct pluripotent states. *J. Cell. Biochem.* **105**, 940–948 (2008).
492. Gifford, C. A. *et al.* Transcriptional and epigenetic dynamics during specification of human embryonic stem cells. *Cell* **153**, 1149–1163 (2013).
493. Tsankov, A. M. *et al.* Transcription factor binding dynamics during human ES cell differentiation. *Nature* **518**, 344–349 (2015).
494. Smukler, S. R., Runciman, S. B., Xu, S. & van der Kooy, D. Embryonic stem cells assume a primitive neural stem cell fate in the absence of extrinsic influences. *J. Cell Biol.* **172**, 79–90 (2006).
495. Muñoz-Sanjuán, I. & Brivanlou, A. H. Neural induction, the default model and beyond. *Nature Reviews Neuroscience* **3**, 271–280 (2002).
496. Loh, K. M. *et al.* Efficient endoderm induction from human pluripotent stem cells by logically directing signals controlling lineage bifurcations. *Cell Stem Cell* **14**, 237–252 (2014).
497. Karbassi, E. *et al.* Targeted CRISPR activation is functional in engineered human pluripotent stem cells but undergoes silencing after differentiation into cardiomyocytes and endothelium. *Cell. Mol. Life Sci.* **81**, 95 (2024).
498. Pawlowski, M. *et al.* Inducible and deterministic forward programming of human pluripotent stem cells into neurons, skeletal myocytes, and oligodendrocytes. *Stem Cell Reports* **8**, 803–812 (2017).
499. Lombardo, A. *et al.* Site-specific integration and tailoring of cassette design for sustainable gene transfer. *Nat. Methods* **8**, 861–869 (2011).
500. González, F. *et al.* An iCRISPR platform for rapid, multiplexable, and inducible genome editing in human pluripotent stem cells. *Cell Stem Cell* **15**, 215–226 (2014).
501. Hong, C. K. Y. *et al.* Massively parallel characterization of insulator activity across the genome. *Nat. Commun.* **15**, 8350 (2024).
502. Sizer, R. E. *et al.* Use of tRNA gene barriers improves stability of transgene expression in CHO cells. *Biotechnol. J.* **19**, e2400196 (2024).
503. Sizer, R. E. & White, R. J. Use of ubiquitous chromatin opening elements (UCOE) as tools to maintain transgene expression in biotechnology. *Comput. Struct. Biotechnol. J.* **21**, 275–283 (2023).

504. Doan, C. C., Le, T. L., Ho, N. Q. C. & Hoang, N. S. Effects of ubiquitous chromatin opening element (UCOE) on recombinant anti-TNFα antibody production and expression stability in CHO-DG44 cells. *Cytotechnology* **74**, 31–49 (2022).
505. Hockemeyer, D. *et al.* Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nat. Biotechnol.* **27**, 851–857 (2009).
506. van Rensburg, R. *et al.* Chromatin structure of two genomic sites for targeted transgene integration in induced pluripotent stem cells and hematopoietic stem cells. *Gene Ther.* **20**, 201–214 (2013).
507. Hyde, S. C. *et al.* CpG-free plasmids confer reduced inflammation and sustained pulmonary gene expression. *Nat. Biotechnol.* **26**, 549–551 (2008).
508. Chen, Z.-Y., He, C.-Y., Ehrhardt, A. & Kay, M. A. Minicircle DNA vectors devoid of bacterial DNA result in persistent and high-level transgene expression in vivo. *Mol. Ther.* **8**, 495–500 (2003).
509. Lienert, F., Lohmueller, J. J., Garg, A. & Silver, P. A. Synthetic biology in mammalian cells: next generation research tools and therapeutics. *Nat. Rev. Mol. Cell Biol.* **15**, 95–107 (2014).
510. Becker, J. S., Nicetto, D. & Zaret, K. S. H3K9me3-dependent heterochromatin: Barrier to cell fate changes. *Trends Genet.* **32**, 29–41 (2016).
511. Long, A. H. *et al.* 4-1BB costimulation ameliorates T cell exhaustion induced by tonic signaling of chimeric antigen receptors. *Nat. Med.* **21**, 581–590 (2015).
512. Eyquem, J. *et al.* Targeting a CAR to the TRAC locus with CRISPR/Cas9 enhances tumour rejection. *Nature* **543**, 113–117 (2017).
513. Chen, C. *et al.* Expanding the frontiers of genome engineering: A comprehensive review of CRISPR-associated transposons. *Biotechnol. Adv.* **78**, 108481 (2025).
514. Abramson, J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024).
515. Xu, X. *et al.* Engineered miniature CRISPR-Cas system for mammalian genome regulation and editing. *Mol. Cell* **81**, 4333-4345.e4 (2021).
516. Ma, S. *et al.* hpCasMINI: An engineered hypercompact CRISPR-Cas12f system with boosted gene editing activity. *Nat. Commun.* **16**, 5001 (2025).
517. Witte, I. P. *et al.* Programmable gene insertion in human cells with a laboratory-evolved CRISPR-associated transposase. *Science* **388**, eadt5199 (2025).
518. Li, F. *et al.* Bridge RNA-guided genetic recombination tools for treating neurodegenerative nucleotide repeat disorders. *Neurosci. Bull.* **41**, 734–736 (2025).
519. Hsu, A. *et al.* Mechanistic machine learning for prediction of prime editing outcomes. *Nat. Biotechnol.* https://doi.org/10.1038/s41587-026-03261-7 (2026) doi:10.1038/s41587-026-03261-7.
520. Hollmann, J. *et al.* Genetic barcoding reveals clonal dominance in iPSC-derived mesenchymal stromal cells. *Stem Cell Res. Ther.* **11**, 105 (2020).

521. Sohn, S. *et al.* Cell barcoding reveals lineage-dependent outcomes in hiPSC cardiac differentiation. *bioRxivorg* (2025) doi:10.64898/2025.12.12.694049.
522. Huang, Y. *et al.* Lipid nanoparticle-mediated CRISPR/Cas9 delivery enables efficient trabecular meshwork gene editing in mice. *J. Control. Release* **389**, 114499 (2026).
523. Wu, F. *et al.* Lipid nanoparticles for delivery of CRISPR gene editing components. *Small Methods* **10**, e2401632 (2026).
524. Wang, C. *et al.* Blood-brain-barrier-crossing lipid nanoparticles for mRNA delivery to the central nervous system. *Nat. Mater.* https://doi.org/10.1038/s41563-024-02114-5 (2025) doi:10.1038/s41563-024-02114-5.

Chapter 7: Software and AI Use Acknowledgements

7.1 Illustration Design and Figure Generation

All illustrations, schematics, and figures presented in this dissertation were designed and generated using BioRender©, in accordance with their academic license agreements.

7.2 AI and Writing Assistance

During the preparation of this thesis, the author used Grammarly Pro © (Version 2.14.7) for proofreading to improve grammar, spelling and sentence fluency.

7.3 Reference Management Software

Bibliography and citation formatting were managed and implemented for the dissertation using Paperpile © (Version 0.10.1).

Chapter 8: Appendix

8.1 Supplementary Figures

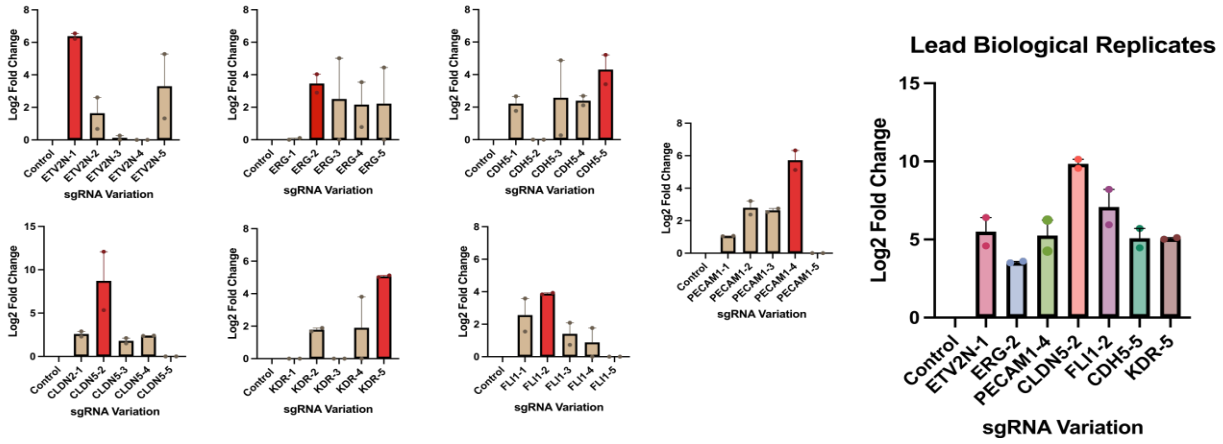


Figure S1: Identifying Lead Candidate sgRNA for Multiplex CRISPR Brain Endothelial Gene Induction.

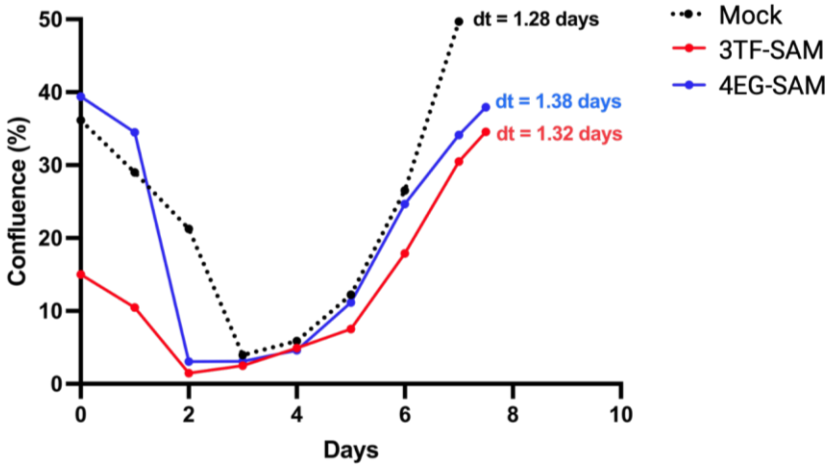


Figure S2: Doubling Time of Multiplexed CRISPRa Transfected hiPSCs During Antibiotic Selection.

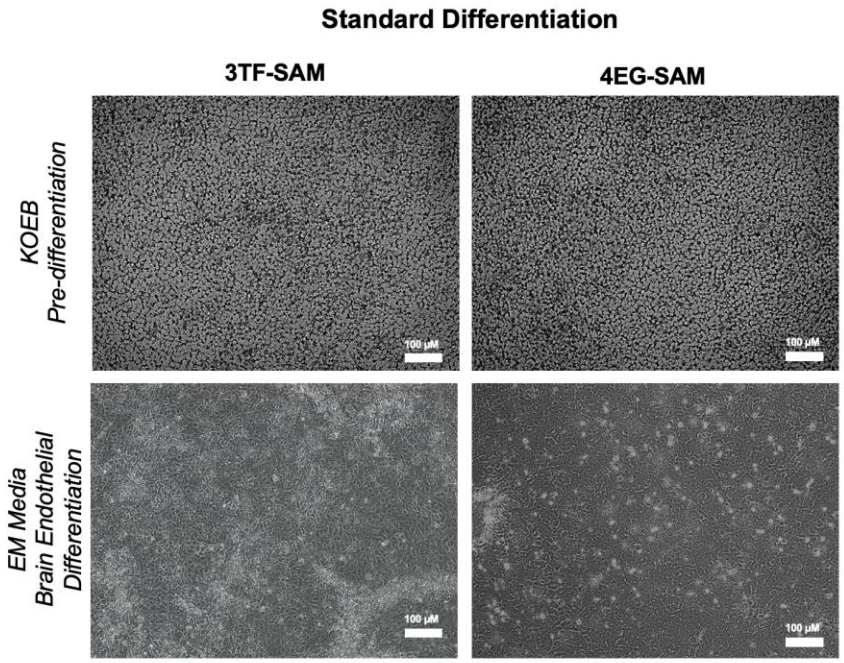


Figure S3: Morphology of 3TF-SAM and 4EG-SAM hiPSCs in Conventional Differentiation Schemes.

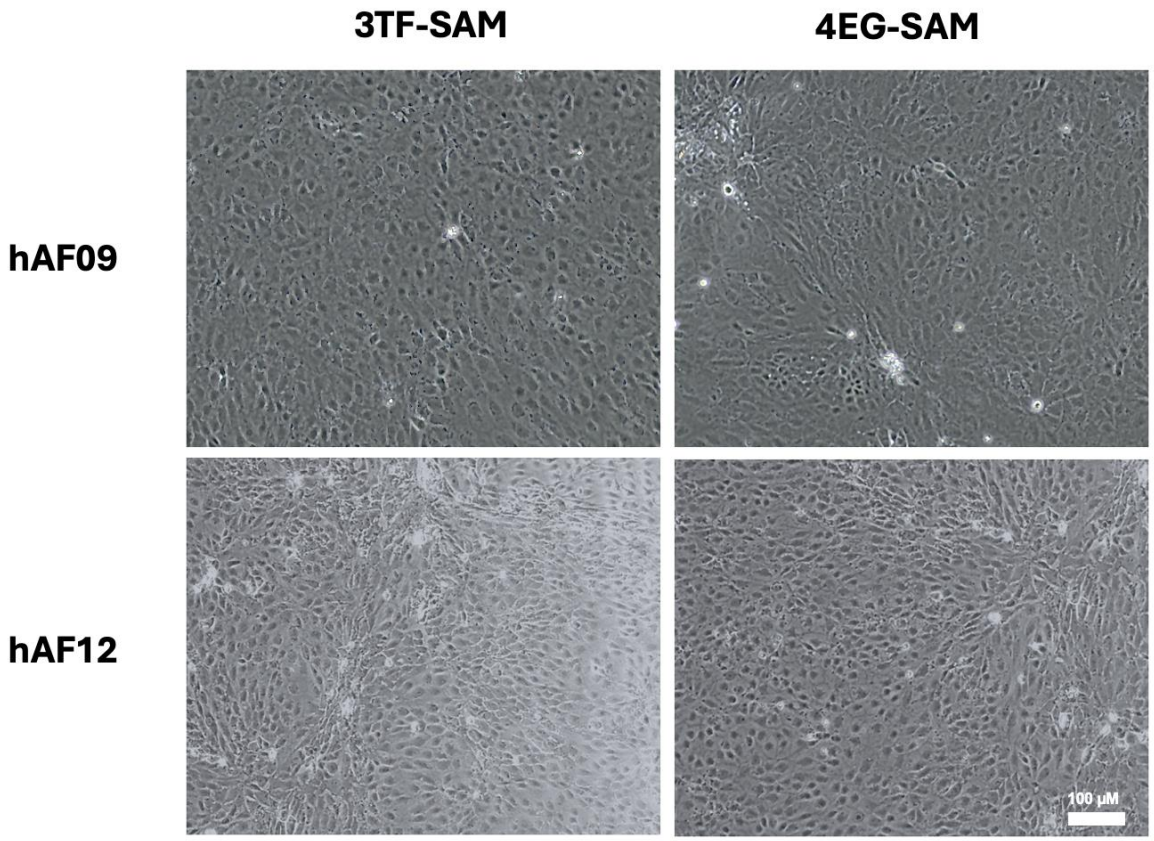


Figure S4: Directed Differentiation in hAF09 and hAF12 hiPSC Cell Lines.

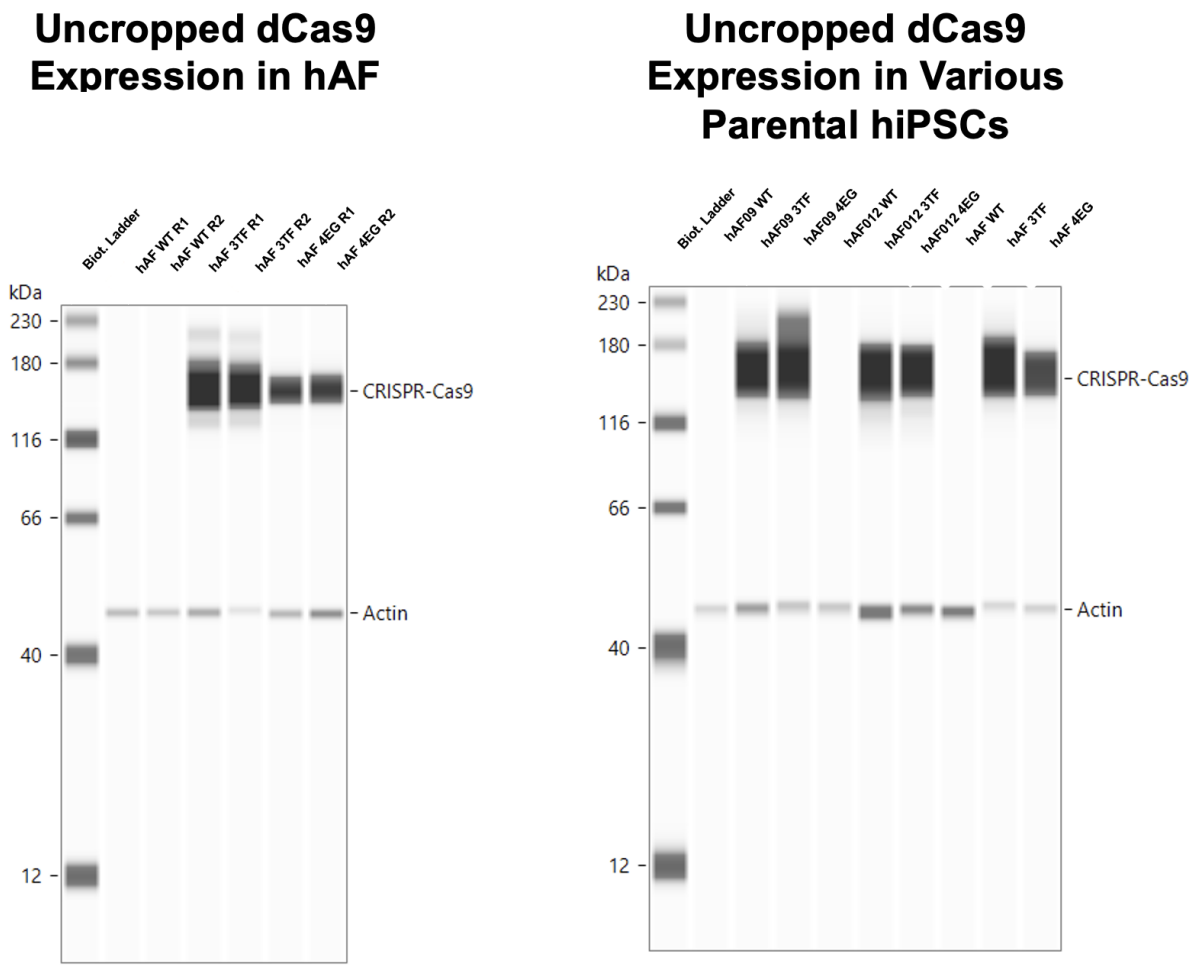


Figure S5: Uncropped Western Blot of CRISPR-Cas9 in Various hiPSCs.

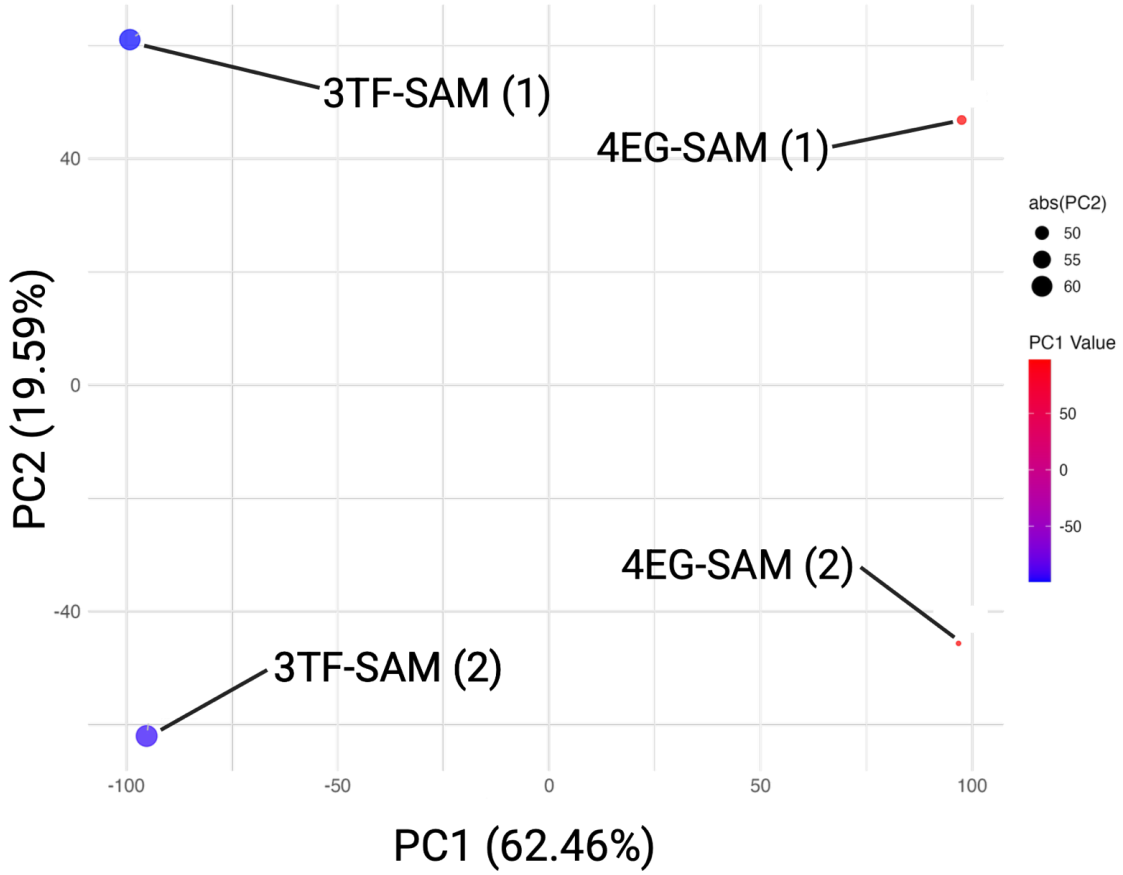


Figure S6: Transcriptomic PCA Analysis Between 3TF-SAM and 4EG-SAM iBECs.

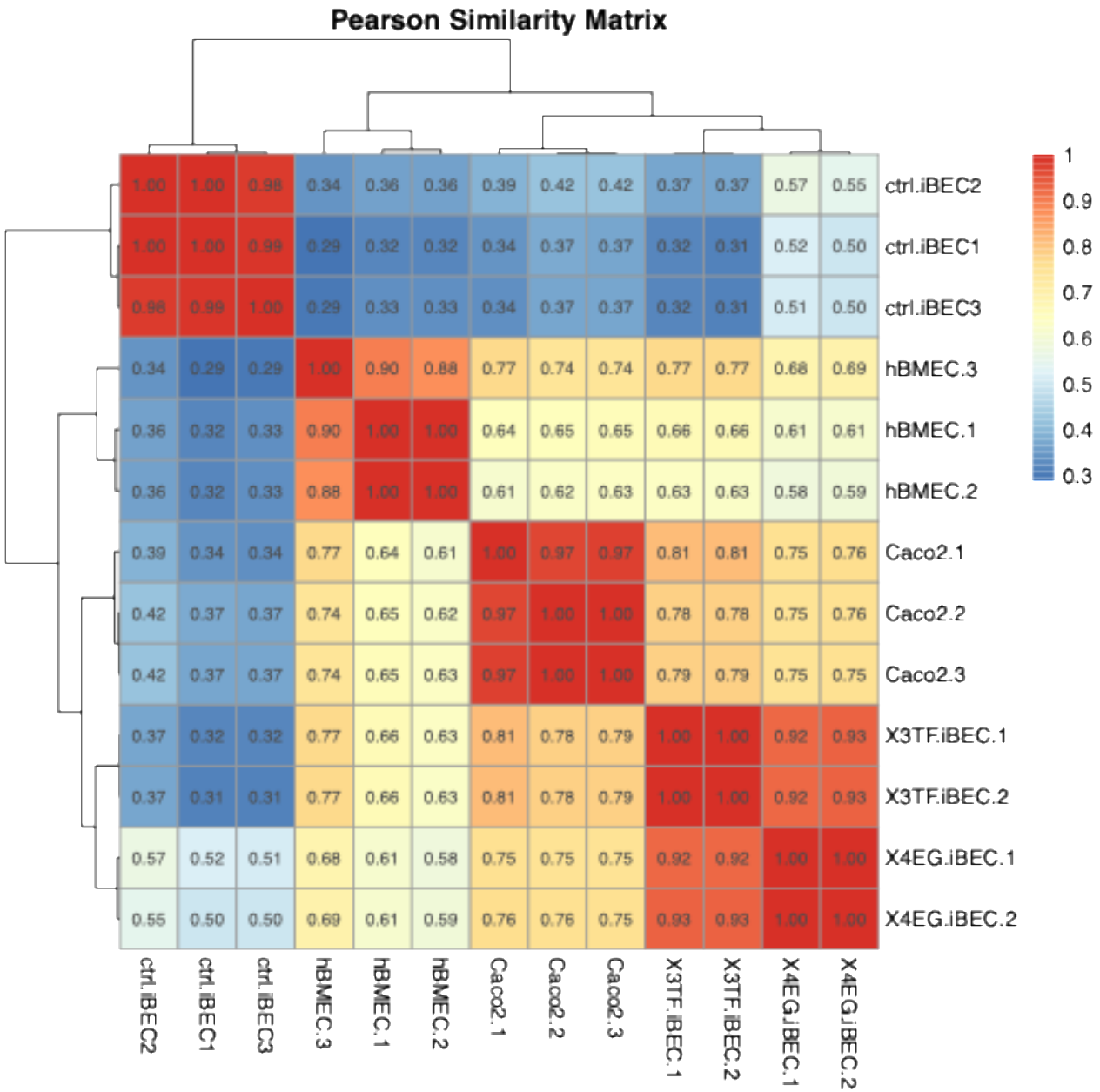


Figure S7: Pearson Correlation Heatmap of Described Cell Lines.

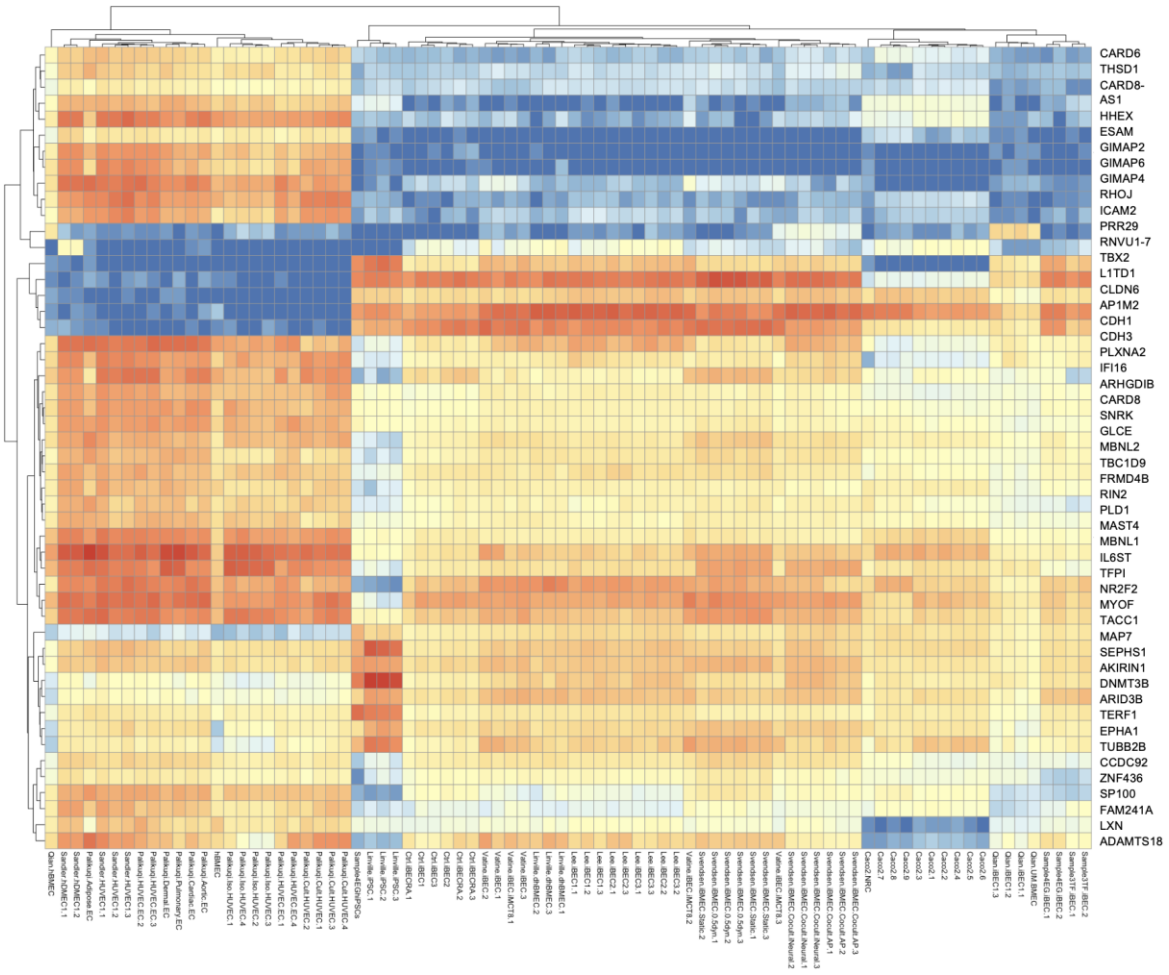


Figure S8: Global Expression Heatmap for Top 20 Differentially Expressed Genes.

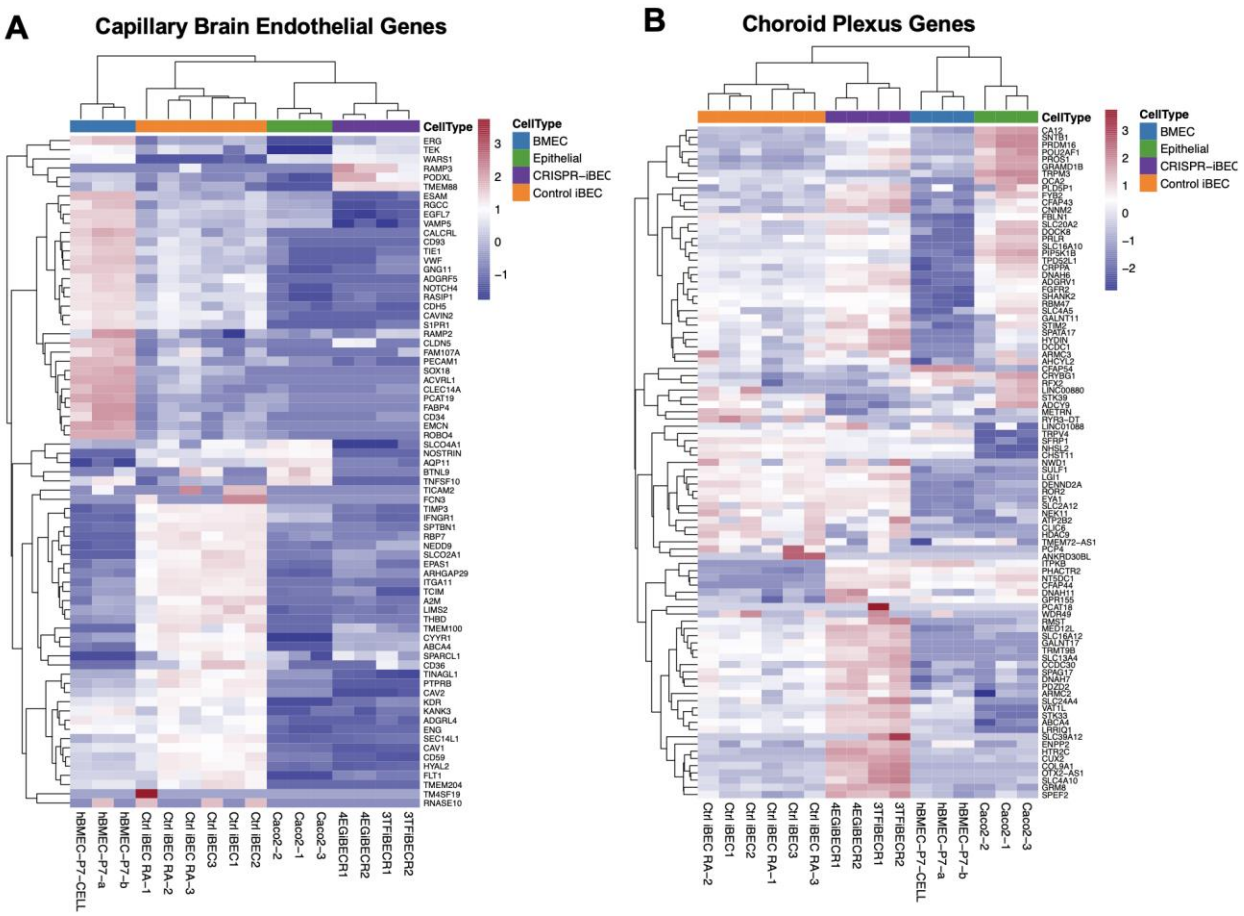


Figure S9: Capillary Brain Endothelial Genes and Choroid Plexus Genes Gene Expression Heatmap.

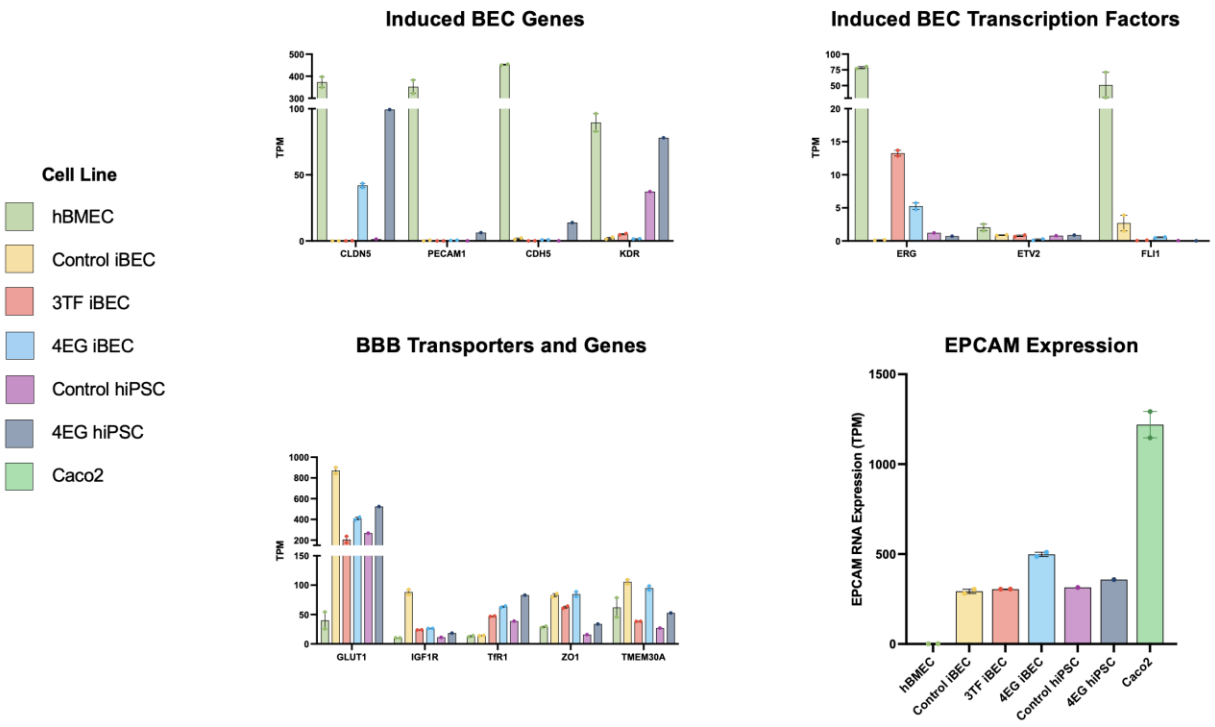


Figure S10: RNA-seq TPM Bar Charts.

8.2 Supplementary Tables

Table S1: Top 20 Most Differentially Expressed Genes in 3TF-SAM iBECs.

List of the top 20 genes ranked by absolute log2 fold change in 3TF-SAM iBECs compared to control and adjusted p-values.

Gene Name	log_2FC	padj
CDC42	-6.2756587	< 2.2e-16
COL12A1	-7.8907305	< 2.2e-16
CTSV	4.73131208	< 2.2e-16
GJA1	5.08771513	< 2.2e-16
KRT7	-7.163648	< 2.2e-16
LUM	-8.7848612	< 2.2e-16
TGM2	-10.060168	< 2.2e-16
TIMP3	-7.8506617	< 2.2e-16
TFPI2	-7.5173176	< 2.2e-16
RGS5	7.52374714	< 2.2e-16
SYT7	5.12779629	< 2.2e-16
CLSTN2	-5.197212	< 2.2e-16
TINAGL1	-7.6455125	< 2.2e-16
TUBB	-3.4483418	< 2.2e-16
H19	-8.69459	< 2.2e-16
ARSI	-6.7431472	< 2.2e-16
GABRP	-5.3016504	2.06E-285
TUBB6	-4.1503542	2.33E-284
FN1	-5.5417892	1.82E-278
ANXA1	-3.9589258	1.45E-272

Table S2: Top 20 Most Differentially Expressed Genes in 4EG-SAM iBECs.

List of the top 20 genes ranked by absolute log_2 fold change in 4EG-SAM iBECs compared to control and adjusted p-values.

Gene Name	log_2FC	padj
CDC42	-6.0320837	< 2.2e-16
COL3A1	-9.4986773	< 2.2e-16
COL12A1	-6.5706148	< 2.2e-16
CTSV	5.46665066	< 2.2e-16
GJA1	4.13614621	< 2.2e-16
KRT7	-5.5302885	< 2.2e-16
NTRK2	-5.9692103	< 2.2e-16
TGM2	-9.4751907	< 2.2e-16
TIMP3	-5.9593921	< 2.2e-16
TFPI2	-6.2216628	< 2.2e-16
RGS5	7.95434299	< 2.2e-16
TINAGL1	-5.7180416	< 2.2e-16
TUBB	-3.4799914	< 2.2e-16
H19	-6.6855521	< 2.2e-16
ARSI	-6.3998317	8.69E-302
SYT7	4.97254238	1.32E-295
TGOLN2	-3.8991555	8.92E-242
CRABP2	-4.7765059	7.29E-229
STOM	3.958206	6.51E-228
CSRP1	-5.6394096	2.95E-221

Table S3: Data Compilation of all Described Studies for Meta-analysis.

Author	GSE
Lee_2018	GSE108012
Linnville_2020	GSE151976
Palikuqi_2020	GSE131039
Qian_2017	GSE97575
Sandler_2014	GSE57662
Svendsen_2019	GSE129290
Vatine_2017	GSE97324
Adam_2022	GSE217504

8.3 Manuscript: Recent Developments in Lipid Nanoparticle-mediated Delivery of Biotherapeutics and Gene Therapy Across the Blood-brain Barrier.

BioDrugs
https://doi.org/10.1007/s40259-026-00782-0

REVIEW ARTICLE

Recent Developments in Lipid Nanoparticle-Mediated Delivery of Biotherapeutics and Gene Therapy Across the Blood–Brain Barrier

Umar Iqbal[1] · Roy W. Hwang[2] · Will J. Costain[1]

Accepted: 8 April 2026
© Crown 2026

Abstract
Peripherally administered therapeutics for neurological indications are challenged with anatomical and physiological barriers that limit their ability to access their site of action in the central nervous system (CNS). This is particularly true for complex therapeutics such as antibodies, immunotherapeutics, and gene therapies. The blood–brain barrier is the specialized structure that functionally regulates the ability of blood constituents to access the CNS. Blood–brain barrier delivery technologies for protein therapeutics have been established in pre-clinical models and are beginning to be verified in clinical studies. Technologies reliant on the transcellular pathway across the blood–brain barrier utilize the receptor-mediated transcytosis mechanism. Research into the use of lipid nanoparticles (LNPs) to deliver complex therapeutics has tremendously expanded in recent years. Lipid nanoparticles represent a compelling alternative to viral vectors for the delivery of various gene therapy modalities, including messenger RNA, small interfering RNA, and antisense oligonucleotides. Functionalization of LNPs with blood–brain barrier-penetrant moieties is being explored as a means to enable CNS delivery of LNP-based therapeutics. The recent innovations and validation of LNP-based delivery systems have hastened the fulfillment of the promise of facile CNS-targeted gene therapies. This review focuses on functional aspects of the blood–brain barrier and how they relate to recent advances in LNP technologies for CNS delivery, as well as their potential impact on gene therapy.

1 Lipid Nanoparticles: A Uniquely Capable Central Nervous System Therapeutic Modality

Solving the challenge of the blood–brain barrier (BBB) for central nervous system (CNS)-targeted therapeutics continues to be an area of intensive research focus. While directly bypassing the BBB using invasive methods (intracranial, intrathecal, or intracerebroventricular [ICV] administration) or temporarily disrupting the BBB (focused ultrasound or osmotic disruption) are viable strategies for a limited number of CNS delivery applications, they are associated with considerable safety and implementation

Key Points

The development and refinement of antibodies that exploit receptor-mediated transcytosis to deliver peripherally administered therapeutics to the central nervous system have enabled new treatment modalities for central nervous system diseases.

The recent clinical validation of the safety and utility of nanoparticle delivery systems has ignited research into the functionalization of lipid nanoparticles for targeted delivery. Similarly, recent developments in clustered regularly interspaced short palindromic repeats (CRISPR) technology have addressed many challenges and presented new opportunities for gene therapy.

The functionalization of lipid nanoparticles with receptor-mediated transcytosis ligands represents a timely solution to the challenge of delivering CRISPR-based gene therapies to the central nervous system.

✉ Will J. Costain
Will.Costain@nrc-cnrc.gc.ca

[1] Human Health Therapeutics Research Centre, National Research Council of Canada, 1200 Montreal Road, Ottawa, ON K1A 0R6, Canada

[2] Department of Cellular Molecular Medicine, Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada

Published online: 30 April 2026

issues [1, 2]. Strategies that focus on crossing the BBB by exploiting endogenous transport mechanisms, such as receptor-mediated transcytosis (RMT), aim to avoid the drawbacks of invasive methods [3–5]. Advances in nanotechnology have led to the development of nanoparticles (NPs) formulated for BBB permeability [4]. Among the various NP formats developed, lipid nanoparticles (LNPs) have undergone extensive clinical validation for safety and efficacy. Indeed, LNPs occupy a unique niche in a vast therapeutic landscape that encompasses small-molecule drugs with simple chemical structures, multi-functional protein biologics, nucleic acids, NPs, viruses, and genetically engineered cells. Each of these therapeutic modalities has its utility and limitations. Herein, we review recent advances in LNP formulation, functionalization, and CNS targeting from the perspective of BBB structure, function, and disease. Moreover, we highlight the versatility of LNP delivery vehicles as a uniquely positioned alternative to viruses in advancing CNS gene therapy.

2 NPs and the BBB

The development of nanocarriers for drug delivery has become a highly active area of research. Concurrently, gene therapy applications, particularly those utilizing adeno-associated virus (AAV), hold great promise and have advanced to clinical trials [6]. As viruses (20–300 nm) and NPs (5–200 nm) exceed the size limit for diffusion across the BBB via the paracellular pathway, they must rely on the endocytic or transcytotic pathways [7]. While the majority of AAV serotypes are not capable of crossing the BBB, certain AAV serotypes exhibit brain tropism and are able to cross the BBB and subsequently transduce neurons [6, 8]. Numerous capsid engineering efforts have been undertaken to further exploit brain tropism and have resulted in further optimizations of brain delivery of AAV-mediated gene therapy [6, 9]. Notably, the recently identified AAVBR1 is an AAV2 variant reported to be highly specific for brain microvascular endothelial cells and may represent a means to locally produce therapeutic proteins or to genetically modulate the BBB to enhance its amenability to NP transport to the CNS [10, 11]. Mechanistically, AAV capsid proteins bind receptors present on brain endothelial cells (BECs) and subsequently undergo RMT in a manner that is distinct from their transduction mechanism [12–14].

Systemically administered NPs are largely excluded from the brain [15], necessitating the engineering of BBB-penetrant properties in nanomaterials to target CNS disorders [7, 16, 17]. While viruses as small as 25 nm can deliver meaningful therapeutic effects, the magnitude of therapeutic effect for NPs is size dependent. Larger NPs (> 100 nm) have greater drug loading capacity, whereas smaller NPs (< 100 nm) exhibit greater cellular uptake and prolonged circulation time [18]. Moreover, very small (< 10 nm) and very large (200–500 nm) NPs are rapidly cleared by the kidney, liver, and spleen [15, 19]. Given that high clearance rates are suboptimal for CNS delivery, optimizing NPs for CNS delivery requires careful attention to a variety of factors, including size, charge, and surface functionalization.

In comparison to viral vectors, LNPs offer several advantages for gene therapy applications. Viruses (adenovirus, AAV and lentivirus) are associated with immunogenicity, payload restrictions and potential genotoxicity due to insertional mutagenesis. Lipid nanoparticle-delivered gene therapy strategies can be designed to avoid these pitfalls while offering the capability to exceed what is possible with a viral vector. Moreover, the cost and complexity of manufacturing therapeutic viruses vastly exceed that of LNPs [20]. Thus, LNP-mediated CNS gene therapy represents a highly promising enabling technology for curative therapeutics targeting neurological disorders.

3 LNPs for CNS Drug Delivery

Lipid nanoparticles represent an innovative advancement in CNS drug delivery, offering new potential to overcome the BBB and deliver diverse therapeutic cargoes directly to the brain. These sophisticated LNP platforms encompass various formulations, including solid lipid nanoparticles, nanostructured lipid carriers, traditional liposomes, and cutting-edge ionizable lipid systems [21]. Each platform offers distinct advantages: solid lipid nanoparticles provide excellent stability and controlled-release profiles, nanostructured lipid carriers offer enhanced drug loading capacity through their imperfect crystalline structure, liposomes enable biomimetic membrane interactions, and ionizable lipid systems facilitate endosomal escape through pH-dependent protonation. The versatility of LNP platforms is high, enabling encapsulation of an extensive array of therapeutic agents that range from small molecules and peptides to complex biologics, including proteins, antisense oligonucleotides (ASOs), small interfering RNA (siRNA), messenger RNA (mRNA), plasmid DNA, and clustered regularly interspaced short palindromic repeats (CRISPR)-Cas gene editing components [22, 23]. The encapsulation process for nucleic acids typically involves ionizable cationic lipids that form electrostatic complexes with the negatively charged payload at acidic pH during NP assembly. Nano-encapsulation protects therapeutic cargoes against in vivo enzymatic degradation and systemic clearance, thereby enabling targeted delivery. The focus of this review will be on systems that incorporate

ionizable cationic lipids and their applicability to targeted CNS delivery.

4 BBB Structure and Function

Peripherally administered neurotherapeutics face a uniquely difficult challenge in accessing their targets in the CNS. This is largely owing to the BBB, which is a highly effective obstacle that uses multiple mechanisms to prevent select blood constituents (solutes and cells) from entering the brain [8]. The BBB is a component of the neurovascular unit (NVU; Fig. 1), which is an anatomical and functional structure that responds to the varying energy demands of the brain by regulating cerebral blood flow [24]. In addition to BECs, the NVU consists of astrocytes, mural cells (i.e., vascular smooth muscle cells and pericytes), microglia, neurons, and extracellular matrix components. The NVU is structured such that the interior of the microvasculature and capillaries is lined with endothelial cells that are attached to the extracellular matrix of the basement membrane (basal lamina) (Fig. 1). The luminal surface of the endothelial cells is covered with a mucosal layer (polysaccharide macromolecules known as the glycocalyx) that acts as a sieve-like barrier to large molecules and NPs [16]. The abluminal surface of the vessels is surrounded by pericytes and astrocytic end feet, and is in contact with neuronal projections. The multi-cellular NVU structure permits neurovascular coupling, which links neural activity and cerebral blood flow, enabling precise spatiotemporal regulation of blood flow that corresponds to localized energy demands [24]. Importantly, astrocytes and pericytes play a well-documented role during the development and maintenance of the BBB [25], and the barrier properties of the BBB rely on the cellular crosstalk between pericytes, astrocytes, and neurons within the NVU [26].

Blood constituents within the lumen of the capillaries are directly exposed to the endothelial cell layer of the NVU. As such, BECs represent the first and primary opportunity to restrict access to the brain. Brain endothelial cells are specialized tight junction (TJ)-forming endothelial cells that effectively restrict paracellular transport of hydrophilic proteins > 500 Da across the BBB [28], which is enhanced by the low rate of pinocytosis of BECs [29, 30]. The intercellular TJs are formed from rows of extensive overlapping occlusions between BECs, with structural integrity being provided by TJ protein complexes [29]. Tight junction complexes are formed by a collection of proteins, including occludins, claudins, junctional adhesion molecules, and ZO-1, with claudins considered to be the principal component in establishing TJs and determining the paracellular barrier between adjacent BECs [31]. Because of the negative charge of the glycocalyx, molecules that are positively charged readily interact with the glycocalyx, whereas negatively charged molecules are repelled [16]. The importance of the glycocalyx in mediating the uptake/transport of LNP nanocarriers was recently demonstrated by Szecskó et al. [32]. These authors showed that altering the surface charge of endothelial cells, by neuraminidase or cationic lipid treatment, significantly reduced uptake of cationic peptide-tagged NPs. Thus, electrostatic interactions must be considered when designing brain-targeted therapeutics.

In addition to the physical TJ barrier, BECs are capable of actively transporting specific solutes across the BBB. This is accomplished by the expression of numerous mono-directional or bi-directional transporter proteins that regulate the transport of specific solutes into or out of the brain [33]. In the case of nutrients, such as glucose, amino acids, nucleosides, and certain neurotransmitters, entry into the brain is accomplished through specific solute transporters such as glucose transporter 1, monocarboxylate transporter 1 (SLC16A1), and large neutral amino acid transporter (LAT1, SLC7A5) [28, 33, 34]. Conversely, efflux transporters, such as the ABC transporters, pump solutes from the brain into the blood, thereby clearing harmful metabolites, as well as playing a role in preventing blood solutes from entering the brain [33–35]. The efflux transporter system is responsible for many small-molecule therapeutics that exhibit lower CNS permeability than would be expected from their lipophilicity [35, 36]. An additional component of the BBB is a metabolic barrier wherein drug-metabolizing enzymes inactivate therapeutics and/or alter their ability to enter the CNS [29]. While the protection the brain receives from potentially harmful blood constituents is necessary for brain function, the BBB also prevents many therapeutics from engaging their CNS targets.

5 Mechanisms of CNS Drug Delivery

Complex therapeutics (e.g., macromolecules, NPs, viruses) targeting CNS pathologies must either be directly administered to the brain or rely on a capacity to be delivered to the brain from the peripheral vasculature [8]. A variety of strategies, including clinically implemented medical procedures such as mannitol-induced osmotic BBB disruption [37], and microbubbles with focused ultrasound [2, 38], have been developed to overcome the BBB and enable delivery of therapeutics into the brain. While BBB disruption mechanisms can effectively enable brain permeation in a limited set of diseases, their physical nature is inherently damaging and is not tenable for diseases requiring repeated long-term dosing [16]. In contrast, transcytotic mechanisms are not associated with BBB disruption or invasive delivery

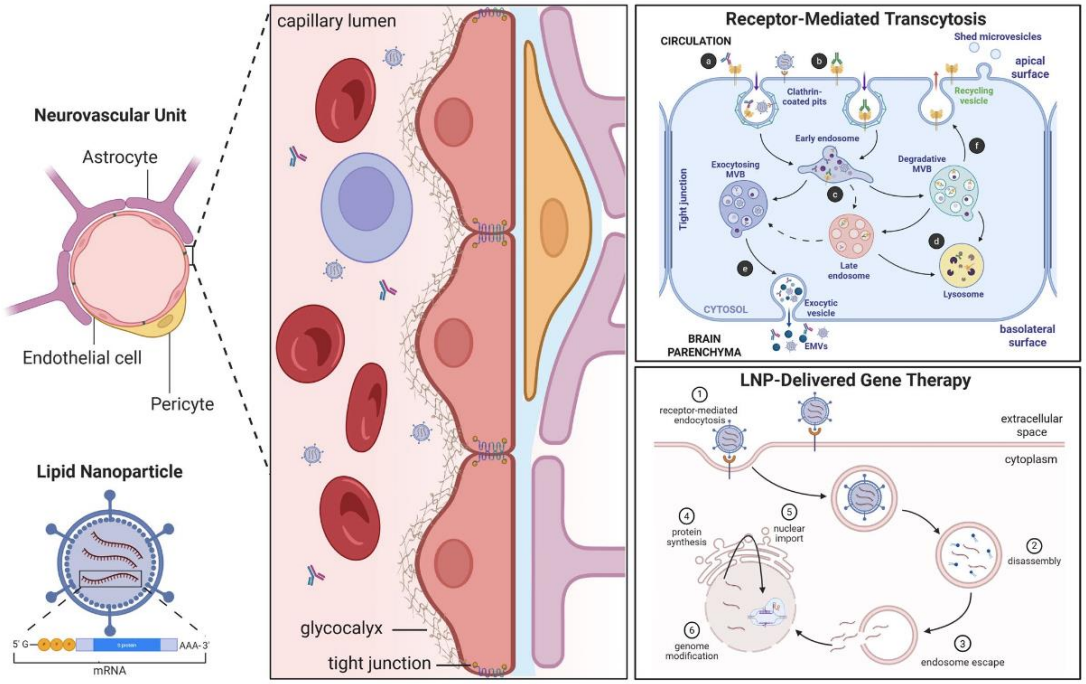


Fig. 1 Schematic of the neurovascular unit structure, the mechanisms of blood–brain barrier transmigration, and lipid nanoparticle (LNP)-RNA delivery in cells. Adjacent brain endothelial cells form tight junctions owing to the presence of adhesion proteins within the plasma membrane. Exposure of therapeutics circulating in the blood to brain endothelial cells can be affected by the presence of the negatively charged glycocalyx. Antibodies and receptor-mediated transcytosis (RMT) ligand-functionalized LNPs can bind receptors, such as the transferrin receptor (TfR), present on the endothelial cell surface and undergo RMT (*top right panel*). **a** Monovalent antibodies or LNPs bind receptors on the luminal side (circulation) of a brain endothelial cell. **b** Bivalent antibody binding to the TfR on a brain endothelial cell. **c** Following clathrin-mediated endocytosis and sorting to the early endosome, antibody ligands or LNPs are further sorted to either exocytosing or degradative multivesicular bodies (MVBs). Dissociation of low-affinity/avidity antibodies can occur, in part due to the reduced endosomal pH (pH sensitivity). **d** High-affinity bivalent interactions that are not pH sensitive are thought to induce TfR dimerization and lead to trafficking toward degradative MVBs and ultimately lysosomal degradation of both the antibody and TfR. **e** Low-affinity, monovalent, or pH-sensitive binding results in trafficking to exocytosing MVBs and ultimately exocytic vesicles where LNPs and antibodies are released, either as free solutes or via exocytosing MVBs (EMVs), into the parenchyma (abluminal side). **f** Non-antibody-bound TfR are trafficked from degradative MVBs to recycling vesicles, where they are redisplayed on the brain endothelial cell surface. Alternatively, partial sorting to late endosomes and subsequently to EMVs may occur [27]. Targeted LNP delivery of gene therapy (*lower right panel*) occurs in cells expressing cell surface receptors that bind the internalization ligand displayed on the LNP. (1) The LNP ligand binds its cell surface receptor and initiates receptor-mediated endocytosis. (2) Endocytosed LNPs are trafficked to endosomes, where the LNP is disassembled into its constituent components. (3) The LNP cargo escapes the endosome and is released into the cytoplasm, where it is imported into the nucleus. (4) The messenger RNA (mRNA) undergoes export to the endoplasmic reticulum for clustered regularly interspaced short palindromic repeats (CRISPR) protein synthesis. (5) The CRISPR protein is then imported into the nucleus, where it can perform genome modifications (6) (created with BioRender.com)

of blood constituents. Transcytosis can be mechanistically subdivided into carrier-mediated transport, RMT, or adsorptive-mediated transcytosis (AMT). The stereospecific transport of smaller cargoes (e.g., glucose, hormones, and fatty acids) is achieved by carrier-mediated transport, while AMT enables the transport of positively charged molecules by virtue of electrostatic interactions with cell surface binding sites [16]. Receptor-mediated transcytosis has proven to be a robust mechanism for delivering a wide variety of therapeutic modalities across the BBB, with multiple receptor systems exploitable for this purpose [8].

Endocytosis is the underlying mechanism for AMT and RMT, and involves internalization of extracellular molecules into intracellular vesicular structures (endosomes). Endosomes are highly heterogeneous, and can be broadly classified as early/sorting endosomes, recycling endosomes, and late endosomes [27, 39, 40]. Classification has been based on size, morphology, cellular localization, and protein

expression. Expression of protein markers is used to differentiate between early/sorting endosomes (Rab5a, Eea1), recycling endosomes (Rab4, Rab11), and late endosomes (Lamp 1/2, M6pr, Rab7a, Rab11a/b) [27, 39]. Nanoscale intraluminal vesicles within the early endosome are formed by the invagination of the limiting membrane of the early endosome, resulting in the creation of multivesicular bodies (MVBs) [41]. These intermediate endosomal structures transition to either degradative MVBs (late endosomes and lysosomes) [42] or exocytosing MVBs, which ultimately form microvesicles that are shed to the extracellular environment [43]. The ability of BECs to shed exocytosing MVBs to their abluminal surface represents the enabling mechanism of RMT (summarized in Fig. 1).

Endocytosis can be classified as clathrin-mediated endocytosis or clathrin-independent endocytosis, including caveolin-dependent endocytosis and pinocytosis. Clathrin-mediated endocytosis (also known as receptor-mediated endocytosis) is a selective process that requires binding to a cell-surface receptor as the first step in the internalization of a specific subset of extracellular molecules. Receptor-mediated endocytosis is a complex controlled process involving numerous factors in the plasma membrane and cytosol [44]. Clathrin-mediated endocytosis and RMT are similar processes, with RMT occurring in a restricted subset of polarized cell types, including neurons, epithelial cells, and endothelial cells [44, 45]. The establishment and maintenance of apicobasal polarity in BECs rely on the function of intercellular tight junctions [46], and BEC function is dynamically regulated by numerous factors, including sheer stress and growth factor signaling [47]. Receptor-mediated transcytosis in barrier-forming cells (epithelial and endothelial cells) is the underlying mechanism responsible for the transport of select macromolecules (such as iron, insulin, and insulin-like growth factors) across the BBB. The initiation of RMT and clathrin-mediated endocytosis is identical, namely the binding of a ligand to a receptor expressed on the luminal (apical) surface of the BEC. Clathrin-mediated endocytosis-mediated internalization of the ligand-receptor complex is followed by sorting into early endosomal vesicles. Multivesicular body formation ensues, with ligands sorted into degradative, recycling, or exocytosing MVBs. Ligands present in exocytosing MVBs are released to the abluminal (basolateral surface, transport to brain parenchyma) or luminal (apical surface, recycling to vessel lumen) sides of the endothelial cell. Notably, most AAV serotypes use glycan moieties (heparan sulfate for AAV2,3,6; N-linked and O-linked sialic acid for AAV1,4,5,6, and terminal galactose for brain trophic AAV9) for initial attachment and engage the AAV receptor KIAA0319L and various co-receptors post-attachment [48–50]. Similarly, brain trophic AAV variants have been found to utilize GPI-linked receptors (lymphocyte antigen 6A, lymphocyte antigen 6C1) or transmembrane proteins (carbonic anhydrase 4, LDL receptor-related protein 6, transferrin receptor (TfR, TFRC), alkaline phosphatase) independently of KIAA0319L to effect brain trophism and BBB transcytosis [14, 50, 51]. Mechanistically, KIAA0319L enables intracellular trafficking (retrograde transport from endosomes to the trans-Golgi network), is required for transduction, and is likely to utilize the above-described mechanism for transcytosis. Similarly, functionalizing NPs to bind known endothelial transcytosis receptors (TfR, insulin receptor, transmembrane protein 30A, insulin-like growth factor 1 receptor [IGF1R], low-density lipoprotein receptor, low-density lipoprotein receptor-related protein 1 [LRP1], CD98, basigin, integrin-beta 1, integral membrane protein 2A, neonatal Fc receptor, P-selectin [SELP]) represents a viable strategy for enabling CNS delivery of therapeutic NPs to the brain parenchyma following peripheral administration [5, 16].

Additionally, there are alternative, less well-understood transcytosis mechanisms, such as AMT [52] and caveolar transcytosis [53], both of which exploit electrostatic interactions between cationic or biomimetic NP surfaces and endothelial membranes to bypass barriers. Adsorptive-mediated transcytosis relies on non-specific electrostatic interactions between positively charged NPs and the negatively charged endothelial glycocalyx, triggering macropinocytosis and subsequent transcellular transport [54]. Caveolar transcytosis utilizes caveolin-1-mediated vesicular transport, which is upregulated in BECs during diseased conditions and avoids the lysosomal degradation pathway [53].

6 BBB and Neuropathology

It has become increasingly evident that structural alterations to the NVU and BBB dysfunction are common observations in numerous neurological diseases and injuries associated with inflammation and neurodegeneration [16, 26]. Blood–brain barrier dysfunction has been observed in animal models and patients with Alzheimer's disease (AD) [55], Parkinson's disease (PD) [56, 57], Huntington's disease (HD) [58–60], amyotrophic lateral sclerosis [61–63], and multiple sclerosis [64, 65], as well as other diseases. Indeed, it is estimated that >40% of brain disorders exhibit dysfunction and/or structural alterations to the BBB [26]. Clinically, BBB dysfunction can be observed using dynamic contrast enhanced-magnetic resonance imaging, by the presence of circulating biomarkers such as claudin-5, occluding, and ZO-1, or by examining the cerebrospinal fluid (CSF)/serum albumin ratio [26].

Post-mortem analysis of brain tissue showing the presence of blood components (fibrinogen, immunoglobulin G, albumin) or pathological immune cell trafficking into the brain parenchyma indicates compromised BBB in neuropathological conditions [26]. Currently, the primary mechanism thought to cause disease-related BBB dysfunction is an altered TJ structure and decreased expression of TJ proteins [25, 26]. Blood–brain barrier dysregulation and disruption under disease conditions represent an underappreciated aspect of neuropathology that may facilitate drug delivery [16], as well as an opportunity for developing BBB-targeted therapeutics that could restore the essential function of the BBB.

7 BBB and Blood–Brain Tumor Barrier Modulation for CNS Gene Delivery

The BBB imposes a major obstacle to CNS drug delivery under physiological conditions, and this limitation persists in glioblastoma (GBM) through the blood–brain tumor barrier, which remains heterogeneously and often incompletely disrupted [66]. While BBB integrity is compromised within hypoxic and angiogenic tumor cores, the blood–brain tumor barrier stays largely intact in low-grade gliomas and at infiltrative margins of high-grade tumors, where glioma cells migrate into adjacent brain tissue without an overt barrier breakdown [66]. This regional barrier continuity contributes to therapeutic resistance and tumor recurrence. Blood–brain tumor barrier properties vary with tumor grade: low-grade gliomas largely preserve BBB-like microvascular architecture, whereas high-grade GBMs exhibit increased vascular surface area, vessel diameter, and permeability that exceed those of normal white matter [66]. Spatial heterogeneity in pericyte coverage, tight junction protein expression (e.g., claudin-5, occludin), and endothelial transcytosis further drives differential permeability and drug accumulation across tumor compartments [67].

To overcome these challenges, LNP-based strategies have been developed that harness ligand-mediated targeting of receptors such as TfR or LRP1, mannose-mediated immune cell uptake, and biomimetic NP coatings to improve penetration and cellular specificity [66, 68]. Complementary to physiological targeting, several physical modulation approaches, including focused ultrasound, optical (light-based) stimulation, and magnetic field-based strategies, enable transient and region-specific increases in BBB and blood–brain tumor barrier permeability [69–72]. Among these, focused ultrasound combined with circulating microbubbles has demonstrated the greatest clinical translatability, achieving a reversible and localized barrier opening that enhances CNS delivery of nucleic acid-loaded LNPs. In preclinical GBM models, this approach increased NP accumulation up to approximately ten-fold in healthy brain regions and approximated six- to seven-fold within tumors while preserving vascular and neurological integrity [69].

Light-based BBB modulation primarily exploits near-infrared photothermal or optogenetic mechanisms to alter endothelial junctions and barrier function with high spatial precision [70, 71]. Near-infrared-absorbing nanomaterials such as gold nanorods generate localized hyperthermia upon irradiation, sufficient to perturb TJ integrity and increase paracellular permeability while limiting widespread tissue damage [70]. Optogenetic platforms expressing light-responsive ion channels or modulating RhoA signaling in endothelial cells have demonstrated reversible control of barrier permeability in in vitro and ex vivo models [71]. Magnetic field-mediated strategies utilize superparamagnetic iron oxide NPs or magnetoelectric composites, where static gradients enhance NP accumulation and alternating fields induce hyperthermia or mechanical stress to transiently open junctions [72]. In LNP-superparamagnetic iron oxide nanoparticles (SPION) hybrids, magnetic guidance has improved intra-tumoral distribution beyond angiogenic cores in rodent GBM models [72]. Emerging data indicate that synchronizing receptor-targeted LNPs with these physical modalities, particularly focused ultrasound or magnetic hyperthermia, extends parenchymal retention and cellular uptake, offering synergistic platforms for precise CNS gene therapy [69, 72].

8 LNP Engineering for CNS Delivery

An emerging area of innovation involves the development of synthetic lipid libraries providing an extensive repertoire of BBB-penetrating lipids (Fig. 2). This approach represents a paradigm shift from empirical formulation development to systematic structure–activity relationship studies that enable the rational design of brain-targeted delivery systems. Rational screening approaches have identified specialized lipids such as MK16 [73], TD5 [74], and NT1 lipidoids [75], which demonstrate dramatically improved brain accumulation compared to conventional US Food and Drug Administration-approved ionizable lipids such as DLin-MC3-DMA, used in first-generation mRNA vaccines such as Onpattro [76]. MK16, featuring a morpholine head group and optimized alkyl chain architecture, achieves up to 7.4-fold higher brain delivery compared with Food and Drug Administration-approved ionizable lipids following intravenous administration [73].

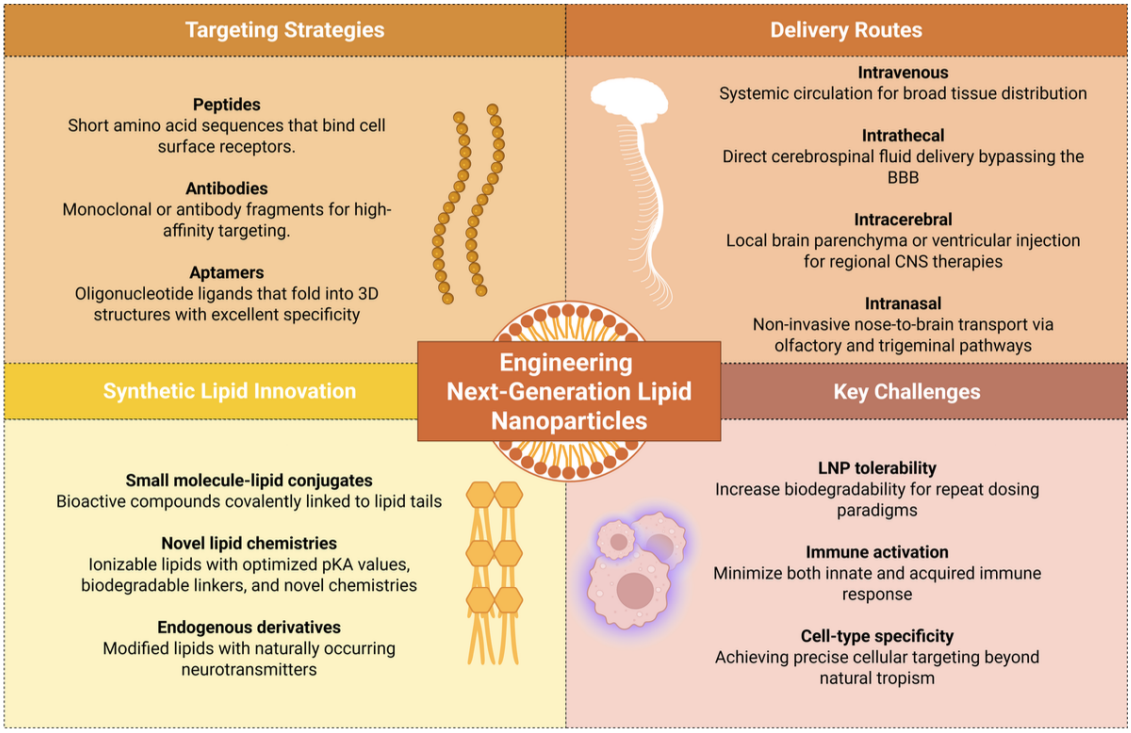


Fig. 2 Considerations for the research and development of next-generation lipid nanoparticles (LNPs) for gene therapy applications (created with BioRender.com). *3D* three-dimensional, *BBB* blood–brain barrier, *CNS* central nervous system

TD5 is a brain-targeting small molecule linked to an amino lipid that resulted in extensive green fluorescent protein expression in 29.6% of neurons and 38.1% of astrocytes across all major brain regions after intrathecal administration [74].

These novel lipids leverage precise chemical modifications to enhance lipid-membrane interactions, promote transcytosis across the BBB, and facilitate transfection in neuronal cells. Additional key design principles include optimization of pKa values to balance circulation stability with endosomal escape (typically targeting pKa values between 6.0 and 7.0), incorporation of biodegradable ester linkages to reduce accumulation toxicity, and fine-tuning of hydrophobic tail structures to optimize membrane fusion kinetics.

High-throughput screening approaches enable the simultaneous evaluation of thousands of lipid candidates, utilizing automated synthesis platforms, robotic handling systems, and high-content imaging to assess BBB penetration, cellular uptake, and toxicity profiles. Machine learning algorithms are increasingly employed to predict optimal lipid structures based on quantitative structure–activity relationships, potentially accelerating the identification of next-generation brain-targeted lipids [77].

The development of neurotransmitter-derived lipidoids represents a particularly elegant engineering strategy wherein lipid molecules mimic structural motifs of dopamine or serotonin, exploiting endogenous vesicular trafficking pathways to traverse the BBB [75]. These biomimetic lipids incorporate neurotransmitter-like head groups that potentially interact with cell surface monoamine transporters or vesicular monoamine transporters, providing a "Trojan horse" approach to BBB crossing. The neurotransmitter 1-lipidoid, which incorporates a dopamine-derived head group, demonstrates preferential accumulation in dopaminergic brain regions and has shown particular promise for delivering therapeutics in Parkinson's disease models [75]. Similarly, linking a synthetic ligand for the 5-HT_3 serotonin receptor, which is known to be transported into the CNS, to the LNP surface was shown to increase brain delivery of encapsulated mRNA [78].

In a notable example, Bian et al. [79] introduced a berberine-inspired ionizable lipid that enhanced nucleic acid

therapeutic delivery across the BBB. By leveraging the structural features of berberine, known to interact with dopamine D_3 receptor, and optimizing lipid architecture, this formulation achieved efficient brain targeting, endosomal escape, and potent gene expression in CNS tissues, further highlighting the potential of transporter-mediated uptake mechanisms. However, because the dopamine D3 receptor is predominantly expressed on neuronal and not endothelial cells, its contribution to direct RMT across the BBB is likely limited, suggesting that the observed brain delivery may primarily result from physicochemical mechanisms rather than true dopamine D3 receptor-driven transcytosis. These findings indicate that in addition to engineering novel lipids, the transporter-mediated mechanism represents a new avenue for BBB-crossing lipid design, potentially exploiting other endogenous transport systems for substrates such as glucose (glucose transporter 1), amino acids (LAT1), choline, and serotonin.

9 Strategies for BBB-Targeted LNP Delivery

9.1 Peptide-Based Targeting

The most extensively developed approach to BBB penetration of LNPs involves RMT, achieved through surface functionalization of LNPs with peptide-based targeting ligands. This mechanism exploits the natural vesicular transport pathways that evolved to facilitate the selective passage of essential nutrients and regulatory molecules across the BBB while maintaining barrier integrity [80]. Peptide-based targeting utilizes specialized sequences, such as T7 or angiopep-2, that specifically bind to receptors highly expressed on BECs, including TfR or LRP1 [81]. As with RMT (Fig. 1), the mechanism of peptide-mediated transcytosis involves several sequential steps: initial binding to surface receptors, triggering receptor clustering and membrane invagination, and subsequent formation of clathrin-coated vesicles that internalize the LNP-peptide complex. These vesicles then traffic across the endothelial cell cytoplasm, avoiding lysosomal degradation, and fuse with the basolateral membrane to release their cargo into the brain parenchyma [32]. Advanced dual-ligand or modular LNP designs can further enhance cell-type selectivity within the brain parenchyma, enabling preferential targeting of specific cell populations such as neurons, astrocytes, or microglia. These sophisticated systems incorporate primary targeting ligands for BBB crossing and the secondary neuronal-targeting ligand, Tet1 peptide, within a single peptide [82].

The use of cell-penetrating peptides (CPPs) has also been explored as a means to enhance CNS targeting with LNPs. In contrast to peptides that engage an RMT-capable receptor, CPPs are permeable to the cell surface membrane and interact with the plasma membrane in a receptor-independent manner [52]. For example, Cao et al. demonstrated that brain delivery of a brain-targeted LNP was enhanced by conjugation of the CPP TAT (derived from the HIV-1 protein TAT) on the surface of the LNP [78]. Similarly, the TAT peptide has been used in combination with transferrin to functionalize liposomes for CNS delivery and enhanced cellular uptake [83]. Notably, CPPs lack tissue specificity and must rely on additional BBB-permeable moieties to effect CNS targeting. However, their ability to enhance cellular uptake may represent a mechanism that can enhance post-BBB cellular uptake.

9.2 Antibody-Based Targeting

Antibody-mediated targeting leverages the exceptional specificity of monoclonal antibodies or antibody fragments against BBB receptors, offering potentially superior selectivity compared with peptide-based approaches. The larger size and complex structure of antibodies provide multiple binding sites and extended serum half-lives, though they also present challenges related to immunogenicity and manufacturing costs [84].

Transferrin receptor antibodies conjugated to LNP surfaces have demonstrated superior BBB transcytosis and brain accumulation in preclinical models, facilitating more effective delivery of therapeutics to brain cells [85]. The transferrin receptor is particularly attractive as a BBB target because of its high expression on BECs, rapid internalization kinetics, and natural recycling pathway that avoids lysosomal degradation. Molecular studies have identified specific anti-TfR antibody clones that bind to distinct epitopes, enabling fine-tuning of transcytosis efficiency and brain penetration patterns [86].

Single-chain variable fragments offer the advantages of reduced size (approximately 25 kDa versus 150 kDa for full antibodies) while maintaining targeting efficacy [87]. The smaller size of single-chain variable fragments enables higher surface density decoration of NPs without steric hindrance, potentially improving avidity through multivalent interactions. Single domain antibodies have also emerged as highly promising targeting moieties for CNS-directed LNPs, offering distinct advantages over traditional antibody formats because of their small size (~ 12–15 kDa), enhanced stability, and reduced immunogenicity [88]. These compact antibodies can be effectively conjugated to LNP surfaces to facilitate BBB transcytosis.

A promising example is IGF1R5, a single domain antibody that targets the IGF1R, which is highly expressed on BECs [89]. IGF1R5 demonstrates sub-nanomolar binding affinity for IGF1R and mediates efficient transcytosis across both in vitro BBB models and in vivo systems, as evidenced by increased brain parenchymal uptake and CSF

exposure. Importantly, IGF1R-targeting single domain antibodies exhibit cross-species reactivity and can be humanized without activating IGF1R signaling pathways, making them well suited for therapeutic applications. Incorporation of an IGF1R5 single domain antibody onto the LNP surface can help enable enhanced CNS penetration of RNA therapeutics.

It is noteworthy that antibody-mediated BBB transcytosis of biomolecules and NPs typically results in a small fraction of the injected dose (ID) being delivered to the CNS. For example, Johnsen et al. [90] reported that < 0.5% of the ID anti-TfR-labeled gold NPs were detected in the brain. However, certain reports have indicated that delivery of up to 4% of the ID was achieved using TfR-targeted NPs. This level of delivery appears to be generally consistent with reports of TfR-targeted antibodies, where 3% of the ID was observed [91]. There are many factors that affect an antibody's capacity to mediate RMT, including affinity, avidity, valency, and pH sensitivity [5]. While efforts to optimize antibody characteristics and prolong the serum half-life can have dramatic effects on the delivery of biotherapeutics to the CNS [92], the inherent capacity of RMT may represent an upper limit to CNS delivery based on this mechanism. If the RMT mechanism is a limiting factor, then the therapeutic capacity of LNPs should be superior to that of a fusion protein. More importantly, it is often observed that peripheral tissues, such as the liver and spleen, accumulate the majority of the ID [90]. Thus, further optimization of RMT-mediated CNS delivery could be achieved by detargeting peripheral tissues, thereby reducing the peripheral sink effect. Recently, Gentry et al. [93] screened lipid libraries to identify an LNP formulation that detargets the liver, thereby altering the tropism toward spleen delivery. This breakthrough represents the possibility for the identification of an LNP formulation that detargets peripheral tissues and becomes increasingly reliant on the antibody–target interactions for tissue uptake. This could be expected to increase the percentage of the ID delivered to the CNS via RMT.

9.3 Aptamer-Based Targeting

Aptamers represent an emerging class of targeting moieties consisting of short single-stranded oligonucleotides (typically 20–100 nucleotides) that fold into unique three-dimensional structures capable of binding specific protein targets with high affinity and specificity [94]. These nucleic acid-based ligands are selected through systematic evolution of ligands by exponential enrichment (SELEX), a process that screens libraries containing up to 10^15 unique sequences to identify optimal binders. These molecules represent a compelling alternative to both peptides and antibodies, maintaining high target specificity (dissociation constants often in the picomolar to nanomolar range) while presenting minimal immunogenicity owing to their nucleic acid composition and a lack of protein epitopes. Unlike antibodies, aptamers can be chemically synthesized with high batch-to-batch consistency for a lower-cost scale-up and can be modified with functional groups for easy conjugation to LNPs.

Aptamer-functionalized NPs have been engineered to target BBB receptors [95]. In this example, a targeting strategy involving LNPs functionalized with anti-platelet-derived growth factor receptor-beta aptamers could both facilitate RMT across the BBB and selectively accumulate in glioblastoma in preclinical models. The small size of aptamers compared to antibodies (molecular weight typically 5–30 kDa vs 150 kDa for antibodies) also enables denser NP surface modification without steric hindrance [96].

10 Alternative Administration Routes to Bypass the BBB

10.1 Intranasal Delivery

Intranasal delivery of LNPs exploits direct neuronal pathways through the olfactory and trigeminal nerves to circumvent the BBB entirely, providing rapid brain access with reduced systemic exposure [97]. Although limited examples exist for ionizable cationic-based LNPs, there is some evidence that cationic liposomal mRNA NPs can be administered intranasally for brain transfection in mice [98]. A single intranasal dose (3 mg/kg) yielded 12-fold higher brain expression in the cortex versus naked mRNA, with dose-dependent luciferase activity distributed across the cortex, striatum, and midbrain and minimal systemic exposure [98]. Although intranasal LNP delivery may enable direct brain access via olfactory and trigeminal pathways while avoiding systemic exposure, it currently faces major limitations, including small dosing volume, mucociliary clearance, variable nasal anatomy, low penetration efficiency with non-viral carriers, potential local toxicity, poor scalability from rodents to humans, and a lack of robust data for ionizable formulations.

10.2 Intracranial Injection

Intracranial injection allows direct delivery of LNPs into the brain parenchyma, completely bypassing the BBB and enabling high local concentrations with minimal systemic exposure. This approach is particularly valuable for treating localized brain pathologies, such as glioma. Studies using MC3-based LNPs demonstrated efficient delivery of mRNA, including Cre recombinase mRNA and Cas9 mRNA/single-guide RNA, to various mouse brain regions, including the striatum and hippocampus. Transfection efficiencies reached approximately 50% in striatal and hippocampal

neurons, astrocytes, and microglia following direct intracranial injection [99]. Generally, intraparenchymal injections achieve high local concentrations, but with limited diffusion. The limited penetration results from the dense extracellular matrix of brain tissue, which restricts NP diffusion, and the lack of convective flow in brain parenchyma.

10.3 Intracerebroventricular Administration

Intracerebroventricular (ICV) injection represents a strategy for delivering LNP-encapsulated RNA therapeutics directly into the CNS, bypassing the BBB and enabling more widespread distribution via CSF [100]. The ventricular system provides a natural distribution network throughout the brain, with CSF flow patterns that facilitate drug dispersion to periventricular and deeper brain structures.

In a pivotal study, LNPs formulated with siRNA targeting CAG repeat expansions were administered via ICV injection in a mouse model of polyglutamine disorders. This approach achieved selective knockdown of mutant mRNA transcripts across key brain regions, including the cortex and striatum, with minimal impact on wild-type alleles. The treatment significantly reduced polyglutamine-expanded protein levels, demonstrating both the specificity and efficacy of ICV-delivered LNPs in silencing disease-causing genes [101]. Although ICV delivery of LNPs allows for direct access to the CNS and broad cerebrospinal fluid distribution, it should be noted that ICV injections can be associated with significant complications, including infectious and non-infectious complications in up to 27% and 33% of patients, respectively [102].

10.4 Intrathecal Delivery

Intrathecal delivery places LNPs into the CSF within the spinal canal, resulting in widespread CNS distribution through CSF circulation. This route offers several advantages over ICV injection, including easier surgical access, reduced risk of brain injury, and potentially more uniform distribution throughout the CNS [103].

Recent research indicates that intrathecal LNP administration can achieve substantial gene transfer, protein expression, and functional gene editing across multiple brain regions [74]. Distribution studies demonstrate that intrathecally injected LNPs reach brain tissue within 30–60 minutes, with peak concentrations occurring 2–4 hours post-injection. This route harnesses CNS fluid dynamics to distribute NPs while circumventing systemic clearance and achieving higher localized CNS exposure compared with intravenous injection [104].

Clinical development efforts utilizing intrathecal LNP delivery for rare CNS diseases and pediatric applications where broad brain exposure is essential are currently underway [104]. Spinal muscular atrophy treatment with ASOs has validated the clinical feasibility of intrathecal nucleic acid delivery, and several gene therapy trials are exploring intrathecal administration of viral vectors [105]. The translation of LNP technology to intrathecal delivery builds upon this established clinical precedent while potentially offering advantages in cargo versatility and repeat dosing.

11 Clinical Translation and Safety Profile

11.1 Therapeutic Efficacy

Modern LNP platforms for CNS delivery have demonstrated promising in vivo performance across multiple validated preclinical and emerging clinical contexts, characterized by effective therapeutic outcomes and favorable safety profiles. The transition from proof-of-concept studies to therapeutic validation represents a critical milestone in the development of CNS nanomedicine. In preclinical disease models including AD, gliomas, and neuroinflammation, LNP formulations carrying mRNA, siRNA, or protein therapeutics have achieved marked improvements in behavioral metrics and significant attenuation of pathological hallmarks, including amyloid-beta aggregation, tumor burden, and inflammatory cytokine levels [106–108].

11.2 Enhanced Tolerability

The tolerability of next-generation LNP platforms has improved substantially through structural and compositional refinements, such as biodegradability. Early LNP formulations often caused dose-limiting toxicities, including complement activation, cytokine release syndrome, hepatotoxicity, and immunogenicity, limiting their therapeutic window and repeat dosing potential [109]. Incorporating biomimetic surface coatings and replacing traditional polyethylene glycol with smart polyethylene glycol alternatives or stealth ligands minimizes complement activation, reduces immunogenicity risks, decreases cytokine release, and attenuates hepatic accumulation and toxicity [110, 111]. These improvements are particularly important for long-term administration scenarios where repeat dosing regimens may be necessary, such as in neurodegenerative and complex genetic disorders. Chronic toxicology studies in non-human primates have demonstrated that optimized LNP formulations can be administered weekly without cumulative toxicity, dose-limiting adverse effects, or significant immune sensitization [112].

12 Clinical Trials Involving LNP Delivery Systems

Lipid nanoparticle-mediated delivery systems have now advanced from conceptual platforms to clinically validated modalities spanning RNA interference, in vivo gene editing, and emerging epigenome-targeting therapies, with several key programs defining the translational landscape [113–118]. The first major milestone was patisiran (Onpattro®; Alnylam), an intravenously administered LNP-formulated siRNA that targets transthyretin (TTR) in hepatocytes for the treatment of hereditary transthyretin-mediated amyloidosis with polyneuropathy, which established the feasibility of repeated systemic LNP dosing in humans [113]. Building on this liver-directed paradigm, in vivo CRISPR-Cas9 gene editing entered the clinic with NTLA-2001 (Intellia/Regeneron), an LNP-encapsulated CRISPR-Cas9 mRNA and guide RNA combination designed for single-dose TTR knockout in hereditary transthyretin-mediated amyloidosis, where first-in-human data demonstrated substantial and durable serum TTR reduction and provided proof-of-concept for permanent gene editing via LNPs [114].

Beyond TTR, a new generation of cardiovascular gene editing candidates use LNPs to deliver base-editing or CRISPR-Cas9 components to hepatocytes to durably reduce atherogenic lipoproteins [115]. Verve 101, an in vivo base-editing therapy targeting PCSK9, is being evaluated for heterozygous familial hypercholesterolemia and high atherosclerotic risk, and early clinical summaries emphasize LNP-mediated liver delivery as a central feature [116, 119]. Additional programs such as CTX310 (ANGPTL3 editing) similarly rely on systemic ionizable LNPs to enable one-time or infrequently dosed interventions for cardiovascular risk reduction, reflecting a broader shift toward LNP-delivered genome editors in common diseases [117]. In parallel, epigenome editing approaches extend the LNP delivery to CRISPR-based transcriptional modulation, illustrating how LNPs can support not only gene knockout or correction but also durable programmable regulation of pathogenic loci in the liver [118].

Clinical trials are also underway investigating LNP delivery for mRNA replacement therapy. ARCT-810 (LUNAR©-OTC) is a systemically (intravenous infusion) delivered LNP-mRNA that aims to treat ornithine transcarbamylase deficiency by supplying an mRNA encoding a functional enzyme to the liver [120]. Similarly, ARCT-032 (LUNAR©-CF) is an LNP-mRNA formulation that is locally administered to the lung, via aerosolization, to deliver functional cystic fibrosis transmembrane conductance regulator mRNA to the lung.

These efforts underscore how incremental advances in LNP design and cargo technology translate directly into improved transfection, safety, and manufacturability in the clinical setting. Together, these approved products and early-stage trials define a rapidly evolving translational landscape in which LNPs function as a modular delivery chassis for siRNA, mRNA, and genome or epigenome editors, and they provide a clinically grounded framework for new LNP designs.

13 Current Challenges and Limitations

Despite significant progress, several challenges persist in LNP-mediated CNS delivery that must be addressed to achieve the full therapeutic potential of this technology. Precise regional and cell-type targeting within the heterogeneous CNS environment remains a challenge, with existing approaches falling short of reliably targeting discrete cell populations such as hippocampal neurons, dopaminergic cells, or microglia with high specificity [121]. The challenge of cellular heterogeneity is compounded by regional differences in BBB properties, vascular density, and extracellular matrix composition [122]. For example, the BBB in the hippocampus exhibits different transporter expression profiles compared with cortical regions, while the BBB in pathological conditions such as AD or multiple sclerosis shows altered permeability and receptor expression patterns [123]. These variations necessitate disease-specific and region-specific optimization of LNP formulations [124].

14 Gene Therapy for the CNS

Central nervous system disorders are frequently caused by monogenic or polygenic mutations that lead to dysfunctional proteins or dysregulated gene expression. Gene therapy offers the potential to directly address the underlying genetic cause of these diseases. This is especially critical where early or timely intervention could prevent irreversible damage. As the symptoms of genetic diseases with CNS pathology, such as lysosomal storage disorders, are typically evident during development, they require early intervention to limit or prevent developmental abnormalities. Moreover, timely intervention in late-onset diseases, such as AD or PD, could prevent lasting damage and enable healthy aging.

The introduction of gene regulatory and editing tools and technologies has enabled researchers to design targeted therapies to treat genetic diseases and conditions. Many tools have been developed for gene editing, including zinc-finger nucleases, transcription activator-like effector nucleases, and CRISPR. While zinc-finger nucleases and transcription activator-like effector nucleases are valid gene editing tools, the comparative ease and versatility of CRISPR have accelerated gene therapy research [125]. Recent clinical successes

Table 1 Central nervous system disorders related to genetic dysregulation and emerging therapeutic strategies

Disease	Gene(s) regulatory pathways involved	Type of dysregulation	Therapeutic interventions	References
Alzheimer's disease	*APP, PSEN1/2, APOE, MAPT*	Mutations, epigenetic changes, misfolded proteins	CRISPRa/i for neuroprotection, siRNA for tau silencing, AAV-based gene therapy	[128–130]
Parkinson's disease	*LRRK2, SNCA, PINK1, PARK7*	Gain/loss-of-function mutations	siRNA/ASOs against *SNCA*, CRISPR knock-in/out, base editing	[131–133]
Huntington's disease	*HTT* (CAG trinucleotide repeat)	Toxic gain-of-function	CRISPR/Cas9, knockdown, base editing, CRISPR/CasRx, ASO therapy	[134–137]
Rett syndrome	*MECP2*	Loss-of-function mutation	Base editing, CRISPR gene replacement	[138–140]
Glioblastoma	Not fully described but prognostic markers include: *NOTCH1, MYC, TOP2A, CD44, PTPRC, SSRP1, CDK4*	Oncogenic mutations, transcriptional dysregulation	CRISPRa/i, epigenetic modifiers	[141–144]
Leukodystrophies	*GALC, ARSA, ABCD1*	Enzyme deficiency due to mutations	ASO, CRISPR/Cas9	[145, 146]
Spinal muscular atrophy	*SMN1* deletion, *SMN2* splicing defects	Loss of functional protein	ASO (e.g., nusinersen), base editing	[147–149]
LSDs	Depending on which LSD various targets can be used, GLA, GBA1, and SMPD1	Enzyme deficiencies	CRISPR/Cas9	[150–153]

LSDs lysosomal storage disorders *ASO* antisense oligonucleotide, *CRISPR* clustered regularly interspaced short palindromic repeats, *siRNA* small interfering RNA, *AAV* adeno-associated virus

include the use of CRISPR-Cas9 to correct CEP290 gene mutations that cause childhood blindness in Leber congenital amaurosis (LCA10), and the disruption of the BCL11A enhancer in sickle cell disease [126, 127]. This highlights the transformative potential of CRISPR-Cas gene therapy and precision medicine to treat previously incurable conditions and emphasizes the need to explore implementation strategies for CNS disorders. As described above, the BBB is a significant obstacle that increases the complexity of implementing gene therapy strategies targeting the CNS. With the recent advances in CNS-targeted LNP technology, the feasibility of gene therapies based on CRISPR-Cas systems and RNA therapeutics has increased, with effective treatment modalities for rare neurological and neurodegenerative diseases on the horizon (Table 1).

15 Application of CRISPR Technologies

CRISPR-Cas systems are quickly expanding and evolving technologies for genetic engineering that are continually being adapted for novel applications. CRISPR systems have been developed for diverse applications, including gene knockout, knock-in, sequence editing, transcriptional modulation, and epigenetic editing (Fig. 3). However, gene therapy based on AAV vectors is limited by several factors, including restricted payload capacity, immunogenicity, inefficient CNS penetration, and patient reluctance. In contrast, LNPs are a promising alternative to deliver gene therapies because of their higher nucleic acid and/or protein payload capacity, which enables the use of larger next-generation CRISPR technologies (such as prime editors and PASTE).

15.1 CRISPR Knockout, Knock-in, Knockdown, and Base/Prime Editing

The most common, well-studied CRISPR technologies employ Cas9 or Cas12a nucleases [154, 155] to create a targeted double-strand break, resulting in either a gene knockout by non-homologous end joining or gene insertion by homology-directed repair [156]. This enables the irreversible silencing of toxic alleles/genes or the insertion or correction of nucleotides. More recent novel experimental approaches, such as PASTE (programmable addition via site-specific targeting elements), can be used to enhance site-specific integration efficiency without the reliance on homology-directed repair [157].

CRISPR knockout and knock-in technologies are increasingly being used for gene therapy. For example, Rett syndrome is a severe X-linked neurodevelopmental disorder caused by loss-of-function mutations in the *MECP2* gene

[158, 159]. Several studies have investigated CRISPR-based approaches to correct *MECP2* mutations in human cell lines and induced pluripotent stem cell lines [160, 161]. Notably, Cho et al. demonstrated NP-assisted CRISPR-Cas9 delivery into iPSC-NPCs, successfully restoring *MECP2* function by homology-directed repair-mediated insertion of a corrective gene sequence and generating neurons that phenotypically reflect wild-type neurons [139]. In another example, HD is a dominant inherited neurodegenerative disorder caused by CAG trinucleotide repeats in the *HTT* gene that result in gradual CNS dysfunction [162, 163]. CRISPR-Cas9-mediated gene knockout was performed to disrupt mutant *HTT* alleles in mice using AAV delivery [164], which resulted in reduced motor deficits and extended survival.

Another major advancement in the CRISPR-Cas system is the development of base and prime editing technology. Base editing and prime editing enable precise single or short nucleotide edits without introducing double-stranded breaks [165]. Cytosine and adenine editors enable conversion of point mutations (C to T or A to G) while prime editors expand the range of possible edits up to 12 types of substitutions [166, 167]. These technologies can address monogenic neurological diseases where single-nucleotide mutations result in disease phenotypes. For example, alternating hemiplegia of childhood is a rare neurodevelopmental disorder associated with mutations in the *ATP1A3* gene, causing motor and cognitive impairment [168]. A recent study employed prime and base editing technology in both human cells and alternating hemiplegia of childhood mouse (*ATP1A3*) models to reverse its associated pathological phenotypes [169]. They were able to demonstrate that in vivo delivery of a prime editor targeting the *ATP1A3* gene was capable of correcting five different mutations, ameliorating both motor and cognitive defects, and drastically increasing lifespan. This represents a compelling example of a single-dose gene therapeutic that can be used as a template for other rare neurological diseases.

In addition to DNA-targeting systems, CRISPR-Cas13 (also known as CasRx) has been recently adapted for post-transcriptional gene modulation through mRNA knockdown [170]. Unlike Cas9 or Cas12, which cleave DNA, CasRx binds to and degrades RNA molecules, offering an alternative avenue for reversible gene suppression. A recent study utilized CasRx to knock down mutant *HTT* mRNA in HD mice (HD140Q-KI) and HD pig (HD-KI) models to address their associated pathology [135]. Consequently, they were able to ameliorate gliosis in mice and delay neurodegeneration in pigs, underscoring the therapeutic potential of RNA-targeting CRISPR tools for treating these disorders.

15.2 CRISPR Transcriptional Regulation and Epigenome Editing

CRISPR-based transcriptional regulation systems, CRISPR activation and CRISPR interference, use a catalytically dead Cas enzyme (such as dCas9 or dCas12a) fused with transcriptional activators or repressors to modulate gene expression (Fig. 3) [171, 172]. These systems do not cut DNA and allow for reversible modulation of gene expression, which is especially valuable in CNS disorders where dynamic regulation of gene expression represents the safest treatment modality.

In a demonstration of the utility of CRISPR activation in AD, Park et al. employed a dCas9-activator nanocomplex in mice to induce the expression of *ADAM10*, a member of the α-secretase family that cleaves APP, to decrease dementia-causing plaques in the brain [173]. Similarly, Colasante et al. demonstrated the use of CRISPR activation to induce *Snca1* gene expression to rescue inhibitory interneuron excitability and halt seizures in Dravet syndrome [174]. These proof-of-concept studies demonstrate the potential of regulating endogenous gene expression as a strategy to correct pathologies arising from conditions associated with aberrant gene expression.

CRISPR-Cas systems have also been developed to modify epigenetic signatures by manipulating histone modifications, DNA methylation, chromatin accessibility, and three-dimensional genome organization [175]. These epigenome-editing systems typically employ a dCas9 fused to effector domains including DNA methyltransferase (e.g., DNMT3A), demethylases (e.g., TET1, LSD1), or histone acetyltransferases (e.g., p300) [176]. Researchers have developed tools capable of repressing gene expression with DNA methylators such as dCas9-DNMT3A [177] and also inducing gene expression through demethylators such as dCas9-TET1 [178, 179]. Patients with AD exhibit increased amyloid precursor protein (*APP*) expression, which is thought to trigger dementia-causing plaque formation in the brain [180]. Park et al. aimed to repress APP expression by hypermethylating the *APP* promoter in *APP-KI* mice with dCas9-DNMT3A [181]. This intervention reduced overall mRNA expression, decreased neuronal cell death, and attenuated cognitive and behavioral impairments. Similarly, overexpression of α-synuclein in PD is linked to elevated H3K4me3 enrichment at the SNCA promoter [182–184]. Guhathakurta et al. developed dCas9 Suntag-JARID1A to remove H3K4me3 in the α-synuclein promoter and successfully decreased its expression in both SH-SY5Y cells and iPSC-derived dopaminergic neurons [184]. Similarly, Liu et al. demonstrated that targeting the CGG repeat expansion in the FMR1 promoter with dCas9-Tet1 reduced methylation of the promoter and restored FMR1 gene expression in FXS-induced pluripotent stem cell lines [185]. These examples showcase how targeted epigenetic editing can correct disease-associated dysregulation without altering genomic sequences.

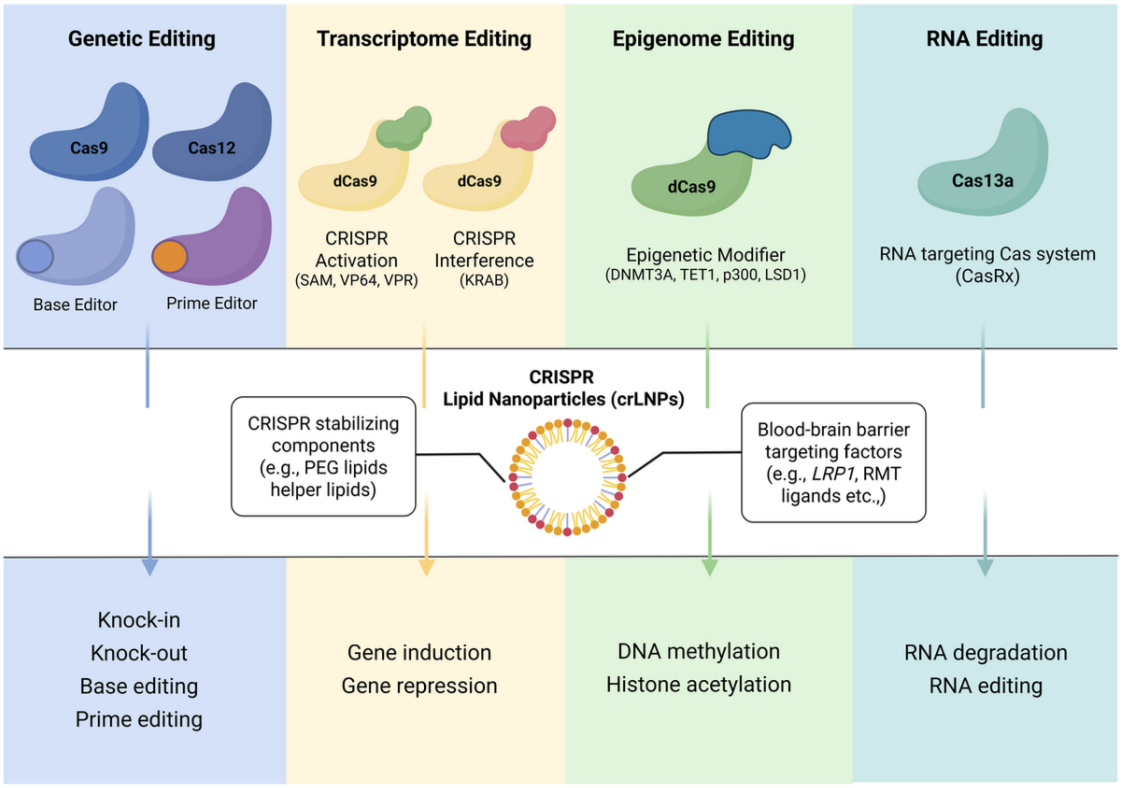


Fig. 3 Clustered regularly interspaced short palindromic repeats (CRISPR)-Cas modalities delivered by brain-targeted CRISPR lipid nanoparticles. DNA targeting editors, including Cas9, Cas12, base editors, and prime editors, enable direct genome modification. Transcriptome modulation is achieved with catalytically inactive dCas9 fused to transcriptional effectors, while epigenome editing employs dCas9 fused to DNMT3A or TET1 to induce DNA methylation or histone modification. CasRx facilitates RNA editing by degrading or modifying transcripts, thereby altering protein expression (created with BioRender.com). *PEG* polyethylene glycol, *RMT* receptor-mediated transcytosis

15.3 LNP Encapsulation of CRISPR Components

Major challenges to the clinical implementation of CRISPR-based therapeutics are the efficient and precise delivery of large molecular payloads across the BBB and nuclear localization following uptake by the target cell. Lipid nanoparticles are well suited to solving these challenges as they are capable of encapsulating large complex CRISPR systems comprised of Cas enzymes, single-guide RNAs, and effector domains. The size and physiochemical properties of ribonucleoprotein complexes add to the challenge of developing LNPs for CRISPR-Cas systems. This is particularly true for Cas enzymes that have been functionalized with additional enzymatic domains, such as the PASTE system, which is a 307-kDa fusion of dCas9, a reverse transcriptase and a serine integrase [157]. Moreover, Cas endonucleases are cationic in nature and pose a barrier to stable encapsulation [186]. Intracellular delivery adds another layer of complexity, requiring that the packaged components escape the endosome and undergo nuclear localization for functional enzymatic activity [125, 187].

Recent developments in CRISPR-loaded LNPs have demonstrated efficient gene editing in cell and animal models [188–191]. These included optimizing ionizable lipids, endosomal release, and incorporation of helper and PEG lipids to mediate distribution and reduce toxicity [188]. Despite these successes, CRISPR-loaded LNPs are designed to target peripheral tissues. A recent study described the development of a BBB-penetrable nanocapsule containing a Cas9-single-guide-RNA complex within a glutathione-polymer sensitive shell with ligands targeting a receptor highly expressed in BECs, such as LRP1 [192]. Zou et al. showed that these CRISPR-loaded LNP effectively downregulated *PLK1* expression in tumor cells and doubled the survival rate in a glioblastoma mouse model [192]. This work demonstrates the potential of BBB-specific RMT ligands and

antibodies to enable the delivery of large complex gene editors based on CRISPR-Cas technology to address neurological and neurodegenerative disorders.

16 Optimizing the Delivery of DNA-Based and RNA-Based Therapies for CNS Disorders

RNA and DNA therapeutics are a modality that can be used for a wide variety of diseases, including neurological and neurodegenerative disorders. Unlike genome-editing tools, these strategies offer transient and or reversible modulation of gene expression [193]. These therapies can be engineered to silence or degrade pathogenic transcripts, upregulate gene expression, encode therapeutic proteins, or modulate disease-associated pathways [193, 194]. This makes nucleic acid-based therapeutics particularly attractive for neurological and neurodegenerative diseases and can be rapidly tailored to both monogenic and polygenic CNS disorders.

16.1 Regulatory RNA Modalities for CNS Therapeutic Intervention

Key classes of regulatory RNA therapeutics include ASOs, siRNAs, and microRNAs (miRNAs), which enable sequence-specific modulation of genes through post-transcriptional mechanisms. Among these, ASOs are short, single-stranded (~12–30 nucleotides) nucleic acid molecules that bind to RNA targets to modulate gene expression by affecting RNA processing or degradation [195]. Antisense oligonucleotides rely on RNAse H-mediated degradation of mRNAs to silence the expression of targeted genes [196, 197]. Several ASO-based drugs have received regulatory approval as gene therapy for neurological disorders such as Duchenne muscular dystrophy, spinal muscular atrophy, hereditary transthyretin-mediated amyloidosis, and amyotrophic lateral sclerosis [198]. Despite their success, the delivery of ASOs to the brain remains a challenge, often relying on invasive intrathecal injections into the CSF with limited tissue distribution [199]. In contrast, siRNAs are double-stranded RNA molecules (~ 20–25 base pairs) that trigger mRNA degradation through the RNA-induced silencing complex [200]. Numerous siRNAs are in the clinical stages of development, such as Mivelsiran and BIIB080 targeting APP mRNA for AD [201]. Other siRNA therapeutics have progressed to phase II trials [202, 203]; however, similar to ASOs, siRNA therapies typically rely on intrathecal delivery. MicroRNAs are small non-coding RNA molecules that play an implicit role in post-transcriptional gene regulation by binding to mRNA targets [204]. Therapeutic strategies using miRNAs include using synthetic miRNA mimics to restore downregulated miRNAs or antagomirs (anti-miRNAs) to inhibit overactive miRNAs [204, 205]. Combinatorial miRNA-based therapies are in development for conditions such as AD with the potential to simultaneously target various components of pathogenic pathways [201]. Despite the distinct molecular mechanisms of ASOs, siRNAs, and miRNAs, they converge functionally as precise tools for post-transcriptional gene regulation, offering an alternative strategy for transient and tunable modulation without permanently altering the genome. Collectively, regulatory RNA modalities show substantial promise, evidenced by the growing pipeline of clinically validated therapeutic interventions. This is mainly owing to their unique properties such as target specificity, reversibility, and dosing flexibility. However, broader translation to the CNS context remains limited by invasive routes of administration, RNA stability and bioavailability at the site of action, as well as challenges in achieving widespread parenchymal distribution across the BBB.

17 mRNA Therapeutics for the CNS: Programmable, Transient, and BBB-Targeted Delivery

Messenger RNA therapeutics have gained traction in recent years owing to the success of the mRNA-based coronavirus disease 2019 vaccine, which demonstrated that synthetic mRNA can be safely delivered to humans [206]. This success, combined with the accompanying LNP development and validation, has invigorated research into mRNA therapeutics for a broad spectrum of diseases. Lipid nanoparticles protect mRNA from degradation, facilitate cellular uptake, and provide a versatile delivery [73, 207] platform through the BBB, making mRNA now an attractive solution for transient protein expression. Moreover, recent efforts in chemical modifications to mRNA, such as nucleoside analogs, optimized untranslated regions, and poly-A tail engineering, enhance RNA stability and reduce innate immune activation, thereby improving their translational efficiency [208, 209]. Collectively, these properties enable a tunable and transient protein production strategy, which can be an advantage when precise temporal control is required. The therapeutic landscape for mRNA therapeutics in the CNS context was broadly explored for immunological and functional protein replacement strategies. For example, LNP-mRNA platforms achieved human trials (NCT04573140) for personalized immunotherapies for high-grade gliomas and adult glioblastomas by stimulating an anti-tumor response [210, 211]. Moreover, preclinical models for PD and AD leveraged mRNA to achieve transient and high expression of neurotrophic factors, such as brain-derived neurotrophic

factor, and anti-inflammatory cytokines [129, 212–214] in the brain. This can be particularly advantageous in the CNS as it can avoid overexpression toxicity and immune activation associated with conventional viral-mediated gene delivery approaches.

Beyond protein replacement therapies, mRNA platforms also support programmability with switch and biosensor technologies leveraging regulatory RNAs [215]. Recent advances in miRNA-responsive RNA switch technology exploit endogenous miRNA expression profiles to selectively regulate mRNA translation. For example, the Hirohide laboratory developed programmable mRNA circuits incorporating miRNA recognition elements flanking a reporter gene to conditionally activate and repress translation based on miRNA recognition [216]. Incorporating miRNA sensors as ON/OFF switches resulted in tunable cell type-specific protein expression, minimized protein leakiness, and suppression of off-target expression. Importantly, this RNA platform was also delivered in vivo, achieving tissue-specific activation of protein production in mouse liver and muscle. In the CNS, where cell-type heterogeneity and specificity are challenges for targeted delivery, solutions such as these pave the way for precise and temporally regulated RNA-based therapeutics targeting specific cell types past the BBB. Integration of miRNA-responsive switches with LNP delivery systems may also further improve translational feasibility by coupling non-viral delivery with endogenous regulatory control.

Despite these advances, mRNA therapeutics are challenged by rapid degradation and limited stability when administered conventionally [217]. Lipid nanoparticle technology has partially addressed these limitations by enhancing bioavailability, reducing immunogenicity, and protecting mRNA payloads from degradation. However, further optimization of LNP-RNA platforms is required to improve their translational performance in clinical settings. Continued innovation in the mRNA and regulatory RNA design alongside LNP engineering expands the scope of CNS-targeted therapies by enabling tunable, transient, and safe interventions, including vaccine and protein replacement strategies for a broad range of neurological disorders.

18 DNA-Based Gene Therapies for the CNS: Sustained Expression and Next-Generation Vectors

While RNA-based therapies emphasize tunability and safety, DNA-based gene delivery remains attractive for applications that require sustained transgene expression in the CNS. Conventional plasmid DNA vectors (3–10 kb) support episomal gene expression lasting days to weeks in dividing cells and weeks to months in non-dividing tissues, with well-established scalability and low manufacturing costs [218, 219]. Expression durability is highly tissue dependent; skeletal muscle can sustain expression for over a year, whereas hepatic expression often declines within days because of promoter inactivation and immune clearance [220]. Plasmid DNA has been evaluated as a treatment for chronic neurodegenerative disorders where long-term therapeutic expression is desired. A prominent example of plasmid DNA-based CNS therapy is the delivery of the protective apolipoprotein E2 (*APOEε2*) isoform to mitigate neurodegeneration associated with AD. *APOEε* genotypes are central to AD risk, where homozygous *APOEε4* carriers develop increased risk for disease onset [221–223]. In contrast, *APOEε2* is protective and can offset the deleterious effects of *APOEε4* [224]. Early preclinical studies using intracerebroventricular lentiviral vector administration in mice demonstrated reduced hippocampal plaque formation [225]. Subsequent studies in larger animal models evaluated intraparenchymal, intracisternal, and intraventricular delivery and revealed that intracisternal administration offered the best balance of efficacy and limited invasiveness [222]. However, these approaches were still limited by sub-optimal distribution, immunogenicity, and invasiveness. Emerging non-viral strategies such as plasmid DNA packaged in liposomes and LNPs tagged with transferrin and penetratin were explored as BBB penetrant alternatives [226, 227]. These approaches achieved successful delivery and transgene expression in mice, highlighting the potential of a less-invasive gene-supplementation strategy in the CNS.

The major limitations faced by plasmid DNA include poor nuclear import of large double-stranded DNA in quiescent cells, CpG-mediated activation of innate immune pathways such as cGAS-STING, and transcriptional silencing caused by bacterial sequences [220]. Although LNPs enhance cellular uptake and endosomal escape, they do not intrinsically overcome nuclear entry barriers and can provoke dose-limiting inflammation. To overcome these barriers, next-generation plasmid technologies are being developed to improve the feasibility of DNA-based gene therapy. These advanced DNA vector formats, including mini-string DNA, closed single-stranded DNA, minicircle DNA, and nanoplasmid DNA, have been developed and are increasingly paired with LNP delivery systems to improve nuclear access, reduce immunogenicity, and prolong expression duration [218, 228–234]. Mini-string DNA and closed single-stranded DNA are substantially smaller vectors that lack bacterial backbone elements, thereby improving nuclear delivery and reducing immunogenicity [231]. Mini-string DNA consists of compact (2-4 kb) backbone-free constructs that exhibit two- to ten-fold higher transgene expression than conventional plasmids across multiple cell types, with minimal innate immune activation [228]. Closed single-stranded

DNA is a closed single-stranded circular format produced enzymatically that has demonstrated high efficiency gene insertion, particularly in hematopoietic stem and progenitor cells, with reduced toxicity compared with double-stranded DNA and greater stability than RNA [229, 230]. Despite their favorable biological profiles, both formats face challenges related to production complexity [228–230].

Minicircle DNA vectors, generated by site-specific recombination to remove the bacterial backbone sequences, typically range from 2-6 kb and exhibit markedly enhanced nuclear import and transcriptional persistence [235]. Minicircles achieve 10- to 1000-fold higher expression than parental plasmids and support sustained expression for weeks in dividing cells and months in non-dividing tissues following an initial decline [236]. Although early production methods were labor intensive and low yielding, advances in producer strain engineering have significantly improved scalability, enabling broader translational use [236]. Lipid nanoparticle-based delivery strategies for minicircle DNA, including targeted formulations, are now under active development.

Nanoplasmid DNA vectors represent an intermediate strategy that preserves plasmid manufacturability while minimizing immunogenicity. These vectors employ ultrashort bacterial backbones ($\leq$ 500 bp) and antibiotic-free RNA-based selection systems, reducing CpG content and transgene silencing [231]. Nanoplasmids maintain high fermentation yields, approaching 2 g/L, and have demonstrated favorable safety and efficacy in clinical studies, with improved expression compared with conventional plasmids [231, 232]. Notably, robust nanoplasmid delivery and expression have been achieved in vitro and in vivo when paired with LNP systems [237, 238].

Across all DNA vector formats, innate immune sensing of cytosolic DNA remains a key limitation for LNP-mediated delivery. Activation of the cGAS-STING pathway drives acute inflammation and restricts expression durability independent of the DNA sequence. Recent studies have shown that incorporation of endogenous anti-inflammatory lipids, such as nitro-oleic acid, into DNA-loaded LNPs can suppress cGAS-STING signaling, enabling prolonged transgene expression exceeding 1 month while preventing severe inflammatory toxicity [220]. This anti-inflammatory co-loading strategy represents a broadly applicable platform for improving the safety and durability of LNP-based DNA therapeutics.

While DNA-based modalities offer the advantage of sustained expression and large cargo capacity, their reliance on nuclear delivery, limited tunability, and widespread CNS distribution presents significant challenges for translational purposes. These constraints have catalyzed increasing interest in RNA-based genome engineering approaches offering complementary advantages in safety, dosing, and flexibility to control protein and gene expression (Table 2). Together, DNA-based and RNA-based modalities illustrate the versatility of nucleic acid-driven interventions for the delivery of therapeutic proteins for CNS diseases. However, their clinical translation remains constrained, and innovation in LNP engineering will be critical to further enhance this as a gene therapy platform.

18.1 Packaging DNA and RNA-Based Therapies in LNPs

Lipid nanoparticles have transformed the prospects of nucleic acid-based therapeutics, overcoming previous limitations associated with delivery and stability in cellular uptake, ushering in an age of novel vaccine therapeutics [217]. Despite these advances, delivery into the CNS remains a bottleneck as current clinical strategies rely on intrathecal and intracerebroventricular administration to bypass the BBB [239]. Thus, a growing interest in engineering BBB-penetrable LNPs capable of delivering diverse RNA and DNA modalities is essential to advance CNS disorder treatments. For example, researchers engineered an LNP with vascular cell adhesion molecule 1 antibodies, a protein expressed in BECs in response to neuroinflammation

Table 2 Comparative design considerations for gene therapy modalities. This table summarizes key biological and delivery-related factors influencing the use of DNA, mRNA, regulatory RNAs, and CRISPR interventions relevant to central nervous system targeting

Design consideration	DNA	mRNA	Regulatory RNAs (ASO, siRNA, miRNA)	CRISPR
Expression duration	Sustained	Transient	Transient	Sustained or long term
Nuclear entry requirement	Required	Not required	Not required	Required
Redosing flexibility	Limited	High	High	Limited
Genomic alteration risk	None	None	None	Present
Safety profile	Moderate	High	High	Context dependent
Cargo capacity	High	Moderate	Low	Moderate to high

ASO antisense oligonucleotide, *CRISPR* clustered regularly interspaced short palindromic repeats, *miRNA* microRNA, *mRNA* messenger RNA, *siRNA* small interfering RNA

[240]. The vascular cell adhesion molecule 1-targeted LNP was delivered in an acute ischemic stroke mouse model and achieved a 100-fold increase in accumulation within cerebral blood vessels when compared with non-targeted LNPs. Moreover, when loaded with interleukin-10 mRNA, researchers observed 62% reduction in infarct size, highlighting the therapeutic potential of precision RNA delivery. Continued optimization of LNP composition and payload design will be critical to efficiently package and safely deliver a diverse repertoire of gene therapies across a wide range of CNS disorders.

19 Future Perspectives

The future of LNP delivery to the brain is positioned to transform CNS therapeutics through advances in gene editing, epigenetic modulation, precision pharmacology, and personalized dosing strategies. The convergence of multiple technological advances suggests that there will be substantial progress in addressing current limitations and expanding therapeutic applications. Next-generation LNPs will likely enable safe, efficient, in vivo delivery of complex gene editors and epigenetic regulators to precisely target neuronal or glial populations. Advanced gene editing approaches such as prime editing, base editing, and epigenome editing offer unprecedented precision in correcting disease-causing mutations or modulating gene expression patterns. Lipid nanoparticle delivery of these molecular tools could enable treatment of monogenic neurological disorders and/or age-related neurodegenerative diseases.

Rational design increasingly employs artificial intelligence-driven lipid and ligand libraries to maximize BBB penetration and cell-type specificity while minimizing off-target effects and immune activation. Machine learning algorithms trained on expanding datasets of LNP performance should help enable the prediction of optimal formulations for specific therapeutic applications. While employing these unsupervised methods will surely accelerate the development of LNP formulations, an understanding of the biological problems remains important to guide design efforts. Evidence that it is possible to optimize against apparently contradictory objectives, such as minimizing peripheral tissue uptake while maximizing endothelial uptake, has been demonstrated in principle with the identification of detargeted LNP formulations [93]. Central nervous system targeting is doubly challenging as therapeutic delivery to endothelial cells is generally not the ultimate objective. To target parenchymal cells with gene therapy, an LNP must remain intact during transcytosis and subsequently fuse with neurons or glia. It can be imagined that multi-functional LNP formulations would be required to detarget peripheral tissues, enable transcytosis of intact LNPs across the BBB, and selectively mediate post-BBB uptake and payload release in neurons or glia.

Emerging data support direct CNS routes, such as intrathecal and intracerebroventricular administration, for broad and durable brain distribution of LNPs, offering promising avenues especially for rare genetic and pediatric disorders requiring single-dose treatments. For systemic administration, safety remains a critical priority. Ongoing refinements in lipid chemistry and surface engineering aim to reduce immunogenicity, complement activation, and hepatic toxicity, enabling the long-term administration necessary for neurodegenerative conditions. The development of biodegradable lipid components, improved stealth coatings, and cell-specific targeting will be essential for enabling repeat dosing regimens required for chronic neurological disorders.

Currently, LNP-mediated delivery systems have been clinically validated to deliver a variety of therapeutic modalities. Strategies targeting peripheral tissues following systemic delivery, such as the liver and lung, have shown great promise in clinical trials. Moreover, pre-clinical studies have demonstrated the utility of LNPs functionalized for BBB permeability. To date, however, clinical trials validating CNS-targeting technologies for LNP therapeutics are lacking. These collective advances suggest that effective treatments for currently intractable neurological conditions may become a reality within the next decade.

Declarations

Funding Open access funding provided by National Research Council Canada library. No funding was received to conduct this study or assist with the preparation of the article.

Conflict of Interest Umar Iqbal and Will J. Costain are employees of the Government of Canada. Roy W. Hwang is a Ph.D. candidate at the University of Ottawa. None has any existing conflict of interest that is directly relevant to the content of this article.

Ethics Approval Not applicable.

Consent to Participate Not applicable.

Consent for Publication Not applicable.

Availability of Data and Material Not applicable.

Code Availability Not applicable.

Authors' Contributions All authors contributed to conceptualizing and writing this manuscript.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes

were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Whelan R, Hargaden GC, Knox AJS. Modulating the blood-brain barrier: a comprehensive review. Pharmaceutics. 2021;13:1980.
2. Poon C, McMahon D, Hynynen K. Noninvasive and targeted delivery of therapeutics to the brain using focused ultrasound. Neuropharmacology. 2017;120:20–37.
3. Ribecco-Lutkiewicz M, Sodja C, Haukenfrers J, Haqqani AS, Ly D, Zachar P, et al. A novel human induced pluripotent stem cell blood-brain barrier model: applicability to study antibody-triggered receptor-mediated transcytosis. Sci Rep. 2018;8:1873.
4. Wu D, Chen Q, Chen X, Han F, Chen Z, Wang Y. The blood-brain barrier: structure, regulation, and drug delivery. Signal Transduct Target Ther. 2023;8:217.
5. Rennie K, Yogi A, Costain WJ. Mechanisms and methods for evaluating drug delivery via transcytosis to the brain. In: Talevi A, editor. CNS drug development and delivery: concepts and applications. Springer; 2024. p. 31–68.
6. Wang J-H, Gessler DJ, Zhan W, Gallagher TL, Gao G. Adeno-associated virus as a delivery vector for gene therapy of human diseases. Signal Transduct Target Ther. 2024;9:78.
7. Senanayake D, Yapa P, Dabare S, Munaweera I. Precision targeting of the CNS: recent progress in brain-directed nanodrug delivery. RSC Adv. 2025;15:25910–28.
8. Stanimirovic DB, Sandhu JK, Costain WJ. Emerging technologies for delivery of biotherapeutics and gene therapy across the blood-brain barrier. BioDrugs. 2018;32:547–59.
9. Luo N, Lin K, Cai Y, Zhang Z, Sui X, Han Z, et al. An engineered adeno-associated virus mediates efficient blood-brain barrier penetration with enhanced neurotropism and reduced hepatotropism. J Control Release. 2025;379:303–15.
10. Rasmussen CLM, Frederiksen SF, Heegaard CW, Thomsen MS, Hede E, Laczek B, et al. Endothelial and neuronal engagement by AAV-BR1 gene therapy alleviates neurological symptoms and lipid deposition in a mouse model of Niemann-Pick type C2. Fluids Barriers CNS. 2025;22:13. https://doi.org/10.1186/s12987-025-00621-4.
11. Körbelin J, Dogbevia G, Michelfelder S, Ridder DA, Hunger A, Wenzel J, et al. A brain microvasculature endothelial cell-specific viral vector with the potential to treat neurovascular and neurological diseases. EMBO Mol Med. 2016;8:609–25. https://doi.org/10.15252/emmm.201506078.
12. Di Pasquale G, Chiorini JA. AAV transcytosis through barrier epithelia and endothelium. Mol Ther. 2006;13:506–16. https://doi.org/10.1016/j.ymthe.2005.11.007.
13. Merkel SF, Andrews AM, Lutton EM, Mu D, Hudry E, Hyman BT, et al. Trafficking of adeno-associated virus vectors across a model of the blood-brain barrier; a comparative study of transcytosis and transduction using primary human brain endothelial cells. J Neurochem. 2017;140:216–30.
14. Moyer TC, Hoffman BA, Chen W, Shah I, Ren X-Q, Knox T, et al. Highly conserved brain vascular receptor ALPL mediates transport of engineered AAV vectors across the blood-brain barrier. Mol Ther. 2025;33:3902–16.
15. De Jong WH, Hagens WI, Krystek P, Burger MC, Sips AJAM, Geertsma RE. Particle size-dependent organ distribution of gold nanoparticles after intravenous administration. Biomaterials. 2008;29:1912–9.
16. Wu JR, Hernandez Y, Miyasaki KF, Kwon EJ. Engineered nanomaterials that exploit blood-brain barrier dysfunction for delivery to the brain. Adv Drug Deliv Rev. 2023;197:114820.
17. Zhou Y, Peng Z, Seven ES, Leblanc RM. Crossing the blood-brain barrier with nanoparticles. J Control Release. 2018;270:290–303.
18. Mehta M, Bui TA, Yang X, Aksoy Y, Goldys EM, Deng W. Lipid-based nanoparticles for drug/gene delivery: an overview of the production techniques and difficulties encountered in their industrial development. ACS Mater Au. 2023;3:600–19.
19. Zhu GH, Gray ABC, Patra HK. Nanomedicine: controlling nanoparticle clearance for translational success. Trends Pharmacol Sci. 2022;43:709–11.
20. Arjunan P, Kathirvelu D, Mahalingam G, Goel AK, Zacharaiah UG, Srivastava A, et al. Lipid-nanoparticle-enabled nucleic acid therapeutics for liver disorders. Acta Pharm Sin B. 2024;14:2885–900.
21. Viegas C, Patrício AB, Prata JM, Nadhman A, Chintamaneni PK, Fonte P. Solid lipid nanoparticles vs. nanostructured lipid carriers: a comparative review. Pharmaceutics. 2023;15:1593.
22. Kaczmarek JC, Kowalski PS, Anderson DG. Advances in the delivery of RNA therapeutics: from concept to clinical reality. Genome Med. 2017;9:60.
23. Wu F, Li N, Xiao Y, Palanki R, Yamagata H, Mitchell MJ, et al. Lipid nanoparticles for delivery of CRISPR gene editing components. Small Methods. 2025;10:e2401632.
24. Iadecola C. The neurovascular unit coming of age: a journey through neurovascular coupling in health and disease. Neuron. 2017;96:17–42.
25. Luissint A-C, Artus C, Glacial F, Ganeshamoorthy K, Couraud P-O. Tight junctions at the Blood Brain Barrier: physiological architecture and disease-associated dysregulation. Fluids Barriers CNS. 2012;9:23. https://doi.org/10.1186/2045-8118-9-23.
26. Lemarchant S, Engelhardt B, Cicchetti F, Bix GJ, Janus A, Godfrin Y, et al. Restoring brain barriers: an innovative approach for treating neurological disorders. Fluids Barriers CNS. 2025;22:72.
27. Haqqani AS, Delaney CE, Brunette E, Baumann E, Farrington GK, Sisk W, et al. Endosomal trafficking regulates receptor-mediated transcytosis of antibodies across the Blood Brain Barrier. J Cereb Blood Flow Metab. 2018;38:727–40. https://doi.org/10.1177/0271678X17740031.
28. Abbott NJ, Patabendige AAK, Dolman DEM, Yusof SR, Begley DJ. Structure and function of the Blood-Brain Barrier. Neurobiol Dis. 2010;37:13–25.
29. Lochhead JJ, Yang J, Ronaldson PT, Davis TP. Structure, function, and regulation of the Blood-Brain Barrier tight junction in Central Nervous System disorders. Front Physiol. 2020;11:914.
30. Reese TS, Karnovsky MJ. Fine structural localization of a Blood-Brain Barrier to exogenous peroxidase. J Cell Biol. 1967;34:207–17.
31. Günzel D, Yu ASL. Claudins and the modulation of tight junction permeability. Physiol Rev. 2013;93:525–69.
32. Szecskó A, Mészáros M, Simões B, Cavaco M, Chaparro C, Porkoláb G, et al. PepH3-modified nanocarriers for delivery of therapeutics across the blood-brain barrier. Fluids Barriers CNS. 2025;22:31.
33. Campos-Bedolla P, Walter FR, Veszelka S, Deli MA. Role of the blood-brain barrier in the nutrition of the central nervous system. Arch Med Res. 2014;45:610–38.
34. Ohtsuki S, Terasaki T. Contribution of carrier-mediated transport systems to the blood-brain barrier as a supporting and protecting

interface for the brain; importance for CNS drug discovery and development. Pharm Res. 2007;24:1745–58.
35. Zhang W, Liu QY, Haqqani AS, Liu Z, Sodja C, Leclerc S, et al. Differential expression of ABC transporter genes in brain vessels vs. peripheral tissues and vessels from human, mouse and rat. Pharmaceutics. 2023;15:1563.
36. Begley DJ. ABC transporters and the blood-brain barrier. Curr Pharm Des. 2004;10:1295–312.
37. Bellavance M-A, Blanchette M, Fortin D. Recent advances in blood-brain barrier disruption as a CNS delivery strategy. AAPS J. 2008;10:166–77.
38. Achar A, Myers R, Ghosh C. Drug delivery challenges in brain disorders across the blood–brain barrier: novel methods and future considerations for improved therapy. Biomedicines. 2021;9:1834.
39. Naslavsky N, Caplan S. Advances and challenges in understanding endosomal sorting and fission. FEBS J. 2023;290:4187–95.
40. Haqqani AS, Thom G, Burrell M, Delaney CE, Brunette E, Baumann E, et al. Intracellular sorting and transcytosis of the rat transferrin receptor antibody OX26 across the blood-brain barrier in vitro is dependent on its binding affinity. J Neurochem. 2018;146:735–52.
41. Arya SB, Collie SP, Parent CA. The ins-and-outs of exosome biogenesis, secretion, and internalization. Trends Cell Biol. 2024;34:90–108.
42. Belcher JD, Hamilton RL, Brady SE, Hornick CA, Jaeckle S, Schneider WJ, et al. Isolation and characterization of three endosomal fractions from the liver of estradiol-treated rats. Proc Natl Acad Sci USA. 1987;84:6785–9.
43. Hurley JH, Odorizzi G. Get on the exosome bus with ALIX. Nat Cell Biol. 2012;14:654–5.
44. Serra ND, Sundaram MV. Transcytosis in the development and morphogenesis of epithelial tissues. EMBO J. 2021;40:e106163.
45. Okamoto C. Endocytosis and transcytosis. Adv Drug Deliv Rev. 1998;29:215–28.
46. Worzfeld T, Schwaninger M. Apicobasal polarity of brain endothelial cells. J Cereb Blood Flow Metab. 2016;36:340–62.
47. Vion A-C, Perovic T, Petit C, Hollfinger I, Bartels-Klein E, Frampton E, et al. Endothelial cell orientation and polarity are controlled by shear stress and VEGF through distinct signaling pathways. Front Physiol. 2020;11:623769. https://doi.org/10.3389/fphys.2020.623769/full.
48. Pillay S, Meyer NL, Puschnik AS, Davulcu O, Diep J, Ishikawa Y, et al. An essential receptor for adeno-associated virus infection. Nature. 2016;530:108–12.
49. Pillay S, Zou W, Cheng F, Puschnik AS, Meyer NL, Ganaie SS, et al. Adeno-associated virus (AAV) serotypes have distinctive interactions with domains of the cellular AAV receptor. J Virol. 2017;91:e00391-17. https://doi.org/10.1128/JVI.00391-17.
50. Huang Q, Chan KY, Tobey IG, Chan YA, Poterba T, Boutros CL, et al. Delivering genes across the blood-brain barrier: LY6A, a novel cellular receptor for AAV-PHP.B capsids. PLoS ONE. 2019;14:e0225206.
51. Hordeaux J, Yuan Y, Clark PM, Wang Q, Martino RA, Sims JJ, et al. The GPI-linked protein LY6A drives AAV-PHP.B transport across the blood-brain barrier. Mol Ther. 2019;27:912–21.
52. Zhang Y, Guo P, Ma Z, Lu P, Kebebe D, Liu Z. Combination of cell-penetrating peptides with nanomaterials for the potential therapeutics of central nervous system disorders: a review. J Nanobiotechnol. 2021;19:255.
53. Sorets AG, Rosch JC, Duvall CL, Lippmann ES. Caveolae-mediated transport at the injured blood-brain barrier as an underexplored pathway for central nervous system drug delivery. Curr Opin Chem Eng. 2020;30:86–95.
54. Hervé F, Ghinea N, Scherrmann J-M. CNS delivery via adsorptive transcytosis. AAPS J. 2008;10:455–72.
55. Yue Q, Leng X, Xie N, Zhang Z, Yang D, Hoi MPM. Endothelial dysfunctions in blood-brain barrier breakdown in Alzheimer's disease: from mechanisms to potential therapies. CNS Neurosci Ther. 2024;30:e70079.
56. Lau K, Kotzur R, Richter F. Blood-brain barrier alterations and their impact on Parkinson's disease pathogenesis and therapy. Transl Neurodegener. 2024;13:37.
57. Khor SLQ, Ng KY, Koh RY, Chye SM. Blood-brain barrier and neurovascular unit dysfunction in Parkinson's disease: from clinical insights to pathogenic mechanisms and novel therapeutic approaches. CNS Neurol Disord Drug Targets. 2024;23:315–30.
58. Wei Z, Liu H, Lin Z, Yao M, Li R, Liu C, et al. Non-contrast assessment of blood-brain barrier permeability to water in mice: an arterial spin labeling study at cerebral veins. Neuroimage. 2023;268:119870.
59. Di Pardo A, Amico E, Scalabrì F, Pepe G, Castaldo S, Elifani F, et al. Impairment of blood-brain barrier is an early event in R6/2 mouse model of Huntington disease. Sci Rep. 2017;7:41316.
60. Drouin-Ouellet J, Sawiak SJ, Cisbani G, Lagacé M, Kuan W-L, Saint-Pierre M, et al. Cerebrovascular and blood-brain barrier impairments in Huntington's disease: potential implications for its pathophysiology. Ann Neurol. 2015;78:160–77.
61. Manora L, Borlongan CV, Garbuzova-Davis S. Cellular and non-cellular approaches for repairing the damaged blood-CNS-barrier in amyotrophic lateral sclerosis. Cells. 2024;13:435.
62. Steinruecke M, Lonergan RM, Selvaraj BT, Chandran S, Diaz-Castro B, Stavrou M. Blood-CNS barrier dysfunction in amyotrophic lateral sclerosis: proposed mechanisms and clinical implications. J Cereb Blood Flow Metab. 2023;43:642–54.
63. Alarcan H, Al Ojaimi Y, Lanznaster D, Escoffre J-M, Corcia P, Vourc'h P, et al. Taking advantages of blood-brain or spinal cord barrier alterations or restoring them to optimize therapy in ALS? J Pers Med. 2022;12:1071.
64. Zierfuss B, Larochelle C, Prat A. Blood-brain barrier dysfunction in multiple sclerosis: causes, consequences, and potential effects of therapies. Lancet Neurol. 2024;23:95–109.
65. Schreiner TG, Romanescu C, Popescu BO. The blood-brain barrier: a key player in multiple sclerosis disease mechanisms. Biomolecules. 2022;12:538.
66. Tang L, Zhang R, Wang Y, Liu M, Hu D, Wang Y, et al. A blood-brain barrier- and blood-brain tumor barrier-penetrating siRNA delivery system targeting gliomas for brain tumor immunotherapy. J Control Release. 2024;369:642–57.
67. Arvanitis CD, Ferraro GB, Jain RK. The blood-brain barrier and blood-tumour barrier in brain tumours and metastases. Nat Rev Cancer. 2020;20:26–41.
68. Zhao C, Zhu X, Tan J, Mei C, Cai X, Kong F. Lipid-based nanoparticles to address the limitations of GBM therapy by overcoming the blood-brain barrier, targeting glioblastoma stem cells, and counteracting the immunosuppressive tumor microenvironment. Biomed Pharmacother. 2024;171:116113.
69. Kwak G, Grewal A, Slika H, Mess G, Li H, Kwatra M, et al. Brain nucleic acid delivery and genome editing via focused ultrasound-mediated blood-brain barrier opening and long-circulating nanoparticles. ACS Nano. 2024;18:24139–53.
70. Hirschberg H, Madsen SJ. Cell mediated photothermal therapy of brain tumors. J Neuroimmune Pharmacol. 2017;12:99–106.
71. Zhang S, Cui N, Wu Y, Zhong W, Johnson CM, Jiang C. Optogenetic intervention to the vascular endothelium. Vascul Pharmacol. 2015;74:122–9.
72. Busquets MA, Espargaró A, Sabaté R, Estelrich J. Magnetic nanoparticles cross the blood-brain barrier: when physics rises to a challenge. Nanomaterials (Basel). 2015;5:2231–48.
73. Wang C, Xue Y, Markovic T, Li H, Wang S, Zhong Y, et al. Blood-brain-barrier-crossing lipid nanoparticles for

mRNA delivery to the central nervous system. Nat Mater. 2025;24:1653–63.
74. Xue Y, Wang C, Li H, Du S, Zhong Y, Zhang Y, et al. Lipid nanoparticles enhance mRNA delivery to the central nervous system upon intrathecal injection. Adv Mater. 2025;37:e2417097.
75. Ma F, Yang L, Sun Z, Chen J, Rui X, Glass Z, et al. Neurotransmitter-derived lipidoids (NT-lipidoids) for enhanced brain delivery through intravenous injection. Sci Adv. 2020;6:eabb4429.
76. Akinc A, Maier MA, Manoharan M, Fitzgerald K, Jayaraman M, Barros S, et al. The Onpattro story and the clinical translation of nanomedicines containing nucleic acid-based drugs. Nat Nanotechnol. 2019;14:1084–7.
77. Wang W, Chen K, Jiang T, Wu Y, Wu Z, Ying H, et al. Artificial intelligence-driven rational design of ionizable lipids for mRNA delivery. Nat Commun. 2024;15:10804.
78. Cao D, Hou X, Wang C, Wang S, Liu Z, Tian M, et al. Lipid nanoparticles for mRNA delivery in brain via systemic administration. Sci Adv. 2025;11:eadw0730.
79. Bian X, Guo Q, Yau L-F, Yang L, Wang X, Zhao S, et al. Berberine-inspired ionizable lipid for self-structure stabilization and brain targeting delivery of nucleic acid therapeutics. Nat Commun. 2025;16:2368.
80. Teixeira MI, Lopes CM, Amaral MH, Costa PC. Surface-modified lipid nanocarriers for crossing the blood-brain barrier (BBB): a current overview of active targeting in brain diseases. Colloids Surf B Biointerfaces. 2023;221:112999.
81. Han EL, Tang S, Kim D, Murray AM, Swingle KL, Hamilton AG, et al. Peptide-functionalized lipid nanoparticles for targeted systemic mRNA delivery to the brain. Nano Lett. 2025;25:800–10.
82. Guo Q, Xu S, Yang P, Wang P, Lu S, Sheng D, et al. A dual-ligand fusion peptide improves the brain-neuron targeting of nanocarriers in Alzheimer's disease mice. J Control Release. 2020;320:347–62.
83. Zheng C, Ma C, Bai E, Yang K, Xu R. Transferrin and cell-penetrating peptide dual-functioned liposome for targeted drug delivery to glioma. Int J Clin Exp Med. 2015;8:1658–68.
84. Loureiro JA, Gomes B, Coelho MAN, do Carmo Pereira M, Rocha S. Targeting nanoparticles across the blood-brain barrier with monoclonal antibodies. Nanomedicine (Lond). 2014;9:709–22.
85. Johnsen KB, Burkhart A, Thomsen LB, Andresen TL, Moos T. Targeting the transferrin receptor for brain drug delivery. Prog Neurobiol. 2019;181:101665.
86. Yu YJ, Zhang Y, Kenrick M, Hoyte K, Luk W, Lu Y, et al. Boosting brain uptake of a therapeutic antibody by reducing its affinity for a transcytosis target. Sci Transl Med. 2011;3:84ra44.
87. Ye Z, Gastfriend BD, Umlauf BJ, Lynn DM, Shusta EV. Antibody-targeted liposomes for enhanced targeting of the blood-brain barrier. Pharm Res. 2022;39:1523–34.
88. Gao Y, Zhu J, Lu H. Single domain antibody-based vectors in the delivery of biologics across the blood-brain barrier: a review. Drug Deliv Transl Res. 2021;11:1818–28.
89. Yogi A, Hussack G, van Faassen H, Haqqani AS, Delaney CE, Brunette E, et al. Brain delivery of IGF1R5, a single-domain antibody targeting insulin-like growth factor-1 receptor. Pharmaceutics. 2022;14:452.
90. Johnsen KB, Bak M, Kempen PJ, Melander F, Burkhart A, Thomsen MS, et al. Antibody affinity and valency impact brain uptake of transferrin receptor-targeted gold nanoparticles. Theranostics. 2018;8:3416–36.
91. Stocki P, Szary J, Rasmussen CLM, Demydchuk M, Northall L, Logan DB, et al. Blood-brain barrier transport using a high affinity, brain-selective VNAR antibody targeting transferrin receptor 1. FASEB J. 2021;35:e21172.
92. Costain WJ, Haqqani AS, Hussack G, van Faassen H, Lessard E, Ling B, et al. In vivo brain delivery of BBB-enabled iduronate 2-sulfatase in rats. Fluids Barriers CNS. 2025;22:7. https://doi.org/10.1186/s12987-024-00617-6.
93. Gentry K, Lian L, Kim H, Celik O, Jones C, Podilapu AR, et al. Glycolipid nanoparticles target the spleen and detarget the liver without charge. Proc Natl Acad Sci U S A. 2025;122:e2409569122. https://doi.org/10.1073/pnas.2409569122.
94. Pang X, Cui C, Wan S, Jiang Y, Zhang L, Xia L, et al. Bioapplications of Cell-SELEX-generated aptamers in cancer diagnostics, therapeutics, theranostics and biomarker discovery: a comprehensive review. Cancers (Basel). 2018;10(2):47.
95. Monaco I, Camorani S, Colecchia D, Locatelli E, Calandro P, Oudin A, et al. Aptamer functionalization of nanosystems for glioblastoma targeting through the blood-brain barrier. J Med Chem. 2017;60:4510–6.
96. Bukari B, Samarasinghe RM, Noibanchong J, Shigdar SL. Non-invasive delivery of therapeutics into the vrain: the potential of aptamers for targeted delivery. Biomedicines. 2020;8:120.
97. Koo J, Lim C, Oh KT. Recent advances in intranasal administration for brain-targeting delivery: a comprehensive review of lipid-based nanoparticles and stimuli-responsive gel formulations. Int J Nanomedicine. 2024;19:1767–807.
98. Dhaliwal HK, Fan Y, Kim J, Amiji MM. Intranasal delivery and transfection of mRNA therapeutics in the brain using cationic liposomes. Mol Pharm. 2020;17:1996–2005.
99. Tuma J, Chen Y-J, Collins MG, Paul A, Li J, Han H, et al. Lipid nanoparticles deliver mRNA to the brain after an intracerebral injection. Biochemistry. 2023;62:3533–47.
100. Tanaka H, Nakatani T, Furihata T, Tange K, Nakai Y, Yoshioka H, et al. In vivo introduction of mRNA encapsulated in lipid nanoparticles to brain neuronal cells and astrocytes via intracerebroventricular administration. Mol Pharm. 2018;15:2060–7.
101. Hirunagi T, Sahashi K, Tachikawa K, Leu AI, Nguyen M, Mukthavaram R, et al. Selective suppression of polyglutamine-expanded protein by lipid nanoparticle-delivered siRNA targeting CAG expansions in the mouse CNS. Mol Ther Nucleic Acids. 2021;24:1–10.
102. Slavc I, Cohen-Pfeffer JL, Gururangan S, Krauser J, Lim DA, Maldaun M, et al. Best practices for the use of intracerebroventricular drug delivery devices. Mol Genet Metab. 2018;124:184–8.
103. Madadi AK, Sohn M-J. Advances in intrathecal nanoparticle delivery: targeting the blood-cerebrospinal fluid barrier for enhanced CNS drug delivery. Pharmaceuticals (Basel). 2024;17:1070.
104. Meyer AH, Feldsien TM, Mezler M, Untucht C, Venugopalan R, Lefebvre DR. Novel developments to enable treatment of CNS diseases with targeted drug delivery. Pharmaceutics. 2023;15:1100.
105. Finkel RS, Mercuri E, Darras BT, Connolly AM, Kuntz NL, Kirschner J, et al. Nusinersen versus sham control in infantile-onset spinal muscular atrophy. N Engl J Med. 2017;377:1723–32.
106. Ahmed T. Lipid nanoparticle mediated small interfering RNA delivery as a potential therapy for Alzheimer's disease. Eur J Neurosci. 2024;59:2915–54.
107. Yaghi NK, Wei J, Hashimoto Y, Kong L-Y, Gabrusiewicz K, Nduom EK, et al. Immune modulatory nanoparticle therapeutics for intracerebral glioma. Neuro Oncol. 2017;19:372–82.
108. Ralvenius WT, Andresen JL, Huston MM, Penney J, Bonner JM, Fenton OS, et al. Nanoparticle-mediated delivery of anti-PU.1 siRNA via localized intracisternal administration reduces neuroinflammation. Adv Mater. 2024;36:e2309225.

146. Gordon-Lipkin E, Fatemi A. Current therapeutic approaches in leukodystrophies: a review. J Child Neurol. 2018;33:861–8. https://doi.org/10.1177/0883073818792313.
147. Maretina MA, Zheleznyakova GY, Lanko KM, Egorova AA, Baranov VS, Kiselev AV. Molecular factors involved in spinal muscular atrophy pathways as possible disease-modifying candidates. Curr Genomics. 2018;19:339–55. https://doi.org/10.2174/1389202919666180101154916.
148. Edinoff AN, Nguyen LH, Odisho AS, Maxey BS, Pruitt JW, Girma B, et al. The antisense oligonucleotide nusinersen for treatment of spinal muscular atrophy. Orthop Rev (Pavia). 2021;13:24934. https://doi.org/10.52965/001c.24934.
149. Alves CRR, Ha LL, Yaworski R, Lazzarotto CR, Christie KA, Reilly A, et al. Base editing as a genetic treatment for spinal muscular atrophy. bioRxiv [Preprint]. 2023:2023.01.20.524978. https://doi.org/10.1101/2023.01.20.524978.
150. Hassan S, Sidransky E, Tayebi N. The role of epigenetics in lysosomal storage disorders: uncharted territory. Mol Genet Metab. 2017;122:10–8. https://doi.org/10.1016/j.ymgme.2017.07.012.
151. Kido J, Sugawara K, Nakamura K. Gene therapy for lysosomal storage diseases: current clinical trial prospects. Front Genet. 2023;14:1064924. https://doi.org/10.3389/fgene.2023.1064924.
152. Ługowska A. Oncological aspects of lysosomal storage diseases. Cells. 2024;13(19):1664. https://doi.org/10.3390/cells13191664.
153. Edelmann MJ, Maegawa GHB. CNS-targeting therapies for lysosomal storage diseases: current advances and challenges. Front Mol Biosci. 2020;7:559804.
154. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 2012;337:816–21. https://doi.org/10.1126/science.1225829.
155. Chen JS, Ma E, Harrington LB, Da Costa M, Tian X, Palefsky JM, et al. CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity. Science. 2018;360:436–9. https://doi.org/10.1126/science.aar6245.
156. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. Nat Protoc. 2013;8:2281–308. https://doi.org/10.1038/nprot.2013.143.
157. Yarnall MTN, Ioannidi EI, Schmitt-Ulms C, Krajeski RN, Lim J, Villiger L, et al. Drag-and-drop genome insertion of large sequences without double-strand DNA cleavage using CRISPR-directed integrases. Nat Biotechnol. 2023;41:500–12. https://doi.org/10.1038/s41587-022-01527-4.
158. Hagberg B. Rett's syndrome: prevalence and impact on progressive severe mental retardation in girls. Acta Paediatr Scand. 1985;74:405–8. https://doi.org/10.1111/j.1651-2227.1985.tb10993.x.
159. Hagberg B, Aicardi J, Dias K, Ramos O. A progressive syndrome of autism, dementia, ataxia, and loss of purposeful hand use in girls: Rett's syndrome: report of 35 cases. Ann Neurol. 1983;14:471–9. https://doi.org/10.1002/ana.410140412.
160. Palmieri M, Pozzer D, Landsberger N. Advanced genetic therapies for the treatment of Rett syndrome: state of the art and future perspectives. Front Neurosci. 2023;17:1172805. https://doi.org/10.3389/fnins.2023.1172805.
161. Le TTH, Tran NT, Dao TML, Nguyen DD, Do HD, Ha TL, et al. Efficient and precise CRISPR/Cas9-mediated MECP2 modifications in human-induced pluripotent stem cells. Front Genet. 2019;10:625. https://doi.org/10.3389/fgene.2019.00625.
162. Gilliam TC, Tanzi RE, Haines JL, Bonner TI, Faryniarz AG, Hobbs WJ, et al. Localization of the Huntington's disease gene to a small segment of chromosome 4 flanked by D4S10 and the telomere. Cell. 1987;50:565–71. https://doi.org/10.1016/0092-8674(87)90029-8.
163. Saudou F, Humbert S. The biology of huntingtin. Neuron. 2016;89:910–26. https://doi.org/10.1016/j.neuron.2016.02.003.
164. Ekman FK, Ojala DS, Adil MM, Lopez PA, Schaffer DV, Gaj T. CRISPR-Cas9-mediated genome editing increases lifespan and improves motor deficits in a Huntington's disease mouse model. Mol Ther Nucleic Acids. 2019;17:829–39. https://doi.org/10.1016/j.omtn.2019.07.009.
165. Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature. 2016;533:420–4. https://doi.org/10.1038/nature17946.
166. Anzalone AV, Randolph PB, Davis JR, Sousa AA, Koblan LW, Levy JM, et al. Search-and-replace genome editing without double-strand breaks or donor DNA. Nature. 2019;576:149–57. https://doi.org/10.1038/s41586-019-1711-4.
167. Gaudelli NM, Komor AC, Rees HA, Packer MS, Badran AH, Bryson DI, et al. Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. Nature. 2017;551:464–71.
168. Masoud M, Prange L, Wuchich J, Hunanyan A, Mikati MA. Diagnosis and treatment of alternating hemiplegia of childhood. Curr Treat Options Neurol. 2017;19:8. https://doi.org/10.1007/s11940-017-0444-7.
169. Sousa AA, Terrey M, Sakai HA, Simmons CQ, Arystarkhova E, Morsci NS, et al. In vivo prime editing rescues alternating hemiplegia of childhood in mice. Cell. 2025;188:4275-94.e23. https://doi.org/10.1016/j.cell.2025.06.038.
170. Konermann S, Lotfy P, Brideau NJ, Oki J, Shokhirev MN, Hsu PD. Transcriptome engineering with RNA-targeting type VI-D CRISPR effectors. Cell. 2018;173:665-76.e14. https://doi.org/10.1016/j.cell.2018.02.033.
171. Larson MH, Gilbert LA, Wang X, Lim WA, Weissman JS, Qi LS. CRISPR interference (CRISPRi) for sequence-specific control of gene expression. Nat Protoc. 2013;8:2180–96. https://doi.org/10.1038/nprot.2013.132.
172. Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell. 2013;152:1173–83. https://doi.org/10.1016/j.cell.2013.02.022.
173. Park H, Hwang Y, Kim J. Transcriptional activation with Cas9 activator nanocomplexes rescues Alzheimer's disease pathology. Biomaterials. 2021;279:121229. https://doi.org/10.1016/j.biomaterials.2021.121229.
174. Colasante G, Lignani G, Brusco S, Di Berardino C, Carpenter J, Giannelli S, et al. DCas9-based Scn1a gene activation restores inhibitory interneuron excitability and attenuates seizures in Dravet syndrome mice. Mol Ther. 2020;28:235–53. https://doi.org/10.1016/j.ymthe.2019.08.018.
175. Fadul SM, Arshad A, Mehmood R. CRISPR-based epigenome editing: mechanisms and applications. Epigenomics. 2023;15:1137–55. https://doi.org/10.2217/epi-2023-0281.
176. Nakamura M, Gao Y, Dominguez AA, Qi LS. CRISPR technologies for precise epigenome editing. Nat Cell Biol. 2021;23:11–22. https://doi.org/10.1038/s41556-020-00620-7.
177. Xiong T, Meister GE, Workman RE, Kato NC, Spellberg MJ, Turker F, et al. Targeted DNA methylation in human cells using engineered dCas9-methyltransferases. Sci Rep. 2017;7:6732. https://doi.org/10.1038/s41598-017-06757-0.
178. Qian J, Liu SX. CRISPR/dCas9-Tet1-mediated DNA methylation editing. Bio Protoc. 2024;14:e4976. https://doi.org/10.21769/BioProtoc.4976.
179. Morita S, Noguchi H, Horii T, Nakabayashi K, Kimura M, Okamura K, et al. Targeted DNA demethylation in vivo using dCas9-peptide repeat and scFv-TET1 catalytic domain fusions. Nat Biotechnol. 2016;34:1060–5. https://doi.org/10.1038/nbt.3658.
180. O'Brien RJ, Wong PC. Amyloid precursor protein processing and Alzheimer's disease. Annu Rev Neurosci. 2011;34:185–204. https://doi.org/10.1146/annurev-neuro-061010-113613.

109. Wang J, Ding Y, Chong K, Cui M, Cao Z, Tang C, et al. Recent advances in lipid nanoparticles and their safety concerns for mRNA delivery. Vaccines. 2024;12:1148.
110. Gulati NM, Stewart PL, Steinmetz NF. Bioinspired shielding strategies for nanoparticle drug delivery applications. Mol Pharm. 2018;15:2900–9.
111. Estapé Senti M, de Jongh CA, Dijkxhoorn K, Verhoef JJF, Szebeni J, Storm G, et al. Anti-PEG antibodies compromise the integrity of PEGylated lipid-based nanoparticles via complement. J Control Release. 2022;341:475–86.
112. Sabnis S, Kumarasinghe ES, Salerno T, Mihai C, Ketova T, Senn JJ, et al. A novel amino lipid series for mRNA delivery: improved endosomal escape and sustained pharmacology and safety in non-human primates. Mol Ther. 2018;26:1509–19.
113. Adams D, Gonzalez-Duarte A, O'Riordan WD, Yang C-C, Ueda M, Kristen AV, et al. Patisiran, an RNAi therapeutic, for hereditary transthyretin amyloidosis. N Engl J Med. 2018;379:11–21.
114. Gillmore JD, Gane E, Taubel J, Kao J, Fontana M, Maitland ML, et al. CRISPR-Cas9 in vivo gene editing for transthyretin amyloidosis. N Engl J Med. 2021;385:493–502.
115. Gurevitz C, Bajaj A, Khera AV, Do R, Schunkert H, Musunuru K, et al. Gene therapy and genome editing for lipoprotein disorders. Eur Heart J. 2025;46:3420–33.
116. Lee RG, Mazzola AM, Braun MC, Platt C, Vafai SB, Kathiresan S, et al. Efficacy and safety of an investigational single-course CRISPR base-editing therapy targeting PCSK9 in nonhuman primate and mouse models. Circulation. 2023;147:242–53.
117. Khirallah J, Bloomer H, Wich D, Huang C, Workman JN, Li Y, et al. In vivo base editing of Angptl3 via lipid nanoparticles to treat cardiovascular disease. Mol Ther Nucleic Acids. 2025;36:102486.
118. Tremblay F, Xiong Q, Shah SS, Ko C-W, Kelly K, Morrison MS, et al. A potent epigenetic editor targeting human PCSK9 for durable reduction of low-density lipoprotein cholesterol levels. Nat Med. 2025;31:1329–38.
119. Wang S, Xu X, Viitala T, Shu Y, Zhang H. Unlocking mRNA-driven CRISPR-Cas9 gene therapy via optimizing mRNA and the delivery vectors. Mol Ther Nucleic Acids. 2025;36:102737.
120. Ah Mew N, Lichter-Konecki U. Current treatment modalities for urea cycle disorders. Pediatr Drugs. 2025;27:723–34. https://doi.org/10.1007/s40272-025-00719-0.
121. Wei M, Qian N, Gao X, Lang X, Song D, Min W. Single-particle imaging of nanomedicine entering the brain. Proc Natl Acad Sci U S A. 2024;121:e2309811121.
122. Pfau SJ, Langen UH, Fisher TM, Prakash I, Nagpurwala F, Lozoya RA, et al. Characteristics of blood-brain barrier heterogeneity between brain regions revealed by profiling vascular and perivascular cells. Nat Neurosci. 2024;27:1892–903.
123. Erdő F, Denes L, de Lange E. Age-associated physiological and pathological changes at the blood-brain barrier: a review. J Cereb Blood Flow Metab. 2017;37:4–24.
124. Sweeney MD, Zhao Z, Montagne A, Nelson AR, Zlokovic BV. Blood-brain barrier: from physiology to disease and back. Physiol Rev. 2019;99:21–78.
125. Kazemian P, Yu S-Y, Thomson SB, Birkenshaw A, Leavitt BR, Ross CJD. Lipid-nanoparticle-based delivery of CRISPR/Cas9 genome-editing components. Mol Pharm. 2022;19:1669–86.
126. Maeder ML, Stefanidakis M, Wilson CJ, Baral R, Barrera LA, Bounoutas GS, et al. Development of a gene-editing approach to restore vision loss in Leber congenital amaurosis type 10. Nat Med. 2019;25:229–33. https://doi.org/10.1038/s41591-018-0327-9.
127. Almasoudi HH. Therapeutic promise of CRISPR-Cas9 gene editing in sickle cell disease and β-thalassemia: a current review. Curr Res Transl Med. 2025;73:103513. https://doi.org/10.1016/j.retram.2025.103513.
128. Calabrò M, Rinaldi C, Santoro G, Crisafulli C. The biological pathways of Alzheimer disease: a review. AIMS Neurosci. 2021;8:86–132. https://doi.org/10.3934/Neuroscience.2021005.
129. Moosavi SG, Rahiman N, Jaafari MR, Arabi L. Lipid nanoparticle (LNP) mediated mRNA delivery in neurodegenerative diseases. J Control Release. 2025;381:113641. https://doi.org/10.1016/j.jconrel.2025.113641.
130. Imran Sajid M, Sultan Sheikh F, Anis F, Nasim N, Sumbria RK, Nauli SM, et al. siRNA drug delivery across the blood-brain barrier in Alzheimer's disease. Adv Drug Deliv Rev. 2023;199:114968. https://doi.org/10.1016/j.addr.2023.114968.
131. Fujita KA, Ostaszewski M, Matsuoka Y, Ghosh S, Glaab E, Trefois C, et al. Integrating pathways of Parkinson's disease in a molecular interaction map. Mol Neurobiol. 2014;49:88–102. https://doi.org/10.1007/s12035-013-8489-4.
132. Dong-Chen X, Yong C, Yang X, Chen-Yu S, Li-Hua P. Signaling pathways in Parkinson's disease: molecular mechanisms and therapeutic interventions. Signal Transduct Target Ther. 2023;8:73–73. https://doi.org/10.1038/s41392-023-01353-3.
133. Li Y, Zhang J, Ma B, Yu W, Xu M, Luan W, et al. Nanotechnology used for siRNA delivery for the treatment of neurodegenerative diseases: focusing on Alzheimer's disease and Parkinson's disease. Int J Pharm. 2024;666:124786. https://doi.org/10.1016/j.ijpharm.2024.124786.
134. Labbadia J, Morimoto RI. Huntington's disease: underlying molecular mechanisms and emerging concepts. Trends Biochem Sci. 2013;38:378–85. https://doi.org/10.1016/j.tibs.2013.05.003.
135. Lin Y, Li C, Chen Y, Gao J, Li J, Huang C, et al. RNA-Targeting CRISPR/CasRx system relieves disease symptoms in Huntington's disease models. Mol Neurodegener. 2025;20:4. https://doi.org/10.1186/s13024-024-00794-w.
136. Alkanli SS, Alkanli N, Ay A, Albeniz I. CRISPR/Cas9 mediated therapeutic approach in Huntington's disease. Mol Neurobiol. 2023;60:1486–98. https://doi.org/10.1007/s12035-022-03150-5.
137. Rook ME, Southwell AL. Antisense oligonucleotide therapy: from design to the Huntington disease clinic. BioDrugs. 2022;36:105–19. https://doi.org/10.1007/s40259-022-00519-9.
138. Ehrhart F, Coort SLM, Cirillo E, Smeets E, Evelo CT, Curfs LMG. Rett syndrome: biological pathways leading from MECP2 to disorder phenotypes. Orphanet J Rare Dis. 2016;11:158–158. https://doi.org/10.1186/s13023-016-0545-5.
139. Cho H-Y, Yoo M, Pongkulapa T, Rabie H, Muotri AR, Yin PT, et al. Magnetic nanoparticle-assisted non-viral CRISPR-Cas9 for enhanced genome editing to treat Rett syndrome. Adv Sci. 2024;11:e2306432. https://doi.org/10.1002/advs.202306432.
140. Coorey B, Haase F, Ellaway C, Clarke A, Lisowski L, Gold WA. Gene editing and Rett syndrome: does it make the cut? CRISPR J. 2022;5:490–9.
141. Farsi Z, Allahyari Fard N. The identification of key genes and pathways in glioblastoma by bioinformatics analysis. Mol Cell Oncol. 2023;10:2246657. https://doi.org/10.1080/23723556.2023.2246657.
142. Li L, Liu X, Ma X, Deng X, Ji T, Hu P, et al. Identification of key candidate genes and pathways in glioblastoma by integrated bioinformatical analysis. Exp Ther Med. 2019;18:3439–49. https://doi.org/10.3892/etm.2019.7975.
143. Romani M, Pistillo MP, Banelli B. Epigenetic targeting of glioblastoma. Front Oncol. 2018;8:448. https://doi.org/0.3389/fonc.2018.00448.
144. Begagić E, Bečulić H, Đuzić N, Džidić-Krivić A, Pugonja R, Muharemović A, et al. CRISPR/Cas9-mediated gene therapy for glioblastoma: a scoping review. Biomedicines. 2024;12:238. https://doi.org/10.3390/biomedicines12010238.
145. Aerts-Kaya F, van Til NP. Gene and cellular therapies for leukodystrophies. Pharmaceutics. 2023;15:2522. https://doi.org/10.3390/pharmaceutics15112522.

181. Park H, Shin J, Kim Y, Saito T, Saido TC, Kim J. CRISPR/dCas9-Dnmt3a-mediated targeted DNA methylation of APP rescues brain pathology in a mouse model of Alzheimer's disease. Transl Neurodegener. 2022;11:41. https://doi.org/10.1186/s40035-022-00314-0.
182. Spillantini MG, Schmidt ML, Lee VM, Trojanowski JQ, Jakes R, Goedert M. Alpha-synuclein in Lewy bodies. Nature. 1997;388:839–40. https://doi.org/10.1038/42166.
183. Polymeropoulos MH, Lavedan C, Leroy E, Ide SE, Dehejia A, Dutra A, et al. Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. Science. 1997;276:2045–7. https://doi.org/10.1126/science.276.5321.2045.
184. Guhathakurta S, Kim J, Adams L, Basu S, Song MK, Adler E, et al. Targeted attenuation of elevated histone marks at SNCA alleviates α-synuclein in Parkinson's disease. EMBO Mol Med. 2021;13:e12188. https://doi.org/10.15252/emmm.202012188.
185. Liu XS, Wu H, Krzisch M, Wu X, Graef J, Muffat J, et al. Rescue of Fragile X syndrome neurons by DNA methylation editing of the FMR1 gene. Cell. 2018;172:979-92.e6.
186. Cheng Q, Wei T, Farbiak L, Johnson LT, Dilliard SA, Siegwart DJ. Selective organ targeting (SORT) nanoparticles for tissue-specific mRNA delivery and CRISPR-Cas gene editing. Nat Nanotechnol. 2020;15:313–20. https://doi.org/10.1038/s41565-020-0669-6.
187. Onuma H, Sato Y, Harashima H. Lipid nanoparticle-based ribonucleoprotein delivery for in vivo genome editing. J Control Release. 2023;355:406–16. https://doi.org/10.1016/j.jconrel.2023.02.008.
188. Im SH, Jang M, Park J-H, Chung HJ. Finely tuned ionizable lipid nanoparticles for CRISPR/Cas9 ribonucleoprotein delivery and gene editing. J Nanobiotechnol. 2024;22:175. https://doi.org/10.1186/s12951-024-02427-2.
189. Chen K, Han H, Zhao S, Xu B, Yin B, Lawanprasert A, et al. Lung and liver editing by lipid nanoparticle delivery of a stable CRISPR-Cas9 ribonucleoprotein. Nat Biotechnol. 2025;43:1445–7. https://doi.org/10.1038/s41587-024-02437-3.
190. Masarwy R, Breier D, Stotsky-Oterin L, Ad-El N, Qassem S, Naidu GS, et al. Targeted CRISPR/Cas9 lipid nanoparticles elicits therapeutic genome editing in head and neck cancer. Adv Sci. 2025;12:e2411032. https://doi.org/10.1002/advs.202411032.
191. Rosenblum D, Gutkin A, Kedmi R, Ramishetti S, Veiga N, Jacobi AM, et al. CRISPR-Cas9 genome editing using targeted lipid nanoparticles for cancer therapy. Sci Adv. 2020;6:eabc9450. https://doi.org/10.1126/sciadv.abc9450.
192. Zou Y, Sun X, Yang Q, Zheng M, Shimoni O, Ruan W, et al. Blood-brain barrier-penetrating single CRISPR-Cas9 nanocapsules for effective and safe glioblastoma gene therapy. Sci Adv. 2022;8:eabm8011. https://doi.org/10.1126/sciadv.abm8011.
193. Zhu Y, Zhu L, Wang X, Jin H. RNA-based therapeutics: an overview and prospectus. Cell Death Dis. 2022;13:644. https://doi.org/10.1038/s41419-022-05075-2.
194. Marino M, Mele E, Pastorino GMG, Meccariello R, Operto FF, Santoro A, et al. Neuroinflammation: molecular mechanisms and therapeutic perspectives. Cent Nerv Syst Agents Med Chem. 2022;22:160–74. https://doi.org/10.2174/1871524922666220929153215.
195. Gomes AQ, Nolasco S, Soares H. Non-coding RNAs: multitasking molecules in the cell. Int J Mol Sci. 2013;14:16010–39. https://doi.org/10.3390/ijms140816010.
196. Mikutis S, Bernardes GJL. Technologies for targeted RNA degradation and induced RNA decay. Chem Rev. 2024;124:13301–30. https://doi.org/10.1021/acs.chemrev.4c00472.
197. Sparmann A, Vogel J. RNA-based medicine: from molecular mechanisms to therapy. EMBO J. 2023;42:e114760. https://doi.org/10.15252/embj.2023114760.
198. Porcari GS, Collyer JW, Adang LA, Rajan DS. Current advances and challenges in gene therapies for neurologic disorders: a review for the clinician. Neurol Genet. 2025;11:e200229. https://doi.org/10.1212/NXG.0000000000200229.
199. Yeoh YQ, Amin A, Cuic B, Tomas D, Turner BJ, Shabanpoor F. Efficient systemic CNS delivery of a therapeutic antisense oligonucleotide with a blood-brain barrier-penetrating ApoE-derived peptide. Biomed Pharmacother. 2024;175:116737. https://doi.org/10.1016/j.biopha.2024.116737.
200. Dana H, Chalbatani GM, Mahmoodzadeh H, Karimloo R, Rezaiean O, Moradzadeh A, et al. Molecular mechanisms and biological functions of siRNA. Int J Biomed Sci. 2017;13:48–57.
201. Bhati V, Prasad S, Kabra A. RNA-based therapies for neurodegenerative disease: targeting molecular mechanisms for disease modification. Mol Cell Neurosci. 2025;133:104010. https://doi.org/10.1016/j.mcn.2025.104010.
202. Kumthekar P, Ko CH, Paunesku T, Dixit K, Sonabend AM, Bloch O, et al. A first-in-human phase 0 clinical study of RNA interference-based spherical nucleic acids in patients with recurrent glioblastoma. Sci Transl Med. 2021;13:eabb3945. https://doi.org/10.1126/scitranslmed.abb3945.
203. Molyneaux M, Xu J, Evans DM, Lu P. Effect on tumor growth by TGF-β_1/COX-2 siRNA combination product (STP705) in a human cholangiocarcinoma (HuCCT-1) xenograft tumor model in nude mice. J Clin Oncol. 2019;37:e14652. https://doi.org/10.1200/jco.2019.37.15_suppl.e14652.
204. Baumann V, Winkler J. miRNA-based therapies: strategies and delivery platforms for oligonucleotide and non-oligonucleotide agents. Future Med Chem. 2014;6:1967–84. https://doi.org/10.4155/fmc.14.116.
205. Asakiya C, Zhu L, Yuhan J, Zhu L, Huang K, Xu W. Current progress of miRNA-derivative nucleotide drugs: modifications, delivery systems, applications. Expert Opin Drug Deliv. 2022;19:435–50. https://doi.org/10.1080/17425247.2022.2063835.
206. Chaudhary N, Weissman D, Whitehead KA. mRNA vaccines for infectious diseases: principles, delivery and clinical translation. Nat Rev Drug Discov. 2021;20:817–38. https://doi.org/10.1038/s41573-021-00283-5.
207. Kiaie SH, Majidi Zolbanin N, Ahmadi A, Bagherifar R, Valizadeh H, Kashanchi F, et al. Recent advances in mRNA-LNP therapeutics: immunological and pharmacological aspects. J Nanobiotechnol. 2022;20:276. https://doi.org/10.1186/s12951-022-01478-7.
208. Jung S-J, Seo JJ, Lee S, Hyun S-I, Lee J, Lee S, et al. RNA stability enhancers for durable base-modified mRNA therapeutics. Nat Biotechnol. 2025. https://doi.org/10.1038/s41587-025-02891-7 **(Online ahead of print)**.
209. Sahin U, Karikó K, Türeci Ö. mRNA-based therapeutics: developing a new class of drugs. Nat Rev Drug Discov. 2014;13:759–80. https://doi.org/10.1038/nrd4278.
210. Zhong H, Liu S, Cao F, Zhao Y, Zhou J, Tang F, et al. Dissecting tumor antigens and immune subtypes of glioma to develop mRNA vaccine. Front Immunol. 2021;12:709986. https://doi.org/10.3389/fimmu.2021.709986.
211. Anthony K. RNA-based therapeutics for neurological diseases. RNA Biol. 2022;19:176–90. https://doi.org/10.1080/15476286.2021.2021650.
212. Li H, Cao Y, Ye J, Yang Z, Chen Q, Liu X, et al. Engineering brain-derived neurotrophic factor mRNA delivery for the treatment of Alzheimer's disease. Chem Eng J. 2023;466:143152.
213. Gao M, Li Y, Ho W, Chen C, Chen Q, Li F, et al. Targeted mRNA nanoparticles ameliorate blood–brain barrier disruption postischemic stroke by modulating microglia polarization. ACS Nano. 2024;18:3260–75. https://doi.org/10.1021/acsnano.3c09817.

214. Olson KE, Namminga KL, Lu Y, Thurston MJ, Schwab AD, de Picciotto S, et al. Granulocyte-macrophage colony-stimulating factor mRNA and neuroprotective immunity in Parkinson's disease. Biomaterials. 2021;272:120786.
215. Schmidt CM, Smolke CD. RNA switches for synthetic biology. Cold Spring Harb Perspect Biol. 2019;11:a032532.
216. Masaki K, Fujita Y, Saito H. MicroRNA-responsive ON-OFF hybrid mRNA switch for precise protein expression control. Mol Ther Nucleic Acids. 2025;36:102609.
217. Pardi N, Hogan MJ, Porter FW, Weissman D. mRNA vaccines: a new era in vaccinology. Nat Rev Drug Discov. 2018;17:261–79. https://doi.org/10.1038/nrd.2017.243.
218. Wolff JA, Malone RW, Williams P, Chong W, Acsadi G, Jani A, et al. Direct gene transfer into mouse muscle in vivo. Science. 1990;247:1465–8.
219. Bai H, Lester GMS, Petishnok LC, Dean DA. Cytoplasmic transport and nuclear import of plasmid DNA. Biosci Rep. 2017;37:BSR20160616.
220. Patel MN, Tiwari S, Wang Y, O'Neill S, Wu J, Omo-Lamai S, et al. Safer non-viral DNA delivery using lipid nanoparticles loaded with endogenous anti-inflammatory lipids. Nat Biotechnol. 2026;44:79–89.
221. Genin E, Hannequin D, Wallon D, Sleegers K, Hiltunen M, Combarros O, et al. APOE and Alzheimer disease: a major gene with semi-dominant inheritance. Mol Psychiatry. 2011;16:903–7.
222. Rosenberg JB, Kaplitt MG, De BP, Chen A, Flagiello T, Salami C, et al. AAVrh.10-mediated APOE2 central nervous system gene therapy for APOE4-associated Alzheimer's disease. Hum Gene Ther Clin Dev. 2018;29:24–47.
223. Mahley RW. Apolipoprotein E: remarkable protein sheds light on cardiovascular and neurological diseases. Clin Chem. 2017;63:14–20.
224. Corder EH, Saunders AM, Risch NJ, Strittmatter WJ, Schmechel DE, Gaskell PC, et al. Protective effect of apolipoprotein E type 2 allele for late onset Alzheimer disease. Nat Genet. 1994;7:180–4.
225. Dodart J-C, Marr RA, Koistinaho M, Gregersen BM, Malkani S, Verma IM, et al. Gene delivery of human apolipoprotein E alters brain Aβ burden in a mouse model of Alzheimer's disease. Proc Natl Acad Sci U S A. 2005;102:1211–6. https://doi.org/10.1073/pnas.0409072102.
226. Zhao L, Gottesdiener AJ, Parmar M, Li M, Kaminsky SM, Chiuchiolo MJ, et al. Intracerebral adeno-associated virus gene delivery of apolipoprotein E2 markedly reduces brain amyloid pathology in Alzheimer's disease mouse models. Neurobiol Aging. 2016;44:159–72.
227. Dos Santos Rodrigues B, Kanekiyo T, Singh J. ApoE-2 brain-targeted gene therapy through transferrin and penetratin tagged liposomal nanoparticles. Pharm Res. 2019;36:161.
228. Nafissi N, Alqawlaq S, Lee EA, Foldvari M, Spagnuolo PA, Slavcev RA. DNA ministrings: highly safe and effective gene delivery vectors. Mol Ther Nucleic Acids. 2014;3:e165.
229. Letort G, Duclert A, Le Clerre D, Chion-Sotinel I, Salvatori R, Dessez E, et al. Circular single stranded DNA potentiates non-viral gene insertion in hematopoietic stem and progenitor cells. Nat Commun. 2025;16:10125.
230. Chen Z-Y, He C-Y, Ehrhardt A, Kay MA. Minicircle DNA vectors devoid of bacterial DNA result in persistent and high-level transgene expression in vivo. Mol Ther. 2003;8:495–500.
231. Williams JA, Paez PA. Improving cell and gene therapy safety and performance using next-generation nanoplasmid vectors. Mol Ther Nucleic Acids. 2023;32:494–503.
232. Kay MA, He C-Y, Chen Z-Y. A robust system for production of minicircle DNA vectors. Nat Biotechnol. 2010;28:1287–9.
233. Luke J, Carnes AE, Hodgson CP, Williams JA. Improved antibiotic-free DNA vaccine vectors utilizing a novel RNA based plasmid selection system. Vaccine. 2009;27:6454–9.
234. Arévalo-Soliz LM, Hardee CL, Fogg JM, Corman NR, Noorbakhsh C, Zechiedrich L. Improving therapeutic potential of non-viral minimized DNA vectors. Cell Gene Ther Insights. 2020;6:1489–505.
235. Munye MM, Tagalakis AD, Barnes JL, Brown RE, McAnulty RJ, Howe SJ, et al. Minicircle DNA provides enhanced and prolonged transgene expression following airway gene transfer. Sci Rep. 2016;6:23125. https://doi.org/10.1038/srep23125.
236. Stenler S, Blomberg P, Smith CIE. Safety and efficacy of DNA vaccines: plasmids vs. minicircles. Hum Vaccin Immunother. 2014;10:1306–8.
237. Mitdank H, Tröger M, Sonntag A, Shirazi NA, Woith E, Fuchs H, et al. Suicide nanoplasmids coding for ribosome-inactivating proteins. Eur J Pharm Sci. 2022;170:106107.
238. Merting AD, Poschel DB, Lu C, Klement JD, Yang D, Li H, et al. Restoring FAS expression via lipid-encapsulated FAS DNA nanoparticle delivery is sufficient to suppress colon tumor growth in vivo. Cancers (Basel). 2022;14:361.
239. Saraiva C, Praça C, Ferreira R, Santos T, Ferreira L, Bernardino L. Nanoparticle-mediated brain drug delivery: overcoming blood-brain barrier to treat neurodegenerative diseases. J Control Release. 2016;235:34–47. https://doi.org/10.1016/j.jconrel.2016.05.044.
240. Nong J, Glassman PM, Shuvaev VV, Reyes-Esteves S, Descamps HC, Kiseleva RY, et al. Targeting lipid nanoparticles to the blood-brain barrier to ameliorate acute ischemic stroke. Mol Ther. 2024;32:1344–58. https://doi.org/10.1016/j.ymthe.2024.03.004.