UNIVERSITY OF CALIFORNIA SAN DIEGO

Engineering Multi-Lineage Vascularized Tissues from Human Pluripotent Stem Cells

A Dissertation submitted in partial satisfaction of the requirements
for the degree Doctor of Philosophy

in

Bioengineering

by

Amir Dailamy

Committee in charge:

Professor Prashant Mali, Chair
Professor Adam Engler
Professor Alysson Muotri
Professor Karl Willert
Professor Kun Zhang

2023

The Dissertation of Amir Dailamy is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2023

DEDICATION

This dissertation is dedicated to my mother, Hasina Dailamy, and father, Ramin Dailamy, who sacrificed everything to give me my opportunities. I am nothing without their love, teachings, and sacrifice. I owe it to them to move with gratitude and purpose every single minute of my life.

This dissertation is also dedicated to all the passionate scientists that sacrifice so much of their lives to contribute to the scientific literature. You all are the true explorers and innovators that will push us into brighter futures. I shed tears thinking about how beautiful and important you all are, and I genuinely love you for your efforts.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to acknowledge Professor Prashant Mali for his support, teachings, and patience. His visionary thinking, perseverance and robust experimental design is what I will carry with me throughout my entire scientific journey. I will always be proud to be his student, and I hope I can eventually be half as good as a scientist, leader, and visionary as he is.

I would like to acknowledge my mentee and main collaborator Hammza Khaliq. He taught me so much about mentorship. I am proud of his growth and see a bright scientific future for him. I admire his passion for science and appreciate his genuine desire to continually improve as a scientist. I hope he eventually takes me to platform 9 and ¾.

I would like to acknowledge Dr. Udit Parekh for being my first mentor in my PhD. His patience and breadth of knowledge is inspirational. Working with him is a joy and pleasure. He has also become something like an older brother to me, as I look to him for any kind of advice. I watch his actions closely, and frequently ask myself "What would Udit do?"

I would like to acknowledge Dr. Daniella McDonald for being my best grad school friend. She is an inspiration and a fighter. She is creative, resourceful and incredibly funny. I think of her as a true sister, and genuinely think we could be long-lost twins.

I would like to acknowledge Kian Kalhor for being another best grad school friend to me. I appreciate all the times that he tolerated my immense stress and anxiety and heard me out when I needed it. He was always there for me and made me feel like I had a true brother in San Diego that I could always count on. He is tremendously intelligent and scientifically curious, which would make our scientific discussions so memorable and fruitful. His calming energy was a great balancing force for me, and he always would remind me to take a step back and take care of my health. I am looking forward to future scientific collaborations and adventures with him.

I would like to acknowledge Dr. Dhruva Katrekar for being such a great friend and scientific collaborator every step of the way. I feel like we are also twins, as we both have an idealistic view of the world and constantly dream of making an impact through science. I love his passion, compassion, and creativity, and it has inspired me when times were very difficult.

I would like to acknowledge Dr. Aditya Kumar for being another great scientific inspiration and great friend to me. He was so meticulous and an amazing scientist. He also truly cared for me and listened to me when I really badly needed some support. I am very thankful that our paths crossed, and he will always be a brother to me.

I would like to acknowledge Dr. Ana Moreno for being another great friend and overall inspiration to me. Her ability to juggle being a mother, wife, scientist and CEO is herculean. I hope to be more like her in the future.

I would like to acknowledge Sammi Lyu for being a tremendous collaborator in my final year, in which she was able to master ADARs and RNA editing and help us whip up an entire manuscript in record time for me. I appreciate her resilience and adaptability throughout the project. She tolerated my computational ignorance and taught me basically everything I know about computational analysis. I am so happy our paths crossed, and I am a much better scientist thanks to her help and collaboration.

I would like to acknowledge Cansu Colbay for always being supportive friend and shining light of energy for me. She inspires me and made grad school exciting and fresh, even when tasks were mundane and routine. Throughout the twists and turns of our friendship, we still maintain a close bond, even from a distance. I couldn't of finished my Ph.D without her, and am excited to explore Indonesian Islands with her in the future.

I would like to acknowledge my brother and life-long best friend Soliemon Raza. He is such a loving, charismatic, and genuine person to everyone around him. He wants the best for his family, and the rest of society, and I love to learn from him. While growing up, I thought I would always be the older brother he looked up to, but throughout grad school, I looked up to him, and learned so much about life from him. He is the one person I know I will always be able to count on, and his brotherhood, friendship, and love is the purest connection I have to another person in this planet. Thank you to the goat Soliemon.

I would like to acknowledge my little sister Summer. Summer has always been a bright light for me. Someone that brings positivity, love, and understanding to every situation. I also thought I'd always be a big brother to her, that could help guide her throughout life, but in recent years, I choose to follow her lead in most situations, as she always seems to act in the purest way, only motivated by love and positivity. I love her so much for that, and I look forward to watching her more closely in the future to learn more from her. Thank you so much Summer jan.

I would like to acknowledge Dr. Daniel Kamei for being my first true scientific mentor. I gained so much experience and confidence through the time I spent in his lab, and without that confidence, I would've failed in grad school. He believed in me and pushed me to be my best. Without him, I would've never discovered my love for research.

I would like to acknowledge all other scientific collaborators and friends I made in grad school, especially Debbie Chen, Amanda Suhardjo, James Yen, Yichen Xiang, Sami Nourreddine, Joseph Rainaldi, Michael Tong and Paul Savarino.

Finally, I would like to acknowledge my family members Ramin Dailamy, Hasina Dailamy, Zuhal Raza, Pooyan "ExtraordinAIRe" Jamneshan, Alex Shapiro, Michael Nemeh, Rupert Gutierrez, Amo Assad, Amo Afshin, Idean, Nazanin, Khala Afu, Khala Atoo, Khala Efat,

Khala Roya, and Khala Getey. They are really the ones I will continue to acknowledge, and they are the true reason I can keep pushing and striving every day. They have been with me since I was one day old, and we will grow old together, supporting each other. You all are the real stars and have really allowed me to accomplish these big goals. I hope I can repay you all in the future.

Chapter 2, in full, is a reprint of the following published manuscript: Amir Dailamy*, Udit Parekh*, Dhruva Katrekar, Aditya Kumar, Daniella McDonald, Ana Moreno, Pegah Bagheri, Tse Nga Ng, Prashant Mali, "Programmatic introduction of parenchymal cell types into blood vessel organoids," ***Stem Cell Reports*** 16, pp. 2432-2441 (2021).. The dissertation author was the primary investigator and author of this paper.

VITA

2017 Bachelor of Science in Bioengineering, University of California Los Angeles

2023 Doctor of Philosophy in Bioengineering, University of California San Diego

PUBLICATIONS

1. Udit Parekh, Daniella McDonald, Amir Dailamy, Yan Wu, Thekla Cordes, Kun Zhang, Ann Tipps, Christian Metallo, Prashant Mali, "Charting oncogenicity of genes and variants across lineages via multiplexed screens in teratomas," *bioRxiv*, ***iScience*** 24, pp. 103149 (2021).
2. Amir Dailamy*, Udit Parekh*, Dhruva Katrekar, Aditya Kumar, Daniella McDonald, Ana Moreno, Pegah Bagheri, Tse Nga Ng, Prashant Mali, "Programmatic introduction of parenchymal cell types into blood vessel organoids," ***Stem Cell Reports*** 16, pp. 2432-2441 (2021).
3. Ana M. Moreno, Fernando Aleman, Glaucilene F. Catroli, Matthew Hunt, Michael Hu, Amir Dailamy, Andrew Pla, Sarah A. Woller, Nathan Palmer, Udit Parekh, Daniella McDonald, Amanda J. Roberts, Vanessa Goodwill, Ian Dryden, Robert F. Hevner, Lauriane Delay, Gilson Goncalves dos Santos, Tony Yaksh#, Prashant Mali#, "Long-lasting analgesia via targeted *in situ* repression of Nav1.7 in mice," *bioRxiv*, ***Science Translational Medicine*** 13, eaay9056 (2021).
4. Daniella McDonald*, Yan Wu*, Amir Dailamy, Justin Tat, Udit Parekh, Dongxin Zhao, Michael Hu, Ann Tipps, Kun Zhang#, Prashant Mali#, "Defining the teratoma as a model for multi-lineage human development," ***Cell*** 183, pp. 1402-1419 (2020).
5. Michael Hu, Amir Dailamy, Xin Lei, Udit Parekh, Daniella McDonald, Aditya Kumar, Prashant Mali, "Facile engineering of long-term culturable *ex vivo* vascularized tissues using biologically derived matrices," ***Advanced Healthcare Materials*** 7, pp. 1800845 (2018).

*Both of these authors contributed equally

PATENTS

1. Prashant Mali, Amir Dailamy, Udit Parekh, "Engineered Vascularized Organoids," ***US 20230174949*** (2023).
2. Prashant Mali, Andrew Portell, Amir Dailamy, Aditya Kumar, "Ribozyme-activated RNA constructs and uses thereof," ***WO 2022150773*** (2022).
3. Prashant Mali, Daniella McDonald, Kun Zhang, Yan Wu, Amir Dailamy, Udit Parekh, Michael Hu, "Novel method to engineer transplantable tissues," ***WO 20200010249*** (2020).
4. Prashant Mali, Udit Parekh, Amir Dailamy, Xin Lei, Michael Hu, "Use of 3D-printed freestanding structures for ex vivo tissues," ***WO 2019226710*** (2019).

FIELD OF STUDY

Major Field: Bioengineering

Studies in human pluripotent stem cell engineering, genetic engineering, high-throughput screening, next-generation sequencing

Professor Prashant Mali

ABSTRACT OF THE DISSERTATION

Engineering Multi-Lineage Vascularized Tissues from Human Pluripotent Stem Cells

by

Amir Dailamy

Doctor of Philosophy in Bioengineering

University of California San Diego, 2023

Professor Prashant Mali, Chair

Human pluripotent stem cells (hPSCs) serve as an excellent starting material for tissue engineering due to their renewability, capacity to differentiate into various tissue types, and remarkable self-organizing capabilities. However, a crucial challenge lies in establishing techniques for cultivating lab-grown organ constructs from hPSCs that accurately mimic the multi-

lineage and vascularized nature of authentic human organs. In my thesis, I concentrated on innovating methods to produce multi-lineage and vascularized tissue, employing both *in vitro* and *in vivo* tissue engineering strategies.

In my first aim, I introduce an *in vitro* engineering approach to construct multi-lineage vascularized human tissue. This involves merging cell reprogramming with chemically directed organoid differentiation. Unlike standard organoid protocols, which often result in single-lineage tissue, this method combines transcription factor overexpression and blood vessel organoid-directed differentiation within a single organoid construct. The overexpression of transcription factors serves as an intrinsic reprogramming signal, overriding extrinsic media cues. This synergy drives parenchymal cell fate within the vascular organoid, essentially allowing for the creation of multi-lineage vascularized tissue constructs. In conclusion, in my first aim I developed a modular tissue engineering platform which allows for the programmable introduction of parenchymal cell types into a vascular organoid scaffold, contributing to the bottom-up engineering of complete human organs *in vitro*.

In my second aim, I present an *in vivo* engineering approach for building multi-lineage, vascularized human tissue by molecularly sculpting teratomas. Teratomas are often used in stem cell research to examine the differentiation potential of hPSCs and are usually formed via the subcutaneous injection of hPSCs in immunodeficient mice. Considering teratomas are a multi-lineage and large-scale source of vascularized human tissue, I chose to leverage them as a launching pad for multi-lineage, vascularized tissue engineering. Although teratoma tissue contains a panoply of self-organized, fetal-like structures from all three germ layers, the inherent randomness of global differentiation patterns precludes its utility in tissue engineering. However, by implementing microRNA (miRNA)-mediated suicide gene circuits, in this aim I demonstrate a

unique teratoma sculpting approach, in which one can modularly select for desired tissue types *in vivo* based on the endogenous expression of a lineage-specific miRNA. All in all, this method of multi-lineage, vascular tissue engineering leverages the teratoma as powerful starting material for tissue sculpting and eventual top-down *in vivo* organ engineering.

In my third and final aim, I utilize the teratoma, along with time-course fetal-to-adult human organ tissue data, to conduct a biological investigation into the role of the RNA editing proteins adenosine deaminases acting on RNA (ADARs) in human development and cell-fate specification. ADARs impact diverse cellular processes and pathological conditions, and while we possess important insights into their roles in adult tissues, their functions in early cell fate specification remain less understood. To address this, we devised a comprehensive framework to investigate ADARs in human development, utilizing time-course human organ tissue, time-course teratoma tissue, and ADAR family CRISPR-KO screens in teratomas. After first establishing the teratoma as a faithful model of human development, we conducted pooled ADAR-KO (*ADAR1*-KO, *ADARB1*-KO, and *ADARB2*-KO) CRISPR-Cas9 screening in teratomas to uncover the functional roles of ADAR proteins in developing tissue, across all three germ layers. Via multi-lineage screening, we discovered that knocking out *ADAR1* led to a significant fitness defect in mesodermal tissues, and an enrichment of adipogenic cells, revealing a novel role for *ADAR1* in mesenchymal differentiation and obesity phenotypes. In brief, in my third and final arm, I demonstrate a modeling application for the multi-lineage and vascularized teratoma tissue, towards investigating the developmental role of ADARs and RNA editing across all germ layers.

Collectively, my thesis work seeks to develop novel techniques for generating multi-lineage vascularized lab-grown organ constructs and demonstrating their utility in biological investigations.

Chapter 1 - Introduction to Pluripotent Stem Cell Based Organ Engineering

The Imperative to Engineer Organ Tissues

In the pursuit of addressing critical challenges in regenerative medicine, this thesis endeavors to explore the development and study of multi-lineage and vascularized tissues derived from pluripotent stem cells.

The imperative to engineer functional organ tissues is propelled by a growing demand for transformative solutions that address the inherent limitations of traditional transplantation methods [1,2]. While conventional organ transplantation has undoubtedly been a life-saving medical intervention, its efficacy is impeded by the persistent shortage of donor organs. This scarcity manifests as prolonged waiting lists for transplantation candidates, amplifying the critical nature of the organ shortage crisis. Therefore, patients often endure prolonged suffering, and in tragic instances, succumb to their ailments before a suitable organ becomes available. The pressing need for a more sustainable and accessible source of transplantable organs has become increasingly evident, prompting the exploration of alternative avenues such as tissue engineering and regenerative medicine [8-12].

Moreover, the reliance on donor organs introduces a heightened risk of immune rejection, a critical concern that necessitates the administration of immunosuppressive drugs to transplant recipients [4]. These medications, while vital for preventing rejection, can have adverse side effects and compromise the recipient's overall health. Tissue engineering, particularly using pluripotent stem cells, offers a potential solution to these challenges by providing a scalable and customizable approach to generate organs that are immune-compatible with the recipient, reducing the risk of rejection and mitigating the dependence on a limited pool of donor organs [3,4,27]. As the imperative to engineer functional organ tissues intensifies, so does the momentum to innovate beyond the

constraints of traditional transplantation, marking a pivotal era in the quest for more accessible and effective solutions to address the organ shortage crisis.

Pluripotent Stem Cells

The advent of pluripotent stem cells marks a watershed moment in regenerative medicine, introducing an unprecedented capacity for self-renewal and differentiation that has revolutionized the landscape of tissue engineering [1-5]. Pluripotent stem cells, including embryonic stem cells and induced pluripotent stem cells, possess the unique ability to divide and produce both identical stem cells (self-renewal) and specialized cells of various lineages (differentiation). This remarkable versatility positions them as an invaluable resource for generating a diverse array of cell types, providing researchers with an unparalleled toolset to address the intricate complexities of organ tissue development.

The process of harnessing pluripotent stem cells for tissue engineering involves precisely guiding their differentiation along specific pathways to yield organ-specific tissues [2]. This entails mimicking the intricate cues present during natural embryonic development to coax these cells into becoming heart muscle cells, liver cells, neurons, and other specialized cell types [1]. Through a combination of carefully designed culture conditions and molecular signals, researchers can orchestrate the transformation of pluripotent stem cells into the desired cell lineages, laying the foundation for the creation of functional tissues. This approach not only holds immense promise for generating replacement tissues for transplantation but also provides a powerful platform for studying developmental biology, disease modeling, and drug screening in a controlled laboratory setting [3]. The convergence of pluripotent stem cell technology and tissue engineering methodologies exemplifies a groundbreaking synergy that propels the field towards novel therapeutic applications and a deeper understanding of complex biological processes [3-5].

Pluripotent stem cells, with their remarkable capacity to differentiate into various cell types, also play a pivotal role in disease modeling, offering a transformative approach to understanding complex diseases at the cellular and molecular levels. The use of pluripotent stem cells, including embryonic stem cells (ESCs) and induced pluripotent stem cells (iPSCs), enables researchers to recreate specific cell types affected by diseases, providing an unprecedented opportunity to study disease mechanisms, progression, and potential therapeutic interventions [1-5]. The implementation of pluripotent stem cells in disease modeling involves the directed differentiation of these cells into specific cell lineages affected by the targeted disease, allowing for the generation of relevant in vitro model systems.

One notable application of pluripotent stem cells in disease modeling is the study of neurodegenerative disorders, such as Alzheimer's and Parkinson's diseases [2,7,9]. By differentiating pluripotent stem cells into neurons or glial cells, researchers can recapitulate key aspects of neural development and study the pathological processes underlying these diseases. These model systems offer insights into the molecular and cellular changes associated with neurodegeneration, paving the way for the development of potential therapeutic strategies.

In the realm of cardiovascular diseases, pluripotent stem cells can be differentiated into cardiomyocytes, endothelial cells, and smooth muscle cells, allowing researchers to investigate the mechanisms underlying heart diseases like hypertrophic cardiomyopathy or atherosclerosis [2,10]. These stem cell-derived cardiac models offer a platform to explore disease-related genetic mutations, test drug responses, and develop personalized medicine approaches for patients with cardiovascular disorders.

The versatility of pluripotent stem cells is also harnessed in cancer research [5]. By introducing specific genetic mutations associated with various cancers into pluripotent stem cells

and differentiating them into relevant cell types, researchers can create in vitro models that closely mimic aspects of tumor development and progression. This approach facilitates the exploration of cancer biology, drug screening, and the identification of potential therapeutic targets.

Furthermore, pluripotent stem cells are instrumental in studying genetic disorders. For instance, iPSCs derived from individuals with genetic diseases can be differentiated into the affected cell types, allowing researchers to investigate the underlying genetic mechanisms and test potential therapeutic interventions. This personalized approach holds promise for advancing precision medicine and developing targeted therapies tailored to individual patients.

The use of pluripotent stem cells in disease modeling provides a powerful and versatile platform for studying a wide range of diseases. By recapitulating specific cell types affected by diseases in vitro, researchers can gain valuable insights into disease mechanisms, screen potential drugs, and develop personalized treatment strategies. The ongoing advancements in stem cell technology continue to enhance the fidelity and applicability of these model systems, making them indispensable tools for unraveling the complexities of various diseases and driving innovation in biomedical research.

Challenges in Pluripotent Stem Cell Based Organ Engineering

While the utilization of pluripotent stem cells for tissue engineering holds great promise, as detailed above, several persistent challenges underscore the complexity of recreating fully functional and mature organ tissues. One major hurdle lies in the incomplete maturation of stem cell-derived tissues. Despite advancements in directing pluripotent stem cells toward specific lineages, the resulting tissues often exhibit characteristics more reminiscent of embryonic or fetal stages rather than fully mature, adult structures [1,9,24]. Achieving the desired level of functional maturity is crucial for the success of engineered tissues, as immature cells may lack the full spectrum of physiological functions and could pose potential risks upon transplantation.

Vascularization poses another formidable obstacle in the development of robust and functional tissues [4]. The intricate network of blood vessels is essential for supplying nutrients and oxygen to cells, facilitating waste removal, and maintaining the overall health and viability of tissues. Constructing a vascular system within engineered tissues is challenging, as it requires mimicking the complexity of natural angiogenesis – the formation of blood vessels. The absence of a proper vascular network can lead to inadequate nutrient and oxygen diffusion, limiting the size and viability of the engineered tissue. Overcoming this hurdle is imperative for the successful translation of stem cell-derived tissues into clinically relevant therapies.

Another prominent limitation in the current landscape of stem cell-derived organ constructs is the challenge in building multi-lineage tissues that accurately represent the diverse cell types found in real human organs [1-4,12]. Most stem cell-derived organoids or tissues tend to originate from a single germ layer, typically endoderm, ectoderm, or mesoderm, and thus lack the intricate cellular diversity inherent in natural human organs that arise from multiple germ layers during embryonic development. This limitation is particularly pronounced when attempting to replicate the complexity of organs with intricate cellular interactions and functions, such as the liver, heart, or kidney.

The consequence of this limitation is that the resulting stem cell-derived constructs may not fully capture the heterogeneous cellular composition and physiological nuances of native tissues. In real human organs, various cell types collaborate to form intricate structures and perform specialized functions, and the absence of this comprehensive cellular diversity in stem cell-derived constructs may impact their functionality and relevance for studying certain diseases or testing therapeutic interventions. Addressing this limitation is crucial for advancing the field of regenerative medicine and tissue engineering, as achieving a more faithful representation of natural

tissue complexity is essential for improving the translational potential of stem cell-based therapies and models. Researchers are actively exploring strategies to coax stem cells into differentiating into multiple cell lineages within a single construct, aiming to create more physiologically relevant and multi-lineage tissues for a wide range of applications, from disease modeling to regenerative medicine.

While pluripotent stem cells have provided an unprecedented platform for tissue engineering, addressing the challenges of incomplete maturation, vascularization, and microenvironmental mimicry is pivotal for realizing the full potential of engineered tissues. Researchers are actively exploring innovative approaches, such as the incorporation of biomimetic scaffolds, advanced bioreactor systems, and sophisticated bioengineering techniques, to overcome these limitations. A holistic understanding of the intricate interplay between cells and their microenvironment is essential to guide the development of next-generation tissue engineering strategies that can bridge the gap between experimental success and clinical applicability.

Organoids: A Step in the Right Direction

The development of organoid technology represents a significant leap forward in the field of biological research and therapeutic development. Organoids are three-dimensional miniaturized versions of organs or tissues grown in vitro, derived from pluripotent stem cells or adult tissue-specific stem cells [1-4]. The motivation behind the development of organoids stems from the inherent limitations of traditional two-dimensional cell cultures. Classic 2D model systems often fail to recapitulate the complex architecture, cellular interactions, and physiological functions of native tissues, leading to a gap in our understanding of various diseases and hampering the development of effective treatments.

Organoids address this limitation by providing a more sophisticated and representative model of organ development and function. Their three-dimensional structure mirrors the in vivo

conditions more closely, allowing for the recreation of the intricate cellular composition and spatial organization found in natural organs. This architectural fidelity is particularly crucial for studying diseases and drug responses, as it enables researchers to observe how cells within an organoid interact, differentiate, and respond to therapeutic interventions in a more physiologically relevant environment. Consequently, the development of organoids has become instrumental in advancing our understanding of complex biological processes, offering a bridge between traditional 2D cell cultures and in vivo models.

The advantages of organoids over 2D model systems extend beyond their structural mimicry to include improved functionality and the ability to capture disease phenotypes more accurately. Organoids can replicate key features of organ physiology, such as the formation of lumens, cell polarization, and tissue-specific functions [6,12,25]. This enhanced functionality is particularly valuable in studying organ-specific diseases and identifying potential drug targets. Additionally, the use of patient-derived organoids allows for personalized medicine approaches, where researchers can test drug responses on a patient's own cells, paving the way for more effective and tailored therapeutic interventions. In essence, the development of organoid technology, represents a paradigm shift in biological research, offering a more holistic and clinically relevant approach to studying organ development, diseases, and drug responses.

Addressing Limitations in 3D, Pluripotent Stem Cell Organ Engineering

Despite the significant advancements in organoid technology, several limitations hinder its broader application and representation of complex biological systems. One notable constraint lies in the difficulty of deriving multi-lineage organoids [1-3]. While organoids excel in capturing the cellular diversity of specific tissues, the simultaneous generation of multiple cell types within a single organoid remains a considerable challenge. Achieving a harmonious integration of diverse cell lineages is crucial for mimicking the complexity of native organs, and current techniques often

fall short in reproducing such intricate cellular compositions. The inability to easily derive multi-lineage organoids limits their capacity to fully replicate the structural and functional heterogeneity found in vivo, posing a challenge for studying tissues with intricate cellular interplay.

Another critical limitation in organoid technology is the complex task of vascularization [4,8,24]. Efficient vascular networks are essential for supplying nutrients and oxygen, removing waste products, and maintaining the overall health and functionality of tissues. While researchers have made strides in generating vascularized organoids, achieving a fully functional and integrated vascular system within these miniature organs remains a formidable obstacle. Mimicking the intricate process of angiogenesis in vitro, where blood vessels form and connect to ensure proper tissue perfusion, remains a complex puzzle. The absence of a robust vascular component in organoids limits their growth size, hampers long-term viability, and constrains their potential for therapeutic applications, particularly in the context of transplantation or modeling diseases involving vascular dysfunction. Overcoming these limitations in the development of multi-lineage and vascularized organoids represents a critical frontier in advancing organoid technology toward more comprehensive and physiologically relevant representations of human tissues.

Thus, in my thesis, I focus on developing 2 novel approaches towards building multi-lineage and vascularized tissues. The first approach (Chapter 2) is an in vitro method, which involves the invention of a novel organoid engineering technique, while the second (Chapter 3) is an in vivo approach, which leverages the multi-lineage and vascularized nature of in vivo derived teratomas, but overlays engineering principles onto the teratoma system. Like organoids, teratomas contain the structures of self-organization of tissue-types from all three germ layers, thus they serve as another powerful starting point for organ engineering. Organoids and teratomas are both powerful systems that reflect the remarkable self-organizing capabilities of stem cells, however

there is a lack of engineering in these systems. Thus, these are two orthogonal approaches that involve genetic engineering of pluripotent stem cells, towards organ engineering of multi-lineage and vascularized tissues. Through implementing clever engineering in the context of self-organizing systems, we can leverage the best of both biology and engineering, towards the formation of true human organs. And finally, in my final aim (Chapter 4), I demonstrate a modelling application of these multi-lineage systems, specifically by investigating the developmental role of the RNA editing proteins adenosine deaminase acting on RNA (ADARs) by charting and probing their activity in teratoma tissue. As a whole, my thesis is an exercise in organ engineering technology development, and a demonstration of the modelling application of a multi-lineage, tissue-engineered system.

Chapter 2 - Programmatic Introduction of Parenchymal Cell Types into Blood Vessel Organoids.

Abstract

Pluripotent stem cell derived organoids have transformed our ability to recreate complex three-dimensional models of human tissue. However, the directed differentiation methods used to create them do not afford the ability to introduce cross-germ layer cell types. Here, we present a bottom-up engineering approach to building vascularized human tissue by combining genetic reprogramming with chemically directed organoid differentiation. As a proof of concept, we created neuro-vascular and myo-vascular organoids via transcription factor overexpression in vascular organoids. We comprehensively characterized neuro-vascular organoids in terms of marker gene expression and composition and demonstrated that the organoids maintain neural and vascular function for at least 45 days in culture. Finally, we demonstrated chronic electrical stimulation of myo-vascular organoid aggregates as a potential path toward engineering mature and large scale vascularized, skeletal muscle tissue from organoids. Our approach offers a roadmap to build diverse vascularized tissues of any type entirely from pluripotent stem cells.

Introduction

The ability to recapitulate organogenesis and create complex human tissue *in vitro* has been a long-standing goal for the stem cell and tissue engineering field. The advent of organoid technology has recently made it possible to create 3D, self-organized, pluripotent stem cell (PSC) derived tissues for *in vitro* developmental and disease modeling which closely mimic the cellular, spatial and molecular architecture of endogenous human tissue [1,2]. These advances have enabled substantial progress in building fully PSC-derived, functional human organs *in vitro* [3]. However, the absence of vasculature in most organoid techniques limits their utility, principally in two ways. Firstly, it is widely accepted that vasculature plays a crucial role in development and disease [4,5].

Secondly, vasculature is necessary to prevent necrosis in tissues that grow beyond 1 mm in size [6,7], which deems vasculature critical for building large-scale tissue models.

To address this, several groups have demonstrated progress on developing methods for vascularizing organoids [8–14]. Some groups have succeeded in vascularizing organoids after transplanting them *in vivo* [13,14], but requiring an *in vivo* host limits experimental control, increases cost, and diminishes its potential for clinical applications. Others transiently overexpressed *GATA6* to introduce a nascent vascular network in liver-bud organoids [8], however there is no current evidence this system can be translated to other organoids.

Organoid platforms leverage knowledge of development to provide temporally appropriate chemical cues to self-assembled PSC-derived embryoid bodies, modulating key organogenesis-specific signaling pathways to drive the directed differentiation of organ-specific cells in physiologically faithful architectures. Thus, cells which do not belong to those specific organ compartments or arise from different germ layers are absent from the final organoid. To introduce cell types outside of those available from directed differentiation, a promising strategy is to combine genetic overexpression with directed differentiation. Recently, vascularization of cerebral organoids has been reported, in which PSCs were engineered to ectopically express human ETS variant 2 (*ETV2*) [9], a known driver of differentiation to endothelial cells from PSCs [15,16]. Although this method has exciting potential, it suffers from two limitations: one, it induces only a low degree of vascularization and two, it does not induce the full panoply of vascular lineages, such as smooth muscle cells (SMCs) and mesenchymal stem cells (MSCs), which are critical for blood vessel development and function [17]. Thus, there is a need to explore alternative methods.

A recently described vascular organoid (VO) differentiation approach [18] yields complete blood vessel networks including SMCs, MSCs, and endothelial cells, but these organoids lack

organ specific parenchymal cells limiting their utility for broader disease modelling and regenerative medicine. Here, we overlay reprogramming with the directed differentiation of VOs to derive cross-germ layer and cross-lineage organoids with complete vascular networks. We introduce a parenchymal cell component into VOs via transcription factor (TF) overexpression and demonstrate this approach by building neuro-vascular and myo-vascular organoids. This is done by the induced overexpression, in developing vascular organoids, of *NEUROD1* (iN) to form neuro-vascular organoids (iN-VOs), and the induced overexpression of *MYOD1* plus *BAF60C* (iMB) to form myo-vascular organoids (iMB-VOs).

This yields a facile method for co-differentiation of tissue-specific parenchymal cells and the entire blood vessel lineage from a single PSC line. We present this approach as a proof-of-concept for the introduction of other parenchymal cell types, via the overexpression of lineage-specifying TFs, in the context of a vascular organoid scaffold.

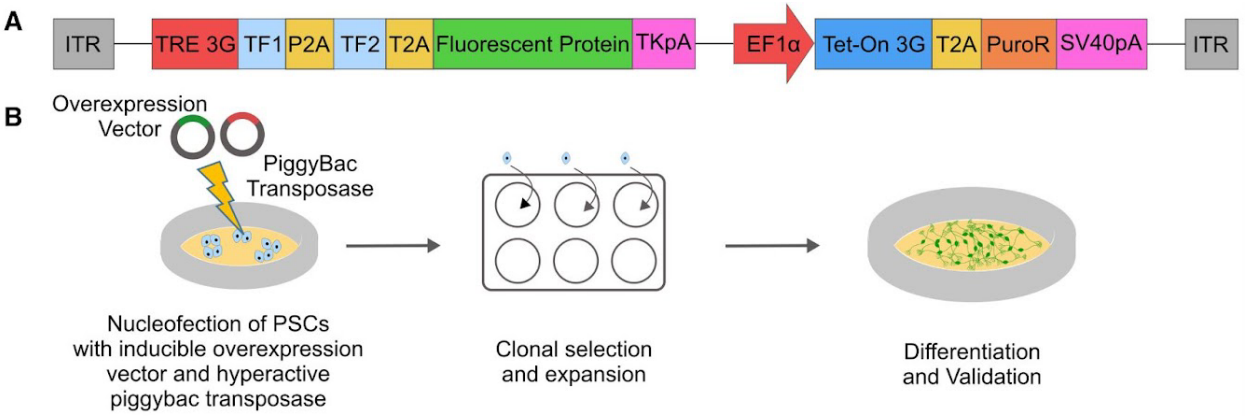
Results

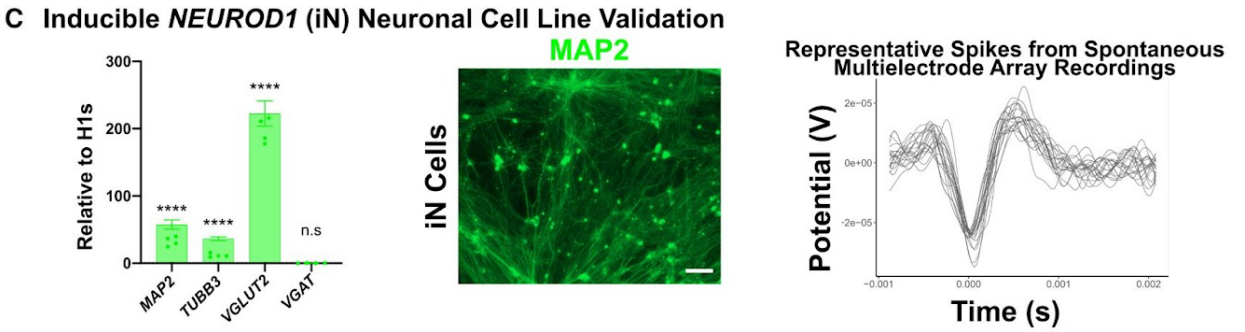
Inducible cell line construction and validation

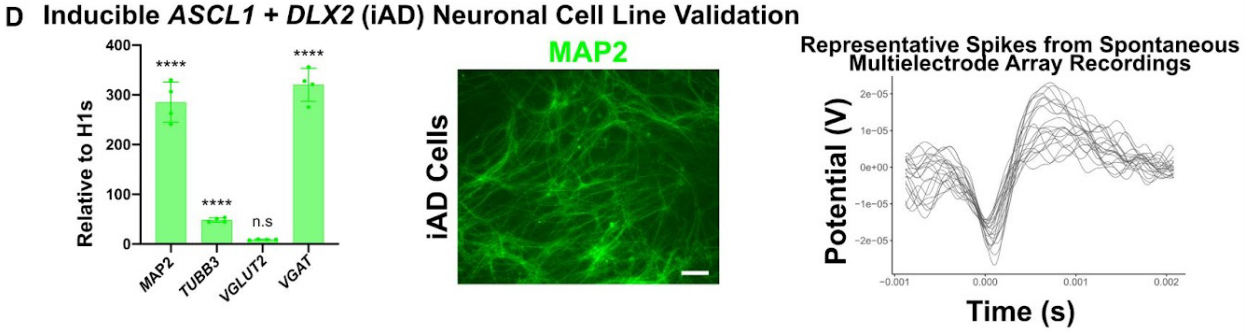
For the inducible expression of TFs, we designed a PiggyBac transposon-based overexpression vector which allowed a single vector to package the complete Tet-On system for doxycycline inducible expression, along with one or more TFs to be overexpressed in conjunction with a reporter fluorescent protein (**Figure 2.1a**). To establish the utility of our PiggyBac overexpression platform, we constructed an array of cell lines and validated their ability to differentiate into functional tissue upon adding doxycycline. For neural differentiation, we chose to overexpress *NEUROD1* to generate glutamatergic excitatory neurons [16,19], and *ASCL1+DLX2* for GABAergic inhibitory neuron differentiation [20]. For mesodermal tissue differentiation, we chose to overexpress *MYOD1+BAF60C* [21] for skeletal muscle differentiation.

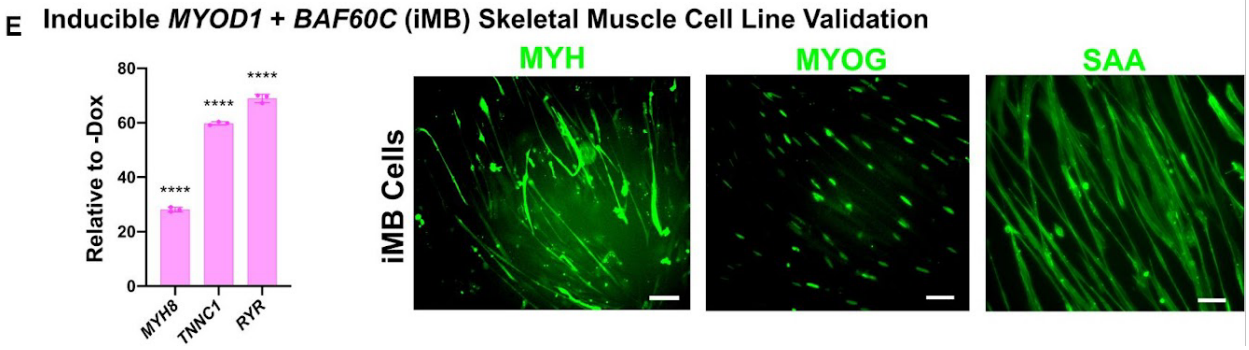
Figure 2.1: Construction and characterization of inducible cell lines.

(A) Schematic of PiggyBac transposon based inducible overexpression vector. (B) Schematic of cell line generation and validation process. (C) Inducible NEUROD1 (iN) cell line validation at 3 weeks post induction via: 1) qRT-PCR analysis of signature neuronal markers MAP2, TUBB3, VGLUT2 and VGAT; Data represent the mean ± s.d. (n = 4 independent experiments) 2) immunofluorescence micrograph of MAP2+ cells (Scale bars = 50 μm); and 3) representative spike plots from MEA measurements of spontaneously firing iN cells. (D) Inducible ASCL1+DLX2 (iAD) cell line validation at 3 weeks post induction via: 1) qRT-PCR analysis of signature neuronal markers MAP2, TUBB3, VGLUT2 and VGAT; Data represent the mean ± s.d. (n = 4 independent experiments) 2) Immunofluorescence micrograph of MAP2+ cells (Scale bars = 50 μm); and 3) representative spike plots from MEA measurements of spontaneously firing iAD cells. (E) Inducible MYOD1+BAF60C (iMB) cell line validation at 2 weeks post induction via: 1) qRT-PCR analysis of signature skeletal muscle markers MYH8, TNNC1, and RYR; Data represent the mean ± s.d. (n = 3 independent experiments) and 2) Immunofluorescence micrograph of MYH+, MYOG+, and SAA+ labelled cells (Scale bars = 50 μm). (c,d,e) (**P ≤ 0.01, ***P ≤ 0.001, and ****P ≤ 0.0001; ns = not significant).


A
ITR
TRE 3G
TF1
P2A
TF2
T2A
Fluorescent Protein
TKpA
EF1α
Tet-On 3G
T2A
PuroR
SV40pA
ITR
B
Overexpression Vector
PiggyBac Transposase
Nucleofection of PSCs with inducible overexpression vector and hyperactive piggybac transposase
Clonal selection and expansion
Differentiation and Validation


C Inducible NEUROD1 (iN) Neuronal Cell Line Validation
Relative to H1s
MAP2
TUBB3
VGLUT2
VGAT
n.s
MAP2
iN Cells
Representative Spikes from Spontaneous Multielectrode Array Recordings
Potential (V)
Time (s)


D Inducible ASCL1 + DLX2 (iAD) Neuronal Cell Line Validation
Relative to H1s
MAP2
TUBB3
VGLUT2
VGAT
n.s
MAP2
iAD Cells
Representative Spikes from Spontaneous Multielectrode Array Recordings
Potential (V)
Time (s)


E Inducible MYOD1 + BAF60C (iMB) Skeletal Muscle Cell Line Validation
Relative to -Dox
MYH8
TNNC1
RYR
MYH
MYOG
SAA
iMB Cells

Stable, dox-inducible *NEUROD1* (iN), *ASCL1+DLX2* (iAD), and *MYOD1+BAF60C* (iMB) human embryonic stem cell (hESC) lines were generated by nucleofecting hESCs with the respective overexpression vector, along with a hyperactive PiggyBac transposase [22].

For validation of neural differentiation of iN and iAD lines, qRT-PCR analysis demonstrated upregulation of neuronal markers *MAP2* and *TUBB3* for both iN and iAD cells (**Figure 2.1c, 2.1d**) compared to undifferentiated hESCs. iN neurons were confirmed to be glutamatergic via upregulation of excitatory marker *VGLUT2* with no detectable expression of the inhibitory marker *VGAT* (**Figure 2.1c**), while iAD neurons were confirmed to be GABAergic, via upregulation of *VGAT* with minimal expression of *VGLUT2* (**Figure 2.1d**). *MAP2*+ cells with classic neuronal morphology were confirmed by immunofluorescence for both iN and iAD cells (**Figure 2.1c, 2.1d**). Finally, iN and iAD neurons were confirmed to be functional by detection of spontaneous firing when co-cultured with glia and measured on a microelectrode array (MEA) at 3-5 weeks post-induction (**Figure 2.1c, 2.1d**).

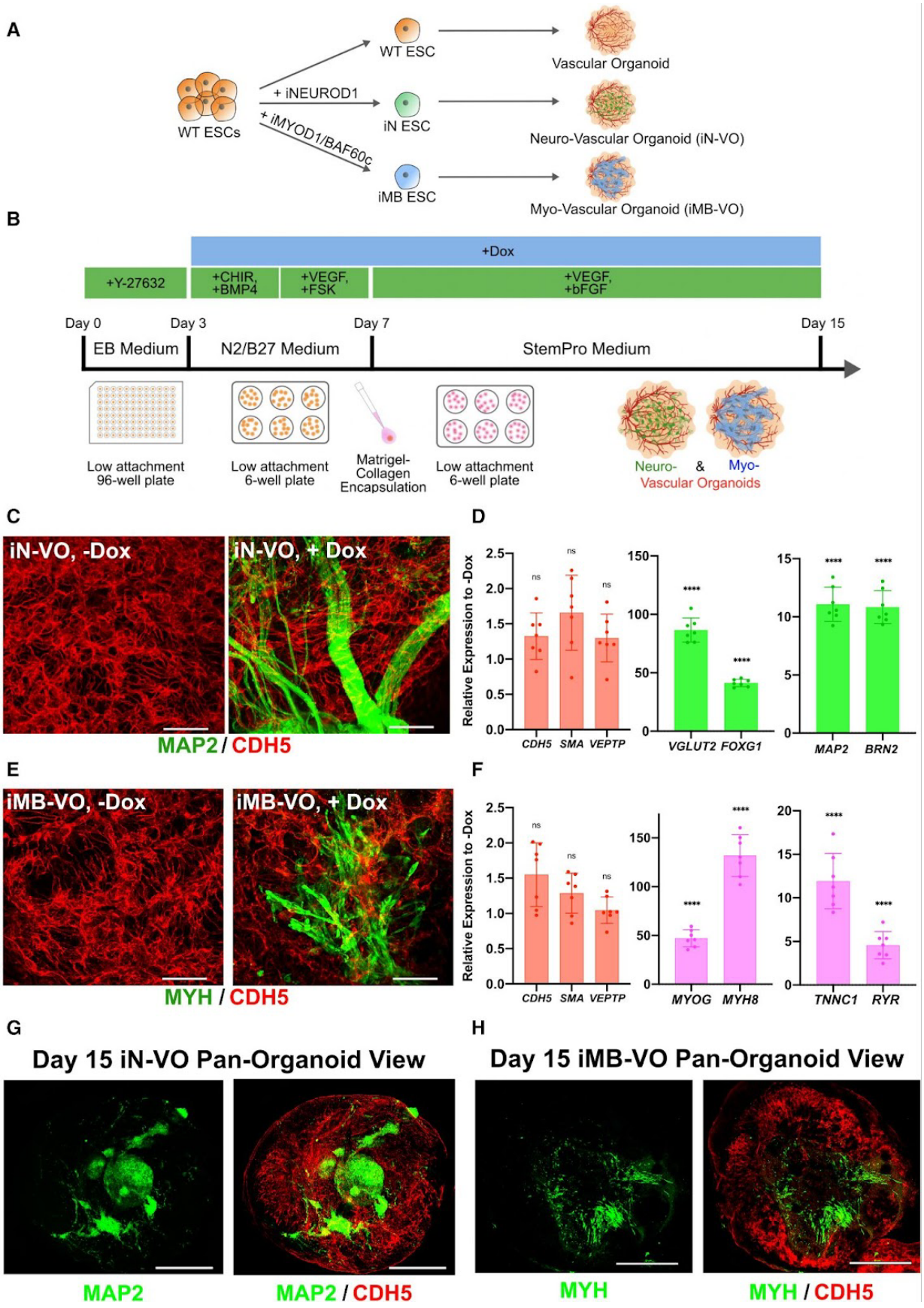
Skeletal muscle differentiation of iMB cells was validated by gene expression and immunofluorescence assays (**Figure 2.1e**). qRT-PCR analysis confirmed upregulation of skeletal muscle markers *MYH8*, *TNNC1* and *RYR*. Immunofluorescence confirmed the presence of characteristic spindle shaped, *MYH*+ and *SAA*+ skeletal muscle morphology, along with *MYOG*+ nuclei.

Constructing neuro-vascular and myo-vascular organoids

After clonal hESC lines were established, we developed a modular, generalizable, TF overexpression-based strategy to differentiate parenchymal cell types in VOs (**Figure 2.2a**). We conducted an optimization experiment to determine if Wild-Type ESCs are required for maintaining the vascular compartment in our organoids (**Figure 2.3a**). Given the higher expression

Figure 2.2: Generation of iN-VOs and iMB-VOs.

(a) General strategy for the generation of vascularized organ tissues via introduction of parenchymal cell types in VOs. **(b)** Schematic of iN-VO and iMB-VO culture protocol. **(c)** Immunofluorescence 100 μm z-stack, maximum projection, confocal micrographs of *MAP2*- and *CDH5*-labelled uninduced (iN-VO, -Dox) and induced (iN-VO, +Dox) day 15 iN-VO organoids (Scale bars = 50 μm). **(d)** qRT-PCR analysis of signature endothelial genes *CDH5* and *VEPTP*; signature smooth muscle gene *SMA*; and signature neuronal genes *MAP2*, *VGLUT2*, *BRN2* and *FOXG1* at day 15 of culture for iN-VO organoids. Data represent the mean ± s.d. (n = 7 organoids, from three independent experiments) **(e)** Immunofluorescence 100 μm z-stack, maximum projection, confocal micrographs of *MYH*- and *CDH5*-labelled uninduced (iMB-VO, -Dox) and induced (iMB-VO, +Dox) day 15 iMB-VO organoids (Scale bars = 50 μm). **(f)** qRT-PCR analysis of signature endothelial genes *CDH5* and *VEPTP*; signature smooth muscle gene *SMA*; and signature skeletal muscle genes *MYOG*, *MYH8*, *TNNC1* and *RYR* at day 15 of culture for iMB-VO organoids. Data represent the mean ± s.d. (n = 7 organoids, from three independent experiments). **(g)** Pan-organoid tile-scan immunofluorescence confocal micrograph of a *CDH5*- and *MAP2*-labelled day 15 neuro-vascular organoid (Scale bars = 500 μm). Image is a 200 μm z-stack maximum intensity projection. **(h)** Pan-organoid tile-scan immunofluorescence confocal micrograph of *CDH5*- and *MYH*-labelled day 15 myo-vascular organoid (Scale bars = 500 μm). Image is a 200 μm z-stack maximum intensity projection. **(d,f)** (***P* ≤ 0.01, ****P* ≤ 0.001, and *****P* ≤ 0.0001; ns = not significant).


A
WT ESCs
+ iNEUROD1
+ iMYOD1/BAF60c
WT ESC
iN ESC
iMB ESC
Vascular Organoid
Neuro-Vascular Organoid (iN-VO)
Myo-Vascular Organoid (iMB-VO)
B
+Dox
+Y-27632
+CHIR, +BMP4
+VEGF, +FSK
+VEGF, +bFGF
Day 0
Day 3
Day 7
Day 15
EB Medium
N2/B27 Medium
StemPro Medium
Low attachment 96-well plate
Low attachment 6-well plate
Matrigel-Collagen Encapsulation
Low attachment 6-well plate
Neuro- & Myo-Vascular Organoids
C
iN-VO, -Dox
iN-VO, + Dox
MAP2 / CDH5
D
Relative Expression to -Dox
CDH5
SMA
VEPTP
VGLUT2
FOXG1
MAP2
BRN2
ns
****
E
iMB-VO, -Dox
iMB-VO, + Dox
MYH / CDH5
F
Relative Expression to -Dox
CDH5
SMA
VEPTP
MYOG
MYH8
TNNC1
RYR
G
Day 15 iN-VO Pan-Organoid View
MAP2
MAP2 / CDH5
H
Day 15 iMB-VO Pan-Organoid View
MYH
MYH / CDH5

of *MAP2* and comparable expression of *CDH5*, subsequent experiments were conducted with iN-VOs formed from 100% iN cells (**Figure 2.3b**). Immunostaining was performed on both induced (+dox, iN-VOs) and uninduced (-dox, iN-VOs) organoids confirming *MAP2*+ cells forming into bundle-like structures of neurons only in induced organoids, along with dense, interpenetrating *CDH5*+ vascular networks in both conditions (**Figure 2.2c**). Furthermore, pan-organoid confocal images show that endothelial networks span all layers of the organoid, as do neurons although they are not uniformly distributed and form clusters and bundles of cells (**Figure 2.2g**). Further analysis via qRT-PCR showed an upregulation of the excitatory marker *VGLUT2,* along with cortical neuron markers *BRN2* and *FOXG1* [19] in day 15 iN-VOs (**Figure 2.2d**). Moreover, we observed expression levels comparable to uninduced organoids of the endothelial markers *CDH5* and *VEPTP*, along with the smooth muscle marker *SMA* (**Figure 2.2d**).

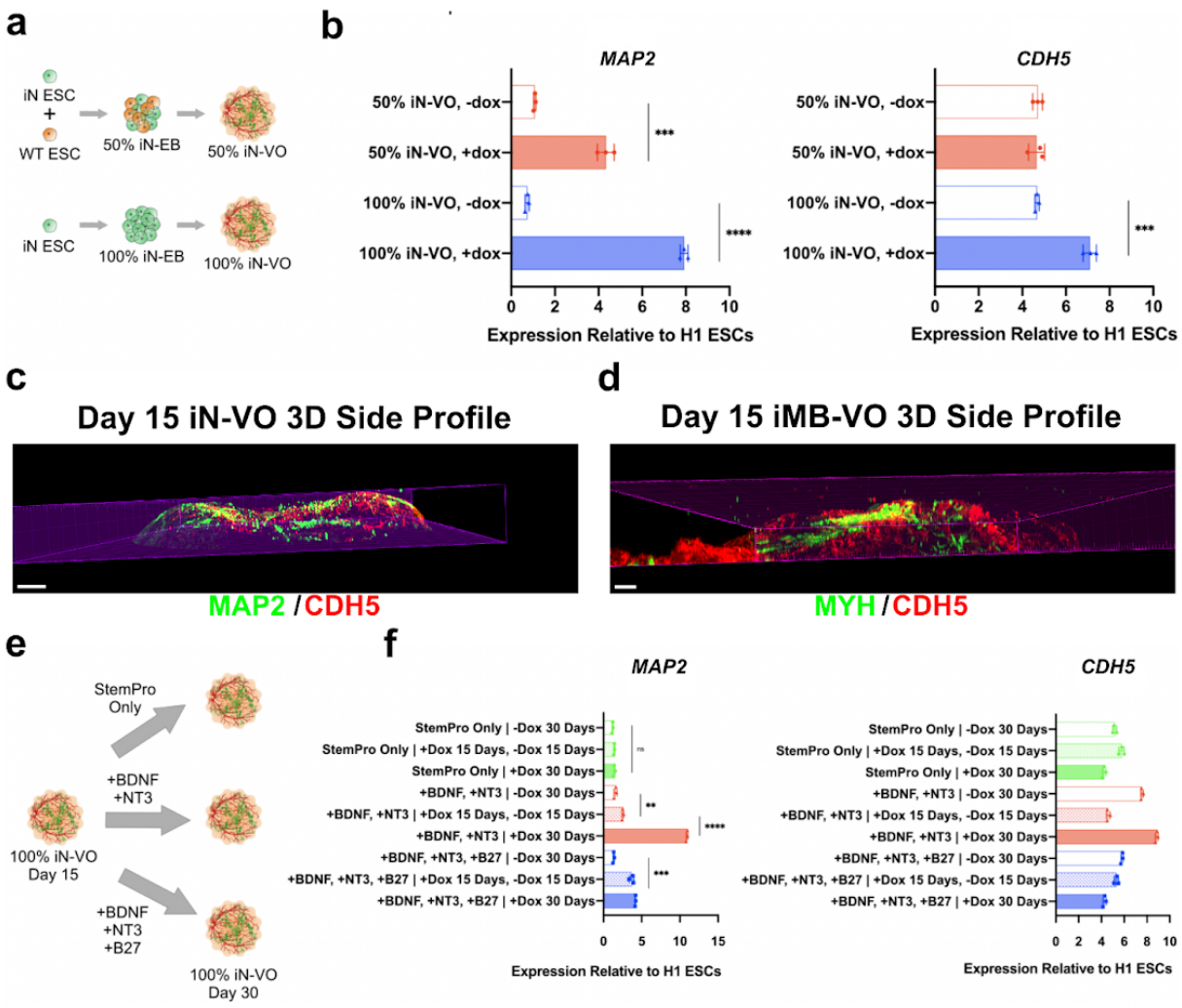
Once the formation of neuro-vascular organoids was established, we then sought to demonstrate if our system was compatible with other germ-layer or cross-lineage cell types. Using 100% iMB cells and the same protocol used to generate neuro-vascular organoids, we grew induced myo-vascular organoids (iMB-VOs) (**Figure 2.2b**). iMB-VOs had *MYH*+ spindle-shaped cells (**Figure 2.2e**), along with expression of skeletal muscle marker genes *MYOG*, *MYH8*, *TNNC1* and *RYR* (**Figure 2.2f**), confirming that our platform could also be used to grow vascularized tissue of mesodermal origin, specifically skeletal muscle. Pan-organoid tile-scan confocal images demonstrated *MYH*+ skeletal muscle cells present in many layers of the myo-vascular organoid, albeit not uniformly distributed throughout the organoid (**Figure 2.2h, 2.3d**).

Comprehensive characterization of neuro-vascular organoids

We then sought to enable further long-term culture of the neuro-vascular organoids beyond the initial 15 days (**Figure 2.3e**). Media optimization experiments revealed that BDNF and NT-3

Figure 2.3: Optimization of organoid culture conditions.

(a) Schematic for experiment to assess the optimal ratio of iN to WT cells for iN-VO formation. **(b)** qRT-PCR analysis of signature neuronal gene *MAP2*, and signature endothelial gene *CDH5*, at day 15 of culture for organoids grown from 50% and 100% iN cells. Data represent the mean ± s.d. (n = 3 organoids) and the unpaired two-tailed t-test was used for all comparisons. **(c)** Side profile 3D rendering of a pan-organoid tile-scan, z-stack immunofluorescence confocal micrograph of a *CDH5*- and *MAP2*-labelled day 15 neuro-vascular organoid (Scale bars = 150 μm). **(d)** Side profile 3D rendering of a pan-organoid tile-scan, z-stack immunofluorescence confocal micrograph of *CDH5*- and *MYH*-labelled day 15 myo-vascular organoid (Scale bars = 150 μm).**(e)** Schematic for experiment to assess optimal media supplements for long term iN-VO culture. **(f)** qRT-PCR analysis of signature neuronal gene *MAP2*, and signature endothelial gene *CDH5*, at day 30 of culture to assess long term neuronal and endothelial survival. Data represent the mean ± s.d. (n = 3 organoids) and the unpaired two-tailed t-test was used for all comparisons. **(b,f)** Statistical significance was attributed to $P < 0.05$ as determined by unpaired two-tailed t-test comparisons. (**$P \leq 0.01$, ***$P \leq 0.001$, and ****$P \leq 0.0001$; ns = not significant).


a
iN ESC
+
WT ESC
50% iN-EB
50% iN-VO
iN ESC
100% iN-EB
100% iN-VO
b
MAP2
50% iN-VO, -dox
50% iN-VO, +dox
100% iN-VO, -dox
100% iN-VO, +dox
***
****
0 2 4 6 8 10
Expression Relative to H1 ESCs
CDH5
50% iN-VO, -dox
50% iN-VO, +dox
100% iN-VO, -dox
100% iN-VO, +dox
***
0 2 4 6 8 10
Expression Relative to H1 ESCs
c
Day 15 iN-VO 3D Side Profile
MAP2 / CDH5
d
Day 15 iMB-VO 3D Side Profile
MYH / CDH5
e
StemPro Only
+BDNF +NT3
100% iN-VO Day 15
+BDNF +NT3 +B27
100% iN-VO Day 30
f
MAP2
StemPro Only | -Dox 30 Days
StemPro Only | +Dox 15 Days, -Dox 15 Days
StemPro Only | +Dox 30 Days
+BDNF, +NT3 | -Dox 30 Days
+BDNF, +NT3 | +Dox 15 Days, -Dox 15 Days
+BDNF, +NT3 | +Dox 30 Days
+BDNF, +NT3, +B27 | -Dox 30 Days
+BDNF, +NT3, +B27 | +Dox 15 Days, -Dox 15 Days
+BDNF, +NT3, +B27 | +Dox 30 Days
ns
**
****
***
0 5 10 15
Expression Relative to H1 ESCs
CDH5
StemPro Only | -Dox 30 Days
StemPro Only | +Dox 15 Days, -Dox 15 Days
StemPro Only | +Dox 30 Days
+BDNF, +NT3 | -Dox 30 Days
+BDNF, +NT3 | +Dox 15 Days, -Dox 15 Days
+BDNF, +NT3 | +Dox 30 Days
+BDNF, +NT3, +B27 | -Dox 30 Days
+BDNF, +NT3, +B27 | +Dox 15 Days, -Dox 15 Days
+BDNF, +NT3, +B27 | +Dox 30 Days
0 2 4 6 8 10
Expression Relative to H1 ESCs

supplementation from day 15 ensured long term survival of neurons with reliable maintenance of vascular lineages (**Figure 2.3f**).

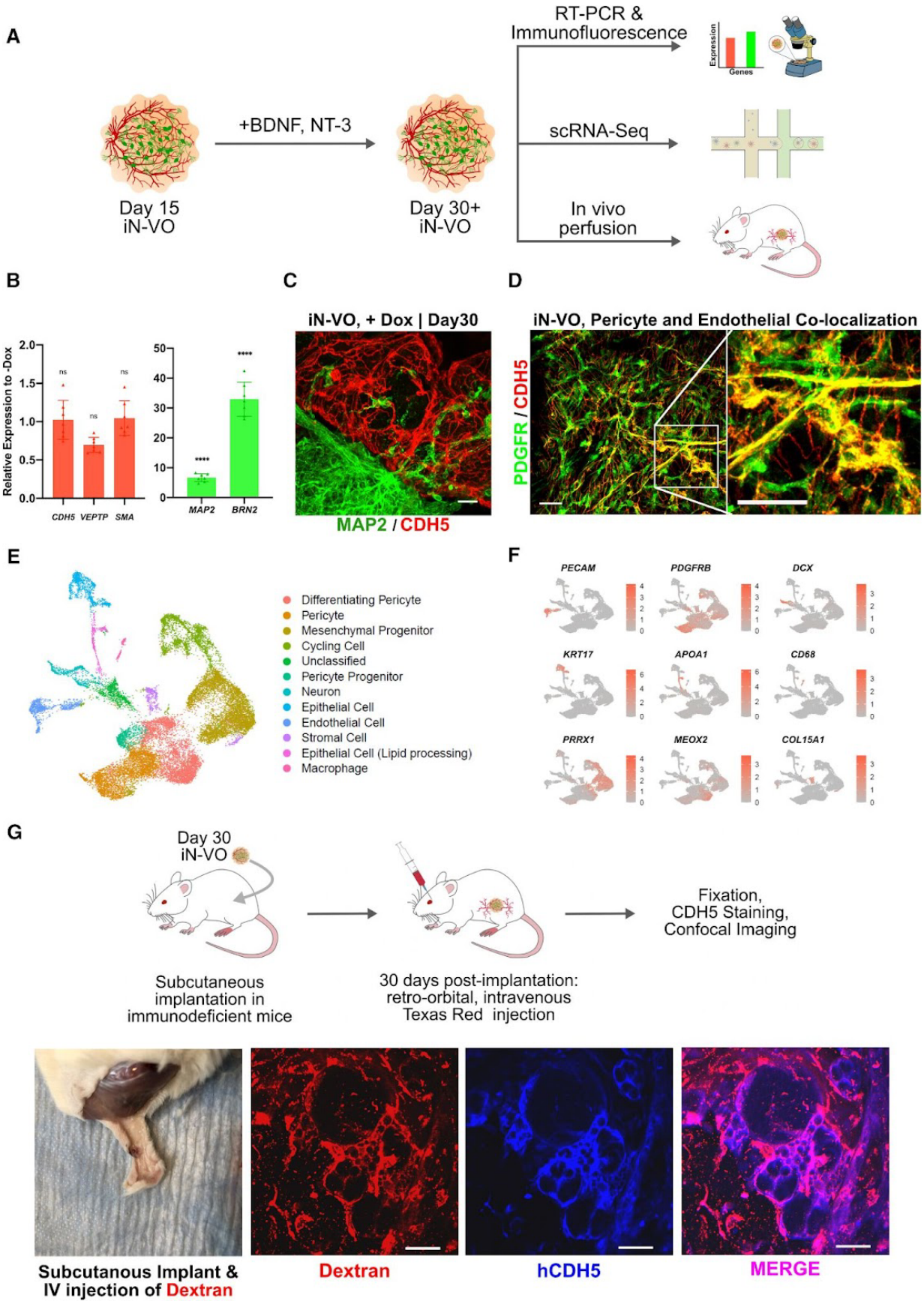
Upon determining an optimized protocol for long-term growth of iN-VOs, we comprehensively characterized these long-term cultured neuro-vascular organoids (**Figure 2.4a**). qRT-PCR analysis showed an upregulation of the neuronal markers *MAP2*, *VGLUT2*, *BRN2* and *FOXG1* in day 30 iN-VOs (**Figure 2.4b**), while maintaining the expression level of endothelial markers *CDH5* and *VEPTP*, and smooth muscle marker *SMA* (**Figure 2.4b**). Immunofluorescence imaging confirmed the presence of *MAP2*+ neurons displaying robust neurite growth, along with interpenetrating *CDH5*+ vascular networks through 30 days in culture (**Figure 2.4c**). These organoids formed capillary networks that consisted of lumen-forming endothelial cells (**Figure 2.6a**) that were tightly associated with pericytes (**Figure 2.4d**).

To characterize the cell-type composition of iN-VOs cultured for up to 45 days, we assayed them using single cell RNA-sequencing (scRNA-seq). We assayed 26,959 cells across day 45 organoids grown under three conditions: 15 days of *NEUROD1* overexpression (iN-VO, 15 Days Induction), 45 days of *NEUROD1* overexpression (iN-VO, 45 Days Induction) and without overexpression (iN-VO, No Induction). Using the Seurat pipeline [23], we identified that *PRRX1*+ mesenchymal progenitors, *MEOX2*+ differentiating pericytes and cycling cells were present in all three types of organoids and constituted a majority of each of the organoids (**Figure 2.4e, 2.4f, 2.5a, 2.5b**). Importantly, functional vascular cell types - *PDGFRβ*+ pericytes and *PECAM1*+ endothelial cells were present in all types of organoids and constituted 16-22% and 2.5-6% of cells respectively (**Figure 2.4e, 2.4f, 2.5a, 2.5b**). Endothelial and neural cells mapped with high fidelity to the endothelium and neurons from the Mouse Cell Atlas, respectively (**Figure 2.5c**). To

further validate their identity, neurons from the neuro-vascular organoids were mapped to *DCX*+ neuronal

Figure 2.4: Molecular and Functional characterization of iN-VOs.

(**a**) Outline of long term cultured iN-VO characterization. **(b)** qRT-PCR analysis of signature endothelial genes *CDH5* and *VEPTP*; signature smooth muscle gene *SMA*; and signature neuronal genes *MAP2* and *BRN2* at day 30 of iN-VO culture. Data represent the mean ± s.d. (n = 7 organoids, from three independent experiments). **(c)** Immunofluorescence 100 μm z-stack, maximum projection, confocal micrographs of *MAP2*+ and *CDH5*+ induced day 30 iN-VOs (Scale bars = 100 μm) (**d**) Immunofluorescence100 μm z-stack, maximum projection, confocal micrographs of *PDGFR*+ and *CDH5*+ induced day 30 iN-VOs (Scale bars = 50 μm) **(e)** UMAP visualization of cell types from day 45 iN-VOs. Two independent induction conditions, along with one non-induction condition. **(f)** Cluster-specific expression of marker genes in day 45 iN-VOs. **(g)** Experimental validation of iN-VO perfusibility in vivo by subcutaneous implantation of iN-VO, showing immunofluorescence micrographs of intravital Dextran, *CDH5* and overlay (Scale bar = 25 μm). Representative image from two independent experiments.


A
RT-PCR &
Immunofluorescence
Expression
Genes
+BDNF, NT-3
scRNA-Seq
Day 15
iN-VO
Day 30+
iN-VO
In vivo
perfusion
B
Relative Expression to -Dox
ns
****
CDH5 VEPTP SMA
MAP2 BRN2
C
iN-VO, + Dox | Day30
MAP2 / CDH5
D
iN-VO, Pericyte and Endothelial Co-localization
PDGFR / CDH5
E
Differentiating Pericyte
Pericyte
Mesenchymal Progenitor
Cycling Cell
Unclassified
Pericyte Progenitor
Neuron
Epithelial Cell
Endothelial Cell
Stromal Cell
Epithelial Cell (Lipid processing)
Macrophage
F
PECAM
PDGFRB
DCX
KRT17
APOA1
CD68
PRRX1
MEOX2
COL15A1
G
Day 30
iN-VO
Subcutaneous
implantation in
immunodeficient mice
30 days post-implantation:
retro-orbital, intravenous
Texas Red injection
Fixation,
CDH5 Staining,
Confocal Imaging
Subcutanous Implant &
IV injection of Dextran
Dextran
hCDH5
MERGE

clusters from a reference dataset profiling 2D differentiation of neurons from hPSCs by NGN2 overexpression [24] (**Figure 2.5d, 2.5e**). After molecular characteristics of the iN-VOs were assayed, we assessed the functionality of both neural and vascular cell-types in the organoid. First we assessed endothelial function *in vitro* by confirming uptake of acetylated low-density lipoprotein by endothelial cells in the organoids (**Figure 2.6b**). Further, we assayed the ability of the vascular networks to form perfusable blood vessels when implanted into mice. Specifically, iN-VOs were grown till day 30 *in vitro*, and then subcutaneously implanted in $Rag2^{-/-};\gamma c^{-/-}$ immunodeficient mice. 30- and 90-days post-implantation, intravenous (IV) injection of a Texas Red dye followed by organoid extraction, fixation and immunofluorescence staining showed co-localization of the dye with human *CDH5*+ vascular cells and *PDGFR*+ pericytes in the iN-VO (**Figure 2.4g, 2.6c**). Endothelium function in vivo was sustained for at least 90 days post implantation, as demonstrated by confocal imaging of lumenized and perfused *CDH5*+ vessels (**Figure 1.6c(ii,iii)**). The absence of *CDH5* signal in wild-type mouse kidneys stained with human-specific anti-*CDH5*, demonstrates the human-specificity of the antibody (**Figure 2.6d**), confirming that the perfused *CDH5*+ vessels in the organoids are of human origin.

To assess functionality of neurons in the iN-VOs, we assayed the organoids for spontaneous electrical activity using MEAs (**Figure 2.7a, 2.7b**). Strikingly, while uninduced organoids displayed no spontaneous activity, spontaneous firing was repeatedly observed in iN-VOs (**Figure 2.7c**), confirming the presence of functional neurons. Taken together, we demonstrate the formation of functional neuro-vascular tissue with long term culture capability.

***In vitro* maturation of myo-vascular organoids**

Finally, while our iMB-VO platform was able to differentiate vascularized skeletal muscle, maturation of this lineage *in vitro* is a longstanding challenge. To further mature the differentiated skeletal muscle, we applied chronic electrical stimulation [25,26].

Figure 2.5: scRNA-seq characterization of neurovascular organoids.

(a) UMAP visualization of iN-VO clusters annotated by sample type. Two independent induction conditions, along with one non-induction condition. **(b)** Expression of marker genes for each classified cell type. **(c)** Mapping of neurovascular organoid cell types to cell types in the mouse cell atlas [36]. (**d**) UMAP visualization of clusters from a reference dataset of cells profiled during 2D differentiation of neurons from pluripotent stem cells by *NGN2* overexpression [24]. Cells were profiled at day 14 and day 35 after induction. Neuronal clusters in the 2D differentiated neurons are highlighted below by overlaying *DCX* expression on the UMAP visualization. (**e**) Mapping of neurovascular organoid neurons to clusters in the 2D differentiated neurons.

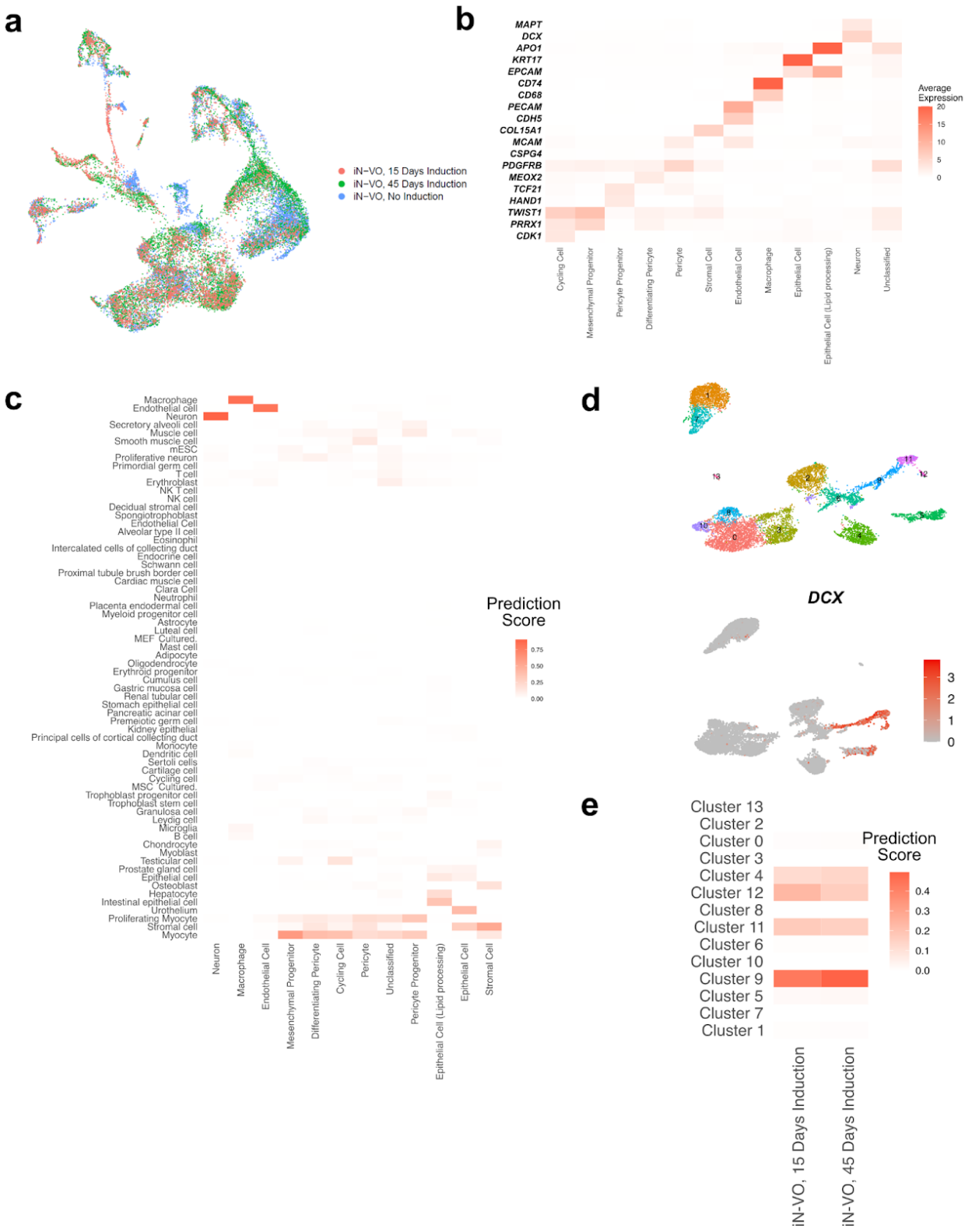
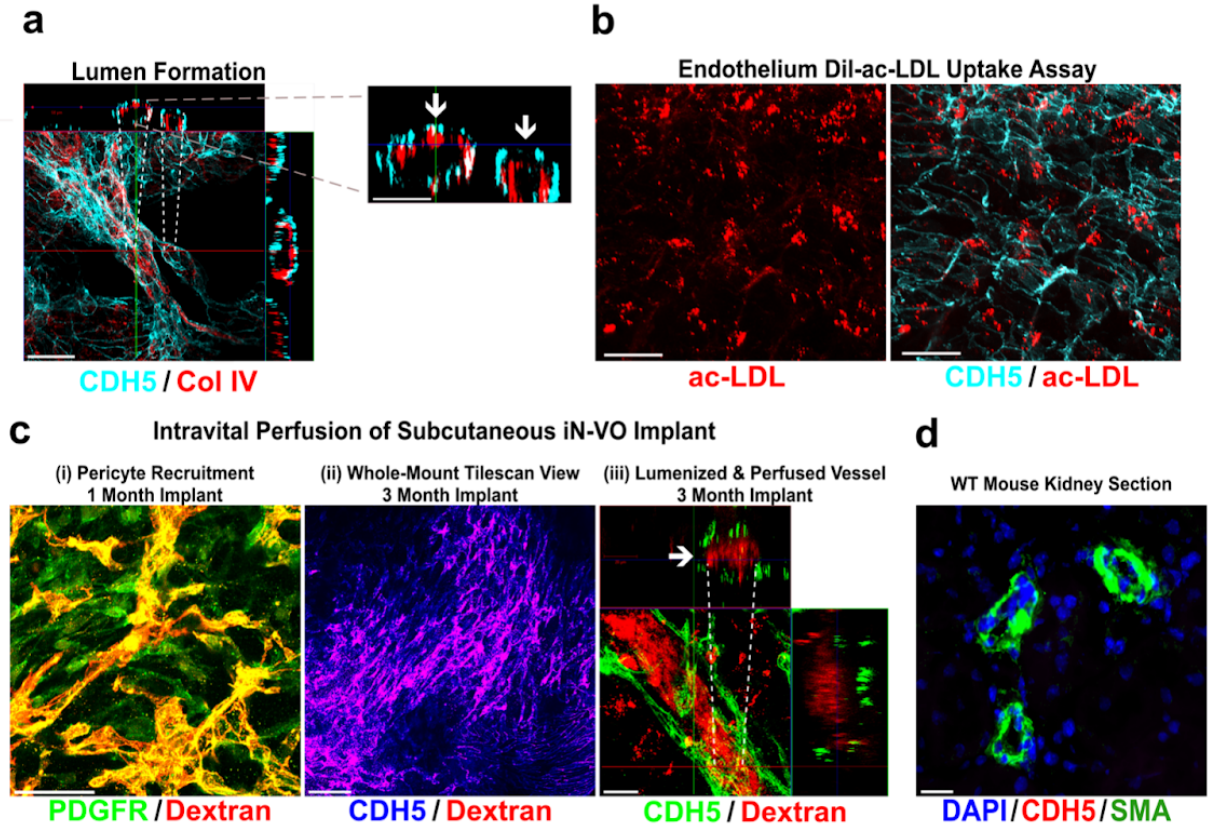

a
iN−VO, 15 Days Induction
iN−VO, 45 Days Induction
iN−VO, No Induction
b
MAPT
DCX
APO1
KRT17
EPCAM
CD74
CD68
PECAM
CDH5
COL15A1
MCAM
CSPG4
PDGFRB
MEOX2
TCF21
HAND1
TWIST1
PRRX1
CDK1
Cycling Cell
Mesenchymal Progenitor
Pericyte Progenitor
Differentiating Pericyte
Pericyte
Stromal Cell
Endothelial Cell
Macrophage
Epithelial Cell
Epithelial Cell (Lipid processing)
Neuron
Unclassified
Average Expression
20
15
10
5
0
c
Macrophage
Endothelial cell
Neuron
Secretory alveoli cell
Muscle cell
Smooth muscle cell
mESC
Proliferative neuron
Primordial germ cell
T cell
Erythroblast
NK T cell
NK cell
Decidual stromal cell
Spongiotrophoblast
Endothelial Cell
Alveolar type II cell
Eosinophil
Intercalated cells of collecting duct
Endocrine cell
Schwann cell
Proximal tubule brush border cell
Cardiac muscle cell
Clara Cell
Neutrophil
Placenta endodermal cell
Myeloid progenitor cell
Astrocyte
Luteal cell
MEF Cultured.
Mast cell
Adipocyte
Oligodendrocyte
Erythroid progenitor
Cumulus cell
Gastric mucosa cell
Renal tubular cell
Stomach epithelial cell
Pancreatic acinar cell
Premeiotic germ cell
Kidney epithelial
Principal cells of cortical collecting duct
Monocyte
Dendritic cell
Sertoli cells
Cartilage cell
Cycling cell
MSC Cultured.
Trophoblast progenitor cell
Trophoblast stem cell
Granulosa cell
Leydig cell
Microglia
B cell
Chondrocyte
Myoblast
Testicular cell
Prostate gland cell
Epithelial cell
Osteoblast
Hepatocyte
Intestinal epithelial cell
Urothelium
Proliferating Myocyte
Stromal cell
Myocyte
Neuron
Macrophage
Endothelial Cell
Mesenchymal Progenitor
Differentiating Pericyte
Cycling Cell
Pericyte
Unclassified
Pericyte Progenitor
Epithelial Cell (Lipid processing)
Epithelial Cell
Stromal Cell
Prediction Score
0.75
0.50
0.25
0.00
d
DCX
3
2
1
0
e
Cluster 13
Cluster 2
Cluster 0
Cluster 3
Cluster 4
Cluster 12
Cluster 8
Cluster 11
Cluster 6
Cluster 10
Cluster 9
Cluster 5
Cluster 7
Cluster 1
iN-VO, 15 Days Induction
iN-VO, 45 Days Induction
Prediction Score
0.4
0.3
0.2
0.1
0.0

Figure 2.6: Extensive molecular and phenotypic characterization of long-term cultured iN-VOs.

(a) The left image shows orthogonal sections of confocal stacks from neruo-vascular organoids stained for *CDH5* (red) and *Col IV* (Scale bars = 50 μm). The right image shows a zoomed in view of formed lumens, as indicated with arrowheads (Scale bars = 25 μm). **(b)** Immunofluorescence 100 μm z-stack, maximum projection, confocal micrographs of Day 30 neuro-vascular organoids labeled for *CDH5* post Dil-ac-LDL endothelial uptake assay (Scale bars = 50 μm). **(c)** Further characterization of endothelial functionality *in vivo* via intravital perfusion shows: (i) Experimental validation of pericyte recruitment of *in vivo* perfused neuro-vascular organoids, showing immunofluorescence micrograph of intravital Dextran and *PDGFRB* overlay (Scale bar = 50 μm); (ii) 100 μm z-stack projection that spans a 650 x 650 μm tilescan region of whole-mount neuro-vascular organoid implant showing intravital Dextran and *CDH5* overlay (Scale bar = 100 μm); and (iii) Orthogonal sections of confocal stacks from neuro-vascular organoids implants 90 days post-implantation, stained for *CDH5* (green) and showing dextran perfused within the vessel. Lumens are indicated with arrowheads (Scale bars = 20 μm). **(d)** Wild-type mouse kidney stained for human-specific anti-*CDH5* and anti-*SMA* antibodies. Note the absence of the *CDH5* signal (Scale bar = 10 μm).


a
Lumen Formation
CDH5 / Col IV
b
Endothelium Dil-ac-LDL Uptake Assay
ac-LDL
CDH5 / ac-LDL
c
Intravital Perfusion of Subcutaneous iN-VO Implant
(i) Pericyte Recruitment
1 Month Implant
PDGFR / Dextran
(ii) Whole-Mount Tilescan View
3 Month Implant
CDH5 / Dextran
(iii) Lumenized & Perfused Vessel
3 Month Implant
CDH5 / Dextran
d
WT Mouse Kidney Section
DAPI / CDH5 / SMA

To subject organoids to stimulation, we encapsulated them in a fibrin and matrigel blend, and placed them in a custom chip between two graphite rods. A pulsed constant voltage stimulation was then applied for a week after encapsulation to drive maturation (**Figure 2.7d**). Organoids were then assayed for gene expression of muscle and calcium handling genes (**Figure 2.7e**). In stimulated organoids embryonic skeletal muscle myosin (*MYH3*) was upregulated, along with a small increase in expression of adult fast skeletal muscle myosin (*MYH2*). Additionally, the calcium handling genes *CASQ2* and *SERCA2* were also highly upregulated. In summary, our iMB-VO approach is a promising method to generate mature, vascularized skeletal muscle tissue *in vitro*.

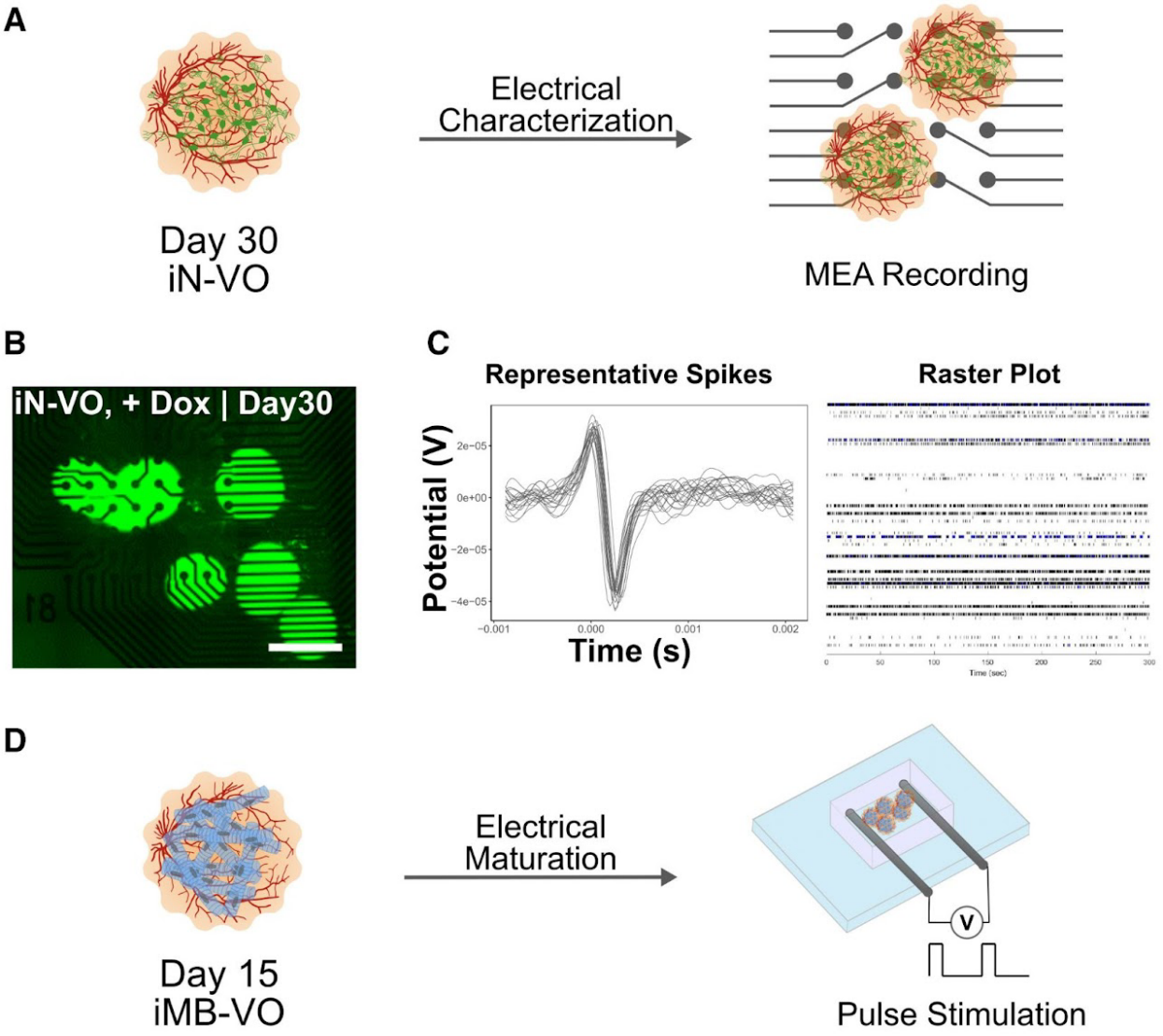
Discussion

In this work we have demonstrated that a combination of directed differentiation and the overexpression of a lineage-specific reprogramming factor can build cross-lineage and cross-germ layer organoid systems. We used this approach to introduce neurons and skeletal muscle into vascular organoids, as a proof-of-principle of such an approach. We then assayed these mixed-lineage organoids to demonstrate that this combined approach yielded neurons only upon induction of *NEUROD1* overexpression, and skeletal muscle only upon induction of *MYOD1* + *BAF60C* overexpression, while retaining the architecture of the vascular organoid. The neuro-vascular organoids thus generated were further optimized for long term culture, and comprehensively characterized for composition and to confirm function of both lineage branches. Finally, for the myo-vascular organoids, we demonstrated the maturation of these organoids via chronic electrical stimulation, which enhanced the expression of skeletal muscle myosins and calcium handling genes.

While providing a modular and powerful approach to engineer vascularized multi-lineage organoids, several challenges remain to be solved. While we have taken advantage of known

Figure 2.7: Electrical characterization and stimulation of organoids.

(a) Schematic of iN-VO electrical characterization via microelectrode array (MEA) recordings. **(b)** Image of day 30 iN-VOs plated on microelectrode array (Scale bars = 500 μm). **(c)** Representative spike plots from MEA measurements of spontaneously firing iN-VOs and corresponding raster plot. Representative image and plot from two independent experiments. **(d)** Schematic of iMB-VO *in vitro* maturation by electrical stimulation. **(e)** qRT-PCR analysis of skeletal muscle myosins: *MYH2, MYH3, MYH7* and *MYH8*; and genes involved in calcium handling, *CASQ1, CASQ2, SERCA1, SERCA2* and *RYR* for stimulated vs unstimulated iMB-VOs. Data represent the mean ± s.d. (n = 6 organoids, from two independent experiments) **(e)** (**$P \leq 0.01$, ***$P \leq 0.001$, and ****$P \leq 0.0001$; ns = not significant).


A
Day 30
iN-VO
Electrical
Characterization
MEA Recording
B
iN-VO, + Dox | Day30
C
Representative Spikes
Raster Plot
Potential (V)
Time (s)
D
Day 15
iMB-VO
Electrical
Maturation
V
Pulse Stimulation
E
Relative to GAPDH Expression
MYH1
MYH3
MYH7
MYH8
-Stimulation
+1 Week Stimulation, 1Hz
CASQ1
CASQ2
SERCA1
SERCA2

reprogramming factors to differentiate neurons and skeletal muscle, such recipes are not known for all cell types and will have to be discovered. Known overexpression recipes may also need to be modified for organoid culture conditions, since recipes are typically optimized and validated only for specific 2D culture and media conditions. While we have restricted the current study to overexpression, knockdown approaches [27] may also be important tools in these engineering efforts. Additionally, improved control of this spatial organization will be critical for accurate tissue engineering, especially as reprogrammed cells transition directly from initial to final lineage states, and one may need to incorporate novel strategies like optogenetic control of gene expression [28–30] or synthetic biology approaches [31].

In summary, our demonstrated approach provides a blueprint for complete bottom-up engineering of vascularized organoids, where genetic perturbation methods for cell reprogramming can be overlaid onto our core protocol to introduce additional cell-types. This framework could be a powerful platform to generate diverse vascularized organoids, thus laying the foundation for completely *in vitro* derived, large-scale tissue systems for regenerative medicine purposes.

Methods

Plasmid Construction

The piggyBac transposon plasmids for inducible overexpression of TFs were constructed using the backbone from 138-dCas9-Dnmt3a (Addgene Plasmid #84570) [32]. The backbone plasmid was digested with NdeI and NsiI to remove the neomycin resistance cassette and was replaced with a puromycin resistance cassette using multi-element Gibson assembly. This puromycin resistant plasmid was then digested with NheI and AgeI to remove the dCas9-DNMT3A fusion sequence and replaced with the TF and fluorescent protein sequences separated

by 2A peptide sequences using a multi-element Gibson assembly. TF and fluorescent protein sequences were amplified from plasmids: EF1a_NEUROD1_P2A_Hygro_Barcode (Addgene Plasmid #120466), EF1a_ASCL1_P2A_Hygro_Barcode (Addgene Plasmid #120427), phDLX2_N174 (Addgene Plasmid #60860) [33], EF1a_MYOD1_P2A_Hygro_Barcode (Addgene Plasmid #120464), pBS-hBAF60C (Addgene Plasmid #21036) [34], EF1a_mCherry_P2A_Hygro_Barcode (Addgene Plasmid #120426), and pEGIP (Addgene Plasmid #26777).

To construct a plasmid expressing the hyperactive piggyBac transposase [22], the sequence for the enzyme was obtained as a synthesized double-stranded DNA fragment (Integrated DNA Technologies). This was cloned into an in-house plasmid using Gibson assembly, such that the expression of the transposase is driven by a CAG promoter.

The Gibson assembly reactions were set up as follows: 100 ng digested backbone, 3:10 molar ratio of insert, 2X Gibson assembly master mix (New England Biolabs), H_20 up to 20 µl. After incubation at 50 °C for 1 h, the product was transformed into One Shot Stbl3 chemically competent Escherichia coli (Invitrogen). A fraction (150 µL) of cultures was spread on carbenicillin (50 µg/ml) LB plates and incubated overnight at 37 ºC. Individual colonies were picked, introduced into 5 ml of carbenicillin (50 µg/ml) LB medium and incubated overnight in a shaker at 37 ºC. The plasmid DNA was then extracted with a QIAprep Spin Miniprep Kit (Qiagen), and Sanger sequenced to verify correct assembly of the vector. Following verification of the vector, larger amounts of plasmid were obtained by seeding 150 µl of bacterial stock into 150 ml of LB medium containing carbenicillin (50 µg/ml) and incubating overnight in a shaker at 37 ºC for 16-18 h. The plasmid DNA was then extracted using a Plasmid Maxi Kit (Qiagen).

Cell Culture

H1 hESCs were maintained under feeder-free conditions in mTeSR1 medium (Stem Cell Technologies). Prior to passaging, tissue-culture plates were coated with growth factor-reduced Matrigel (Corning) diluted in DMEM/F-12 medium (Thermo Fisher Scientific) and incubated for 30 minutes at 37 ºC, 5% CO_2. Cells were dissociated and passaged using the dissociation reagent Versene (Thermo Fisher Scientific).

Generation of Clonal Inducible Overexpression Lines

hESC cells at 50-75% confluency from 3 wells of a 6-well plate were passaged using Versene. The cells were spun at 300 rcf for 5 minutes to obtain a cell pellet and this pellet was resuspended in a buffer containing 100 µl of P3 Nucleofector Solution (Lonza) and up to a maximum of 15 µl of a 1:1 mix of transposon vector plasmid to transposase plasmid by mass. This solution was loaded into a single Nucleovette (Lonza) and electroporated using the CB-150 pulse program on the 4D Nucleofector system (Lonza). After nucleofector run completion, 500 µl of pre-warmed mTeSR1 containing 10 µM Y27632 (Tocris Bioscience) was added to the Nucleovette and incubated at room temperature for 5 minutes. The cells were then removed from the Nucleovette using a Pasteur pipette and transferred dropwise into a 10 cm plate coated with growth-factor reduced Matrigel as previously described and containing pre-warmed mTeSR1 with 10 µM Y27632.

Medium was then changed daily, and 48 hours after nucleofection cells were maintained under puromycin (Thermo Fisher Scientific) selection at 0.75 µg/ml. After approximately 7-10 days of culture post-nucleofection, colonies were large enough for clonal selection. To pick clonal lines, cells were treated with Versene for 3 minutes, Versene was aspirated and the plate was filled with DMEM/F-12 with 1% antibiotic-antimycotic (Thermo Fisher Scientific). Individual colonies were then carefully scraped under a microscope and transferred into individual wells of a 24-well

plate coated with growth-factor reduced Matrigel and containing pre-warmed mTeSR1. These individually picked clones were expanded, aliquots were frozen in mFreSR (Stemcell Technologies) and validated by differentiation to relevant cell types. One validated clone from each line was chosen for further experiments. All clones were maintained in mTeSR1 under selection with puromycin at 0.75 μg/ml.

2D Differentiation of Inducible Overexpression hESC lines

Clonal lines overexpressing *NEUROD1* were differentiated following a previously described protocol [19]. Briefly, cells were passaged as single cells using Accutase (Innovative Cell Technologies) and plated in mTeSR1 at a density of 4-5x10^5 cells per well of a 6-well plate. The following day medium was changed to DMEM/F12 containing N2 supplement (Thermo Fisher Scientific), MEM non-essential amino acids (Thermo Fisher Scientific), 0.2 μg/ml mouse laminin (Invitrogen), 10 ng/ml BDNF (Peprotech), 10 ng/ml NT3 (Peprotech), 0.75 μg/ml puromycin and 1 μg/ml doxycycline (Sigma Aldrich) and cells were maintained in this medium for 2 days. On day 3 of differentiation cells were re-plated on Matrigel coated wells along with mouse glial cells in Neurobasal medium (Thermo Fisher Scientific) containing Glutamax (Thermo Fisher Scientific), B27 supplement (Thermo Fisher Scientific), 10 ng/ml BDNF, 10 ng/ml NT3 and 1 μg/ml doxycycline. From day 5 onward, 2 μM Ara-c (Sigma Aldrich) was added to the medium to inhibit astrocyte proliferation. 50% of the medium was subsequently changed every 2-3 days. Cells were maintained in culture for up to 3 weeks. For functional characterization and electrical measurements, cells were plated on Matrigel coated 6-well multi-electrode arrays (Axion Biosystems) with mouse glial cells and maintained in culture for up to 3 weeks.

Clonal lines overexpressing *ASCL1* and *DLX2* were differentiated following a previously described protocol [20]. Briefly, cells were passaged as single cells using Accutase (Innovative Cell

Technologies) and plated in mTeSR1 at a density of 4-5x10^5 cells per well of a 6-well plate. The following day medium was changed to DMEM/F12 containing N2 supplement, MEM non-essential amino acids, 0.75 μg/ml puromycin and 1 μg/ml doxycycline and cells were maintained in this medium for 7-8 days, with the medium being changed every 2-3 days. 2 μM Ara-C was added to the medium on day 5 of differentiation. On day 7-8, the cells were passaged with Accutase and re-plated on Matrigel coated plates at a density of 4x10^5 cells per well of a 6-well plate in Neurobasal medium containing Glutamax, B27 supplement, 2 μM Ara-c and 1 μg/ml doxycycline. 50% of the medium was subsequently changed every 2-3 days. From day 15 onwards, medium was supplemented with 20 ng/ml BDNF and doxycycline was removed. For functional characterization and electrical measurements, cells were plated on Matrigel coated 6-well multi-electrode arrays (Axion Biosystems) with mouse glial cells and maintained in culture for up to 5 weeks.

Clonal lines overexpressing *MYOD* and *BAF60C* were differentiated to skeletal muscle following a process similar to a previously described protocol [21]. Briefly, cells were passaged as single cells using Accutase and plated at a density of 4-5x10^5 cells per well of a 6-well plate. The following day, medium was changed to DMEM/F12 containing 15% fetal bovine serum (FBS, Thermo Fisher Scientific) and 1% anti-anti (Thermo Fisher Scientific). Medium was exchanged every 2 days. On day 5 of differentiation, medium was changed to DMEM/F12 containing 2% horse serum (Hyclone) and 1% anti-anti. Differentiating cells were cultured for 3 weeks.

iMB-VO and iN-VO Generation

hESCs were grown in one well of a 6-well plate till they were 80% confluent. This was sufficient to seed one ultra-low attachment (ULA) 96-well plate of embryoid bodies. To passage the cells, mTeSR was aspirated and the cells washed with PBS. 1 mL of Accutase was then added

to the well and incubated at 37 °C incubator for 4-6 minutes. Cells were detached by tapping the sides of the plate. 1 mL of mTesr was then added to the well, and the detached cells were titurated with a 200 µl pipette to break up cell clumps and to obtain a single cell suspension. Cells were then spun down at 300 rcf for 5 minutes. Once the cells were pelleted, , the supernatant was removed, and cells resuspended in EB medium - (DMEM-F12, 20% KOSR) + 50 µM Y-27632 - at a concentration of 72,000 cells/ml. 125 µL of this cell suspension, was dispensed into each well of an ULA 96 well plate. hESCs were cultured overnight at 37 ºC, 5% CO_2, allowing them to aggregate into embryoid bodies (EBs).

EBs were grown for 1-3 days till 200-400 µm in diameter. Once this size, EBs were transferred into an ULA 6-well plate using a cut 200 µL pipette tip, with a maximum of ~24 EBs per well of the ULA 6 well plate. Excess EB medium was carefully removed and 2 mL of Mesoderm Induction Media: N2/B27 medium - (1:1 DMEM/F12-Neurobasal, 100x N2, 50x B27) - + 3 µM CHIR (Tocris), 30 ng/mL BMP4 (Peprotech), 1 µg/ml of doxycycline (dox); was added to each well. These were then cultured at 37 ºC, 5% CO_2 for 3 days.

Three days later, the medium was replaced with 2 mL of Vascular Induction Media: N2/B27 medium + 100 ng/ml VEGF (Peprotech) + 10 µM Forskolin (Sigma-Aldrich) + 1 µg/ml dox per well; and cultured at 37 ºC, 5% CO_2. 24 hours later, medium was removed and replaced with 2 mL of N2/B27 media + 100 ng/ml VEGF + 10 µM Forskolin + 1 µg/ml dox per well. Organoids were then cultured at 37 ºC, 5% CO_2 for 24 hours.

Organoids were then encapsulated in a blend of matrigel and collagen (Mat-Col gel: 2 mg/mL Collagen (Advanced Biomatrix) + 20% Matrigel). Briefly, parafilm wells were prepared by placing a piece of parafilm onto an empty 200 µL pipette tip box, and pressing into the tip cavities to create dimples. A 7 x 7 grid of wells was found to be optimal since it prevented

organoids from drying and allowed sufficient organoids to be encapsulated for culturing in one 10 cm dish. The dimpled parafilm was then placed in a 10 cm dish. The Mat-Col Gel Blend was prepared and placed on ice.

Using a cut 200 µL pipette tip, organoids were transferred from the ULA 6 well plate, and placed individually into the parafilm wells. A maximum of ~30 µL of media was transferred with each organoid to avoid overfilling of the parafilm well. Once all organoids were placed in individual wells, excess media was removed, leaving only the organoid in the well. 30 µL of the Mat-Col gel was added to each parafilm well. Individual organoids were checked to ensure encapsulation in the gel solution, and the 10 cm dish was then incubated at 37 ºC, 5% CO_2 for 45 minutes for the gel blend to completely gelate.

Once gelation was complete, the encapsulated organoids were washed off the parafilm using Vascular Maturation Media: StemPro 34 Media + 15% FBS + 100 ng/ml VEGF + 100 ng/ml bFGF + 1 µg/ml dox**.** Once gel droplets were completely washed off from the parafilm, a cut 1000 µL pipette tip was used to transfer the organoids back to the original ULA 6 well plate used in the previous steps of the experiment. Organoids were then cultured at 37 ºC, 5% CO_2 and medium replaced every 3 days using Vascular Maturation Media until day 15.

For long term culture of iN-VOs, at day 15, medium was changed to Vascular Maturation Media + 20 ng/ml BDNF + 20 ng/ml NT3 + 1 µg/ml dox. Medium was replaced every 3-5 days.

Animals

Housing, husbandry and all procedures involving animals used in this study were performed in compliance with protocols (#S16003) approved by the University of California San Diego Institutional Animal Care and Use Committee (UCSD IACUC). Mice were group housed (up to 4 animals per cage) on a 12:12 hr light-dark cycle, with free access to food and water in

individually ventilated specific pathogen free (SPF) cages. All mice used were healthy and were not involved in any previous procedures nor drug treatment unless indicated otherwise. All studies performed in NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ (NSG) mice and maintained in autoclaved cages.

In Vivo Perfusion

iN-VOs were cultured for 30 days before being implanted subcutaneously into Rag2$^{-/-}$;γc$^{-/-}$ immunodeficient mice. To prepare the mice for subcutaneous implantation, the right hind-flank region was shaved and wiped down with povidone-iodine. Then, a one-inch, subcutaneous incision was made, and Day 30 iN-VOs suspended in Matrigel were placed inside the incision region using a cut pipette tip. These organoids were then matured for an extra 30 days in-vivo. To test for proper perfusion of the vasculature, mice were given an intravenous injection of lysine fixable Texas-Red Dextran (1.25 mg per mouse, Thermo Fisher Scientific) and sacrificed after 15 minutes of allowing the dextran to pass through circulation. Organoids were retrieved from the subcutaneous region, fixed and whole-mount stained, as described below.

MEA Measurements

For 2D differentiated neurons, cells were plated on Matrigel coated 6-well multi-electrode arrays (CytoView MEA 6, Axion Biosystems) with mouse glial cells and maintained in culture for up to 3 weeks.

iN-VOs were not encapsulated in Mat-Col gel when preparing them for MEA measurements. MEA electrodes (CytoView MEA 6, Axion Biosystems) were spot-coated with 2% Matrigel and placed in a cell-culture incubator to incubate at 37 ℃ overnight. Because gel encapsulation prevented proper adhesion between the organoid and MEA well, the following day, one day-25 un-encapsulated iN-VO was carefully put in the center of the MEA well with ~50

μL of media. The organoid was left untouched for 2 hours, and then flooded with 0.5 mL of media. PBS was filled in the side compartments of the MEA plate to prevent cell media evaporation. The MEA plates were then left undisturbed for 5 days to ensure robust attachment to the well. MEA measurements were taken on day 30, 5 days after seeded onto the plate. To collect measurements, MEA plates were placed in the reader with the reader plate heater set to 37 ºC and under 95% O_2/5% CO_2 air flow. Plates were allowed to equilibrate under these conditions for a minimum of 5 minutes before collecting spontaneous recordings for 4 minutes.

Electrical signals were collected and analyzed using AxIS Software (Axion Biosystems) with Spontaneous Neural configuration. Signals were filtered with a band-pass filter of 200 Hz – 3 kHz. Spikes were detected with AxIS software using an adaptive threshold crossing set to 5.5 times the standard deviation of the estimated noise for each electrode.

***In vitro* electrical stimulation**

To create chambers for electrical stimulation, custom designed chips consisting of a porous inner well and a solid outer well were fabricated via extrusion printing of a silicone elastomer (Dow Corning Dowsil SE 1700) on glass slides, on a custom 3D printer consisting of a three-axis gantry (AGB 10000, Aerotech) and pneumatic dispensers (Nordson Ultimus I). Chips were cured for at least two hours at 80 ºC to fully crosslink the elastomer. Graphite rods were then inserted into these chips such that they were located at either end of the inner well and gaps were sealed with PDMS (Dow Corning, Sylgard 184). After sealing, chips were again cured for at least two hours at 80 ºC.

To encapsulate iMB-VOs, a solution composed of Fibrin (3 mg/ml, Sigma Aldrich) + 20% Matrigel was prepared similarly to a previously described protocol [25]. Up to five individual organoids were transferred into the inner well, excess medium removed and the space filled with

the hydrogel mixed with thrombin (1 U/ml, Sigma-Aldrich). The hydrogel was allowed to fully gelate and crosslink at 37 ºC for one hour, after which the outer well was filled with VO culture medium containing 1 μg/ml dox. For stimulated samples, wires were attached to the graphite rods and routed to Arduino Uno microcontrollers equipped with Motor Shields. The microcontrollers were programmed to provide chronic electrical stimulation at 0.4 V/mm, 1 Hz with a 2 ms on time. Encapsulated organoids were then cultured in 37 ºC, 5% CO_2 for eight days. On the second day after encapsulation, electrical stimulation was started and maintained for one week.

Immunostaining

Organoids were removed from the culture dish and added to a 1.5 mL centrifuge tube. Up to 20 organoids could be combined into one tube and used in subsequent steps. Excess medium was removed and organoids were washed once with 1 mL PBS. PBS was removed and 500 μL of 4% PFA solution was added to the microcentrifuge tube. Organoids were fixed at room temperature for 1 hour, protected from light. After 1 hour, PFA solution was removed and exchanged with 500 μL PBS. At this point, organoids could be stored in PBS at 4 °C protected from light for up to 1 month.

To block the organoids, and prepare them for immunostaining, PBS was removed and 500 μL of blocking buffer (3% FBS, 1% BSA, 0.5% Triton-X, 0.5% Tween) was added. The tube was placed into a tube rack and then onto an orbital shaker, shaking at 150 rpm for 2 hours to fully block and permeabilize the organoids. Blocking buffer was then removed and 100 μL of primary antibody diluted in blocking buffer was added. All antibodies used were diluted 1:100 in blocking buffer. The tube then was placed back onto the tube rack and onto an orbital shaker (LSE Orbital Shaker, Corning) at 4 °C. The orbital shaker was set to 12 rpm and organoids incubated at 4 °C overnight.

After overnight incubation, blocking buffer was removed, and organoids washed with PBS-T (PBS + 0.05% Tween) three times for 20 minutes. Organoids were placed on an orbital shaker set to 150 rpm during each PBS-T wash.

After washing in PBS-T, 100 µL of secondary antibodies diluted in blocking buffer were added. Organoids were incubated with the secondary antibodies at room temperature for 2 hours, while keeping the samples protected from light. After secondary staining was complete, organoids were washed with PBS-T three times for 20 minutes. Organoids were placed on an orbital shaker set to 150 rpm during each PBS-T wash.

Once secondary staining was complete, a coverslip was prepared for whole-mounting of the organoids. This was done by applying epoxy (Loctite Epoxy) to the non-adhesive surface of an iSpacer (Sunjin Lab) and then attaching the iSpacer to a coverslip. Within 5 minutes the iSpacer was bound to the coverslip. Using a cut 1000 µL pipette tip, 2-4 organoids were transferred to each iSpacer well. Excess PBS was removed and 50 µL of Fluromount G was added to each well. iSpacer cover was then peeled off and a second coverslip attached the exposed sticky side. Whole-mount samples could be stored in 4 °C protected from light for up to 8 months. Confocal images were taken using a LSM 880 with Airyscan Confocal Microscope (Zeiss).

All of the Primary and Secondary Antibodies used in this protocol are diluted in Blocking Buffer at a 1:100 dilution factor. anti-VE-Cadherin (D87F2, Cell Signaling Technologies), anti-MAP2 (HM-2, Sigma-Aldrich), anti-MYH (MF-20, DSHB), anti-PDGFR (AF385, R&D Systems) and anti-SMA (MAB1420, R&D Systems) were used for primary antibody staining. anti-Rabbit Alexa 405 (Thermo Fisher Scientific, A-31556), anti-Rabbit DyLight 550 (Thermo Fisher Scientific, 84541), and anti-Mouse Alexa 647 (Thermo Fisher Scientific, PIA32728) were used for secondary antibody staining.

For endothelial function Dil–acetylated low-density lipoprotein (Dil-ac-LDL) uptake assay, neuro-vascular organoids were incubated with with 10 μg ml−1 Dil-ac-LDL (Thermofisher, L3484) for 6 hours and then washed several times in medium before immunostaining and imaging.

RNA Extraction and qRT-PCR

RNA was extracted from cells using the Qiazol and RNeasy Mini Kit (Qiagen) as per the manufacturer's instructions. The quality and concentration of the RNA samples was measured using a spectrophotometer (Nanodrop 2000, Thermo Fisher Scientific). cDNA was prepared using the Protoscript II First Strand cDNA synthesis kit (New England Biolabs) in a 20 μl reaction and diluted up to 1:2 with nuclease-free water.

qRT-PCR reactions were setup as: 2 μl cDNA, 400 nM of each primer, 2X iTaq Universal SYBR Supermix (Bio-Rad), H2O up to 20 μl. qRT-PCR was performed using a CFX Connect Real Time PCR Detection System (Bio-Rad) with the thermocycling parameters: 95 °C for 3 min; 95 °C for 3 s; 60 °C for 20 s, for 40 cycles. All experiments were performed in triplicate and results were normalized against a housekeeping gene, GAPDH. Relative mRNA expression levels, compared with GAPDH, were determined by the comparative cycle threshold ($\Delta\Delta C_T$) method.

The following primers were used for qPCR reactions:

Table 2.1: qPCR primers

MAP2 Forward primer	CTCAGCACCGCTAACAGAGG
MAP2 Reverse primer	CATTGGCGCTTCGGACAAG
TUBB3 Forward primer	GGCCAAGGGTCACTACACG
TUBB3 Reverse primer	GCAGTCGCAGTTTTCACACTC

Table 2.1: qPCR primers

VGLUT2 Forward primer	GGGAGACAATCGAGCTGACG
VGLUT2 Reverse primer	TGCAGCGGATACCGAAGGA
VGAT Forward primer	ACGTCCGTGTCCAACAAGTC
VGAT Reverse primer	AAAGTCGAGGTCGTCGCAATG
BRN2 Forward primer	AAGCGGAAAAAGCGGACCT
BRN2 Reverse primer	GTGTGGTGGAGTGTCCCTAC
FOXG1 Forward primer	CCGCACCCGTCAATGACTT
FOXG1 Reverse primer	CCGTCGTAAAACTTGGCAAAG
CDH5 Forward primer	AAGCGTGAGTCGCAAGAATG
CDH5 Reverse primer	TCTCCAGGTTTTCGCCAGTG
VEPTP Forward primer	ACAACACCACATACGGATGTAAC
VEPTP Reverse primer	CCTAGCAGGAGGTAAAGGATCT
SMA Forward primer	GTGTTGCCCCTGAAGAGCAT
SMA Reverse primer	GCTGGGACATTGAAAGTCTCA
MYH1 Forward Primer	ATCTAACTGCTGAAAGGTGACC

Table 2.1: qPCR primers

MYH1 Reverse Primer	TAAGTAAATGGAGTGACAAAG
MYH2 Forward Primer	GCCGAGTCCCAGGTCAACAAG
MYH2 Reverse Primer	TGAGCAGATCAAGATGTGGCAAAG
MYH7 Forward Primer	CTGTCCAAGTTCCGCAAGGT
MYH7 Reverse Primer	TCATTCAAGCCCTTCGTGCC
CASQ1 Forward Primer	ACATTGTGGCCTTCGCAGAG
CASQ1 Reverse Primer	CCATACGCTATCCGCATCAGT
SERCA1 Forward Primer	GAAGGGAGCACAATGGAGGC
SERCA1 Reverse Primer	CAGGCCAGCACGAAGGAAAT

Statistics

All statistics on gene expression qPCR plots were assayed via an unpaired two-tailed t-test. P values were assessed as significant as follows: $^{**}P \leq 0.01$, $^{***}P \leq 0.001$, and $^{****}P \leq 0.0001$; ns = not significant. Statistical analysis was carried out using GraphPad Prism 8.

Single cell RNA-seq Processing

To dissociate organoids for single cell RNA-seq, 5-6 organoids were incubated in a 1 mL 20 U/mL Papain solution (Worthington, LS003126) for 30 minutes, passed through a 40 μm filter, spun down at 300 rcf for 5 minutes, and resuspended in 0.04% BSA solution. Cells were then loaded into the Chromium Chip B (10x Genomics) and single cell libraries were made using

Chromium Single Cell 3′ Reagent Kits v3 workflow (10x Genomics). Fastq files were aligned to a hg19 reference and expression matrices generated using the count command in cellranger v3.0.1 (10X Genomics).

Data Integration and Clustering

Data integration was performed on the expression matrices from all 3 organoids: iN-VOs induced for 15 days; iN-VOs induced for 45 days; and iN-VOs which were not induced. Integration was done using the Seurat v3 pipeline [23]. Expression matrices were filtered to remove any cells expressing less than 200 genes or expressing greater than 10% mitochondrial genes. DoubletFinder [35] was used to detect predicted doublets, and these were removed for downstream analysis. The expression matrix was then normalized for total counts, log transformed and scaled by a factor of 10,000 for each sample, and the top 4000 most variable genes were identified. We then used Seurat to find anchor cells and integrated all data sets, obtaining a batch-corrected expression matrix for subsequent processing. This expression matrix was scaled, and nUMI as well as mitochondrial gene fraction was regressed out. Principal component analysis (PCA) was performed on this matrix and 22 PCs were identified as significant using an elbow plot. The 22 significant PCs were then used to generate a k-nearest neighbors (kNN) graph with k=10. The kNN graph was then used to generate a shared nearest neighbors (sNN) graph followed by modularity optimization to find clusters with a resolution parameter of 0.8.

To classify cell types, the integrated dataset was mapped to annotated cell types in the Microwell-seq Mouse Cell Atlas [36] using Seurat label transfer on the intersection of genes in the mouse and organoid datasets, and further refined using cell type-specific marker genes. We finally visualized the results using UMAP dimensionality reduction on the first 22 PCs.

To assess tissue-specificity of endothelial cells, the endothelial cell cluster was subsetted and mapped using Seurat label transfer to tissue-specific endothelial cells from the Tabula Muris consortium. To confirm neuronal character of neurons from the neurovascular organoids, the neuron cluster was subsetted and mapped via Seurat label transfer to cells profiled during differentiation of neurons from pluripotent stem cells in 2D using *NGN2* overexpression. For all mapping to a reference data set, the prediction score is assessed as the mean of the Seurat predicted identity score across all cells with a particular identity.

Acknowledgments

Chapter 2, in full, is a reprint of the following published manuscript: Amir Dailamy*, Udit Parekh*, Dhruva Katrekar, Aditya Kumar, Daniella McDonald, Ana Moreno, Pegah Bagheri, Tse Nga Ng, Prashant Mali, "Programmatic introduction of parenchymal cell types into blood vessel organoids," ***Stem Cell Reports*** 16, pp. 2432-2441 (2021). The dissertation author was the primary investigator and author of this paper.

Chapter 3 - Engineering *in vivo* Vascularized Tissues via Molecular Sculpting of Teratomas.

Abstract

A human pluripotent stem cell-derived (hPSC) teratoma is a lab-grown, tumor-like mass composed of a variety of cell types, including tissues from all three germ layers (ectoderm, mesoderm, and endoderm). Teratomas are often used in stem cell research to examine the differentiation potential of hPSCs and are usually formed via the subcutaneous injection of hPSCs in immunodeficient mice. By using microRNA (miRNA)-mediated suicide gene circuits, we assessed the feasibility of organ sculpting in the teratoma via enriching for desired cell-types by depleting undesired ones. This approach relies on a gene circuit composed of a drug-inducible Caspase9 (iCaspase9) that contains a lineage-specific microRNA (miRNA) target site in its 3' UTR. The miRNA target site allows for cell-type level control of the genetic switch, as endogenous expression of a lineage-specific miRNA will silence the iCaspase9 triggered apoptosis in desired cells, allowing for cell-type specific enrichment in a multi-lineage tissue environment. As a proof of concept, we demonstrate miRNA-mediated neural enrichment in teratomas.

Introduction

hPSCs have been used to model human developmental processes for the past two decades, and there has been significant progress made in understanding early-stage development using 2D monolayer culture systems [37,38,39]. However, development occurs in 3 dimensions, and thus better model systems have since been developed to recapitulate key developmental phenomena in 3D [40].

Organoids have played a major role in this transition to 3 dimensions, as they have been used to make complex organ mimics that faithfully recapitulate the cell-type composition, cytoarchitecture and functions of their in vivo counterparts [1,2,40]. Organoids have been used to generate a diverse array of lineage-specific model systems [8,10,18,41–43], however they suffer in that the extrinsic media conditions can typically only dictate the tissue towards one particular lineage

at a time. This is in contrast to how organogenesis occurs in utero [44,45], as it has been extensively studied that neighboring lineages in the developing embryo can guide the proper differentiation and maturation of many organ systems through paracrine and mechanical signaling [46]. Thus, a tissue engineering platform in which all germ layers can differentiate simultaneously is ideal to allow for any cross-germ layer interactions that are necessary for proper organ development.

The teratoma is currently the gold-standard assay for hPSC pluripotency and is formed by subcutaneous injection of a 1:1 mixture of hPSCs and Matrigel in immunodeficient mice [47]. The resulting tumor is a large (~3 cm after 8 weeks of growth) mass of vascularized human tissue that contains cell-types from all three germ layers. The teratoma is very easy to grow and contains many complex and self-organized tissue structures that all arise spontaneously 86. This qualifies the teratoma as a unique and powerful launching pad for organ engineering.

Teratomas form from random differentiation of hPSCs to an assortment of cell-types; thus, to successfully use the teratoma for organ engineering, we need a way to sculpt the tissue towards a particular lineage. MicroRNAs (miRNAs) are small non-coding RNAs that are typically 20-26 nucleotides in length [48]. They act post-transcriptionally, targeting specific messenger RNAs (mRNAs) in a sequence-specific manner [49]. miRNAs have extensively been studied as powerful developmental regulators, and they typically act in feed-back or feed-forward loops, reinforcing specific gene programs that define a particular lineage [50]. Knowing miRNAs are endogenously expressed developmental regulators, here, we designed a suicide gene circuit that is based on miRNA gene expression. The presence of an endogenous miRNA will protect desired cells in the teratoma, while other undesired cells will trigger the apoptosis pathway, a process we dub "Teratoma Sculpting." As a proof of context, we demonstrate *in vivo* neural enrichment of hPSCs via mir9-mediated teratoma sculpting.

Results

hPSCs have a strong propensity to silence transgenes randomly integrated into the genome, especially in the context of *in vivo* differentiation. The *AAVS1* locus is known to serve as a safe-harbor genetic locus and is thus frequently targeted when performing hPSC genetic engineering. Consequently, we chose to knock-in our miRNA-mediated suicide genetic circuit into the *AAVS1* locus of hPSCs. However, considering the multitude of potential miRNA target combinations worth trying in the suicide gene circuit, we established a platform that allows for rapid *AAVS1* knock-in cell-line generation. This was accomplished by inserting an attP landing-pad BXB1 recombinase site, followed by an mCherry fluorescent reporter, into the *AAVS1* locus via CRISPR-Cas9 mediated gene targeting and homology directed repair (HDR) (**Figure 3.1A**). The BXB1 recombinase can conduct scar-free gene insertion upon recognition of both the attP landing-pad and attB donor sequence. In comparison to site-directed HDR via CRISPR-Cas9 gene targeting, the BXB1 recombinase-mediated gene engineering is highly efficient and reproducible. After successful *AAVS1* locus targeting of the attP landing-pad (**Figure 3.1A**), we were able to rapidly engineer cell lines by transfecting cells with a BXB1 recombinase encoding plasmid, and an attB donor plasmid (**Figure 3.2A**). Any gene circuit of choice can be cloned into the attB donor plasmid and used for rapid AAVS1 transgene integration (**Figure 3.2B**).

To molecularly sculpting teratoma using a miRNA-mediated suicide circuit we added miRNA target sequences to the 3' UTR of a suicide gene [51]. This suicide gene directs the expression of a drug inducible Caspase9 (iCapase9). When in the presence of the chemical dimerizer AP20187, the iCaspase9 will dimerize and trigger apoptosis in the cell [52] (**Figure 3.3A,C**). However, when a lineage-specific endogenous miRNA with the complementary miRNA

Figure 3.1: attP-mCherry Cell Line Generation.

(A) Schematic showing the targeting of the genomic AAVS1 locus in ESCs via electroporation of Cas9, gRNA, and a AAVS1 donor plasmid. These cells are single cell cloned, and then genotyped to check for proper transgene insertion. (B) Gel electrophoresis depicting the left and right arm amplicons of the AAVS1 integrated transgene.

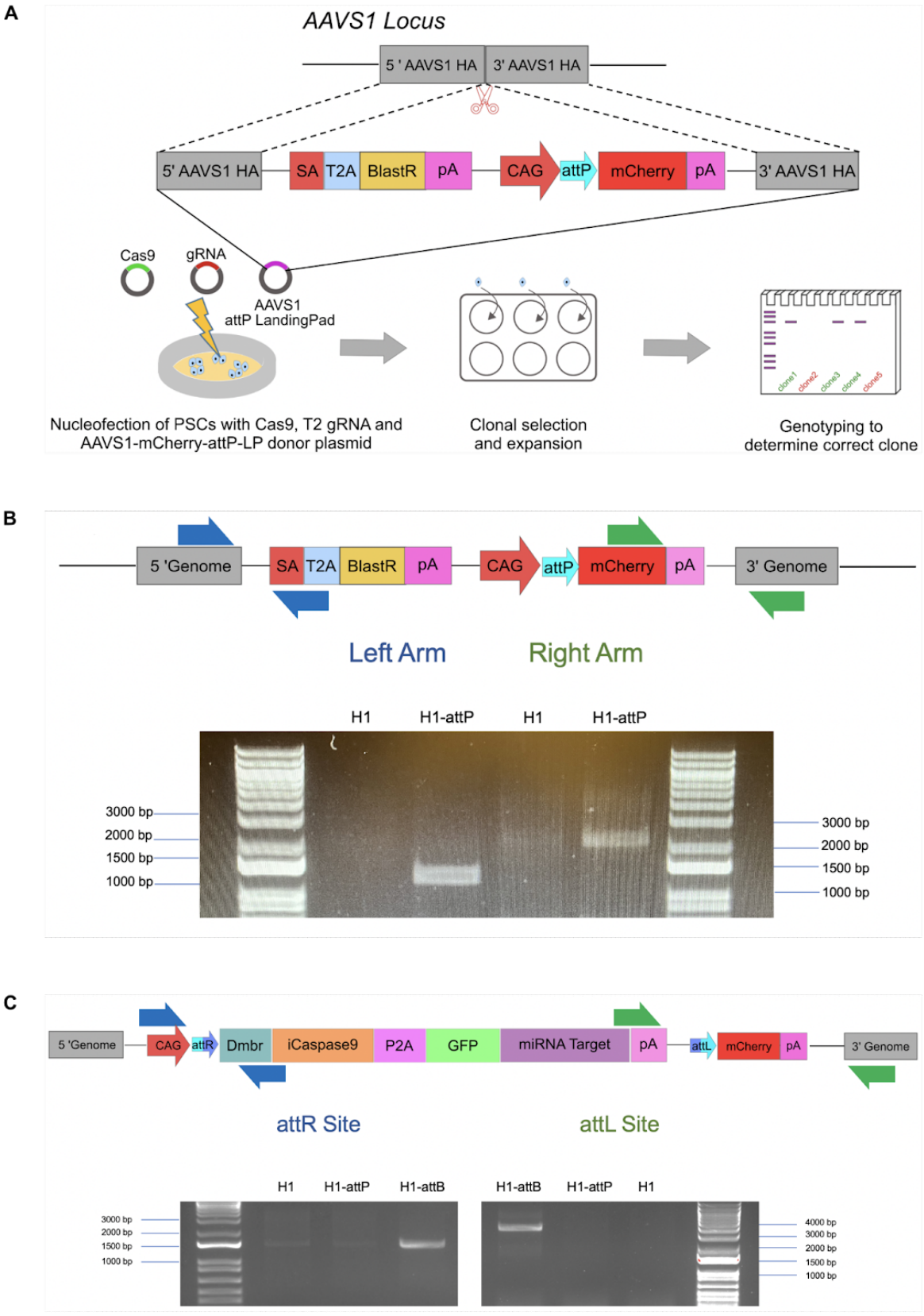
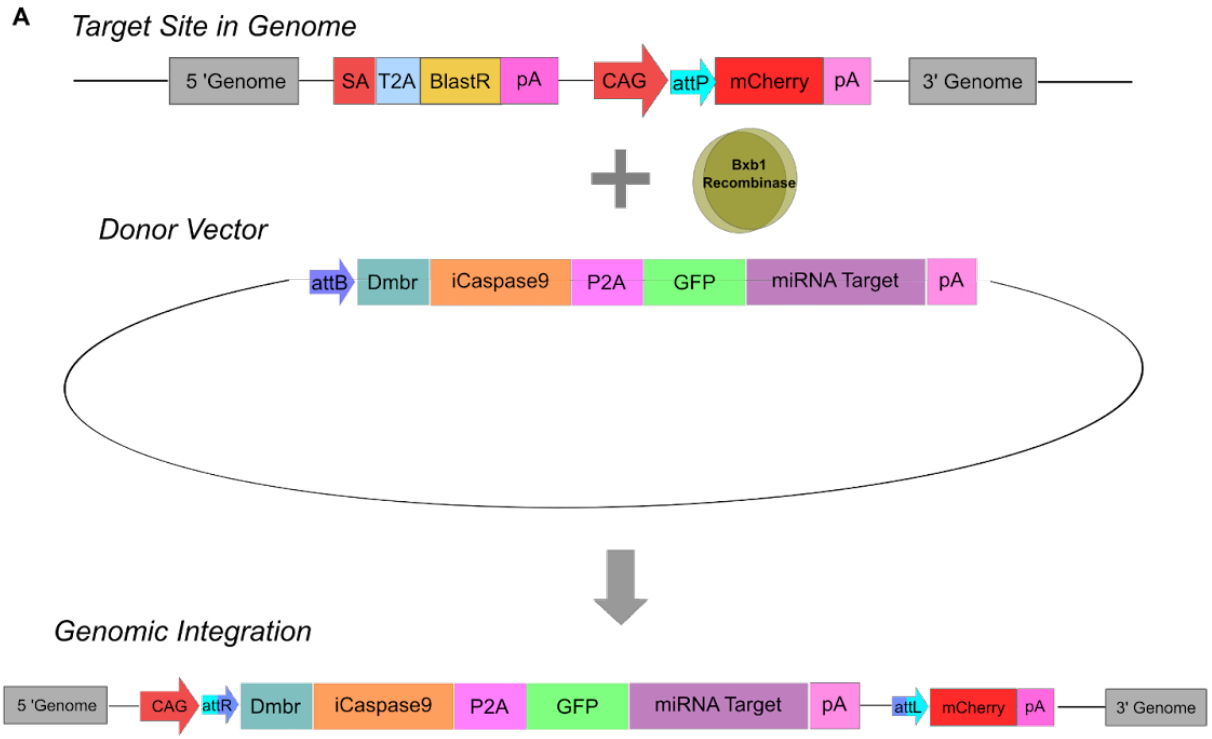

A
AAVS1 Locus
5 ' AAVS1 HA
3' AAVS1 HA
5' AAVS1 HA
SA
T2A
BlastR
pA
CAG
attP
mCherry
pA
3' AAVS1 HA
Cas9
gRNA
AAVS1
attP LandingPad
Nucleofection of PSCs with Cas9, T2 gRNA and AAVS1-mCherry-attP-LP donor plasmid
Clonal selection and expansion
clone1
clone2
clone3
clone4
clone5
Genotyping to determine correct clone
B
5 'Genome
SA
T2A
BlastR
pA
CAG
attP
mCherry
pA
3' Genome
Left Arm
Right Arm
H1
H1-attP
H1
H1-attP
3000 bp
2000 bp
1500 bp
1000 bp
3000 bp
2000 bp
1500 bp
1000 bp
C
5 'Genome
CAG
attR
Dmbr
iCaspase9
P2A
GFP
miRNA Target
pA
attL
mCherry
pA
3' Genome
attR Site
attL Site
H1
H1-attP
H1-attB
H1-attB
H1-attP
H1
3000 bp
2000 bp
1500 bp
1000 bp
4000 bp
3000 bp
2000 bp
1500 bp
1000 bp

Figure 3.2: attB Recombination Mechanism.

(A) Schematic showing the Bxb1-mediated recombination event, following introduction of Bxb1 recombinase and an attB donor plasmid. (B) Gel electrophoresis depicting the attR and attL recombination site amplicons of the after BXB1-mediated recombination into the AAVS1 locus.


A
Target Site in Genome
5 'Genome
SA
T2A
BlastR
pA
CAG
attP
mCherry
pA
3' Genome
Bxb1 Recombinase
Donor Vector
attB
Dmbr
iCaspase9
P2A
GFP
miRNA Target
pA
Genomic Integration
5 'Genome
CAG
attR
Dmbr
iCaspase9
P2A
GFP
miRNA Target
pA
attL
mCherry
pA
3' Genome

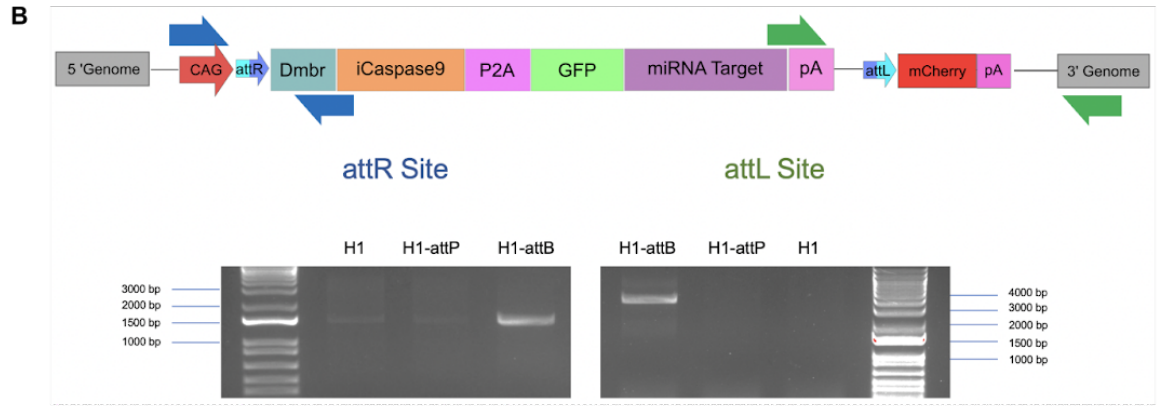
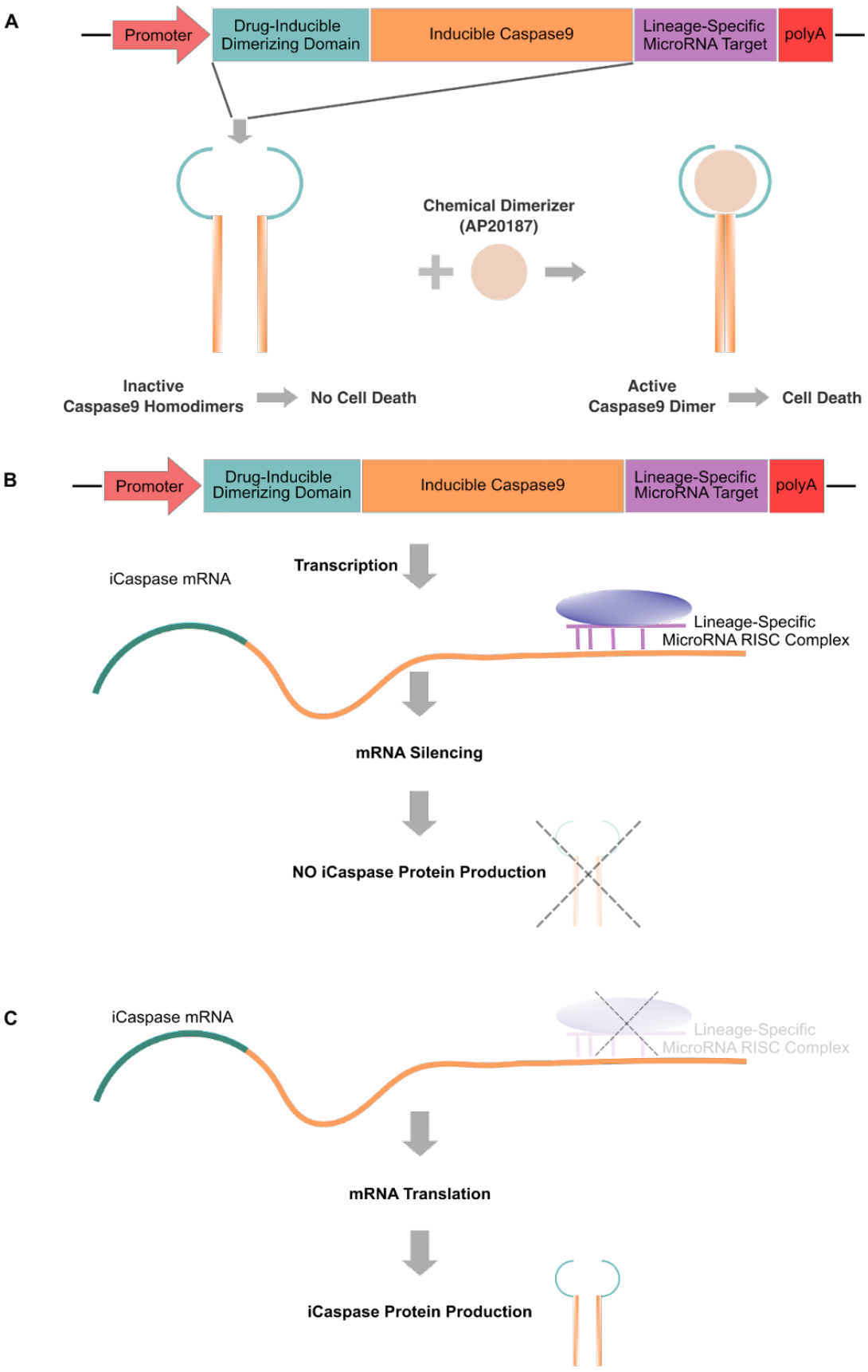

B
5 'Genome
CAG
attR
Dmbr
iCaspase9
P2A
GFP
miRNA Target
pA
attL
mCherry
pA
3' Genome
attR Site
attL Site
H1
H1-attP
H1-attB
H1-attB
H1-attP
H1
3000 bp
2000 bp
1500 bp
1000 bp
4000 bp
3000 bp
2000 bp
1500 bp
1000 bp

Figure 3.3: miRNA-mediated Inducible Caspase9 Construct Mechanism.

(A) miRNA-mediated inducible Caspase9 gene circuit. (B-C) Schematic showing the function of the microRNA-mediated Caspase9 circuit with (B) and without (C) an endogenous lineage-specific microRNA-RISC complex.


A
Promoter
Drug-Inducible Dimerizing Domain
Inducible Caspase9
Lineage-Specific MicroRNA Target
polyA
Chemical Dimerizer (AP20187)
Inactive Caspase9 Homodimers
No Cell Death
Active Caspase9 Dimer
Cell Death
B
Promoter
Drug-Inducible Dimerizing Domain
Inducible Caspase9
Lineage-Specific MicroRNA Target
polyA
iCaspase mRNA
Transcription
Lineage-Specific MicroRNA RISC Complex
mRNA Silencing
NO iCaspase Protein Production
C
iCaspase mRNA
Lineage-Specific MicroRNA RISC Complex
mRNA Translation
iCaspase Protein Production

target sequence is expressed, the miRNA-RISC complex will downregulate iCaspase9 protein expression, deeming the desired cell resistant to AP20187-induced apoptosis (**Figure 3.3B**). Thus, our miRNA-mediated apoptosis gene circuit allows for enrichment of desired cell-types within a multi-lineage environment, such as the teratoma.

Initially, we tested the killing mechanism of our iCaspase system after integrating the genetic construct into PSCs (**Figure 3.2B**). The first generated cell line lacks any miRNA target sites and is thus detonated as iCaspase-mirEmpty (**Figure 3.4A**). AP20187 was added to iCaspase-mirEmpty cells at various concentrations, and viability was assessed at 24 and 48 hours. We observed that concentrations above 1 nM AP20187 caused complete killing of PSCs after 48 hours of AP20187 incubation (**Figure 3.4B**).

Upon validating the effective of iCaspase9 killing upon addition of AP20187, we next assessed if our system could be selectively silenced via the addition of lineage-specific miRNAs in the 3'UTR of the iCaspase transgene. We transduced PSCs, which express mir302b at high levels, with a GFP-mir302bTarget encoding lentivirus, and compared GFP repression levels to PSCs transduced with GFP-mirEmpty control lentivirus via FACS (**Figure 3.5A**). This experiment was run in non-mir302b expression 293T cells in parallel. After lentivirus transduction, we significant GFP repression in GFP-mir302bTarget ESCs (**Figure 3.5B**), compared to GFP-mirEmpty ESCs (**Figure 3.5C**). As expected, GFP repression was not seen in 293Ts (**Figure 3.5D, E**), demonstrating that the addition of the mir302bTarget sequence into the 3' UTR of the GFP transgene can trigger transgene repression in a cell-type specific manner.

We next sought to demonstrate the functionality of our miRNA-mediated iCaspase apoptosis circuit in the content of *in vitro* cell differentiation. The pan-neural miRNA mir-9 has been extensively studied as a very important regulator in neural development and neurogenesis [53].

Figure 3.4: iCaspase-mirEmpty Kill Curve.

(A) Schematic of the iCaspase9-mirEmpty gene construct integrated into the AAVS1 locus of attp-mCherry cells after BXB1-mediated recombination. (B) Assessing cell viability after 24 and 48 hours of AP20187 exposure to iCaspase-mirEmpty cells.

A


No microRNA site
5 'Genome
CAG
attR
Dmbr
iCaspase9
P2A
GFP
pA
attL
mCherry
pA
3' Genome

B

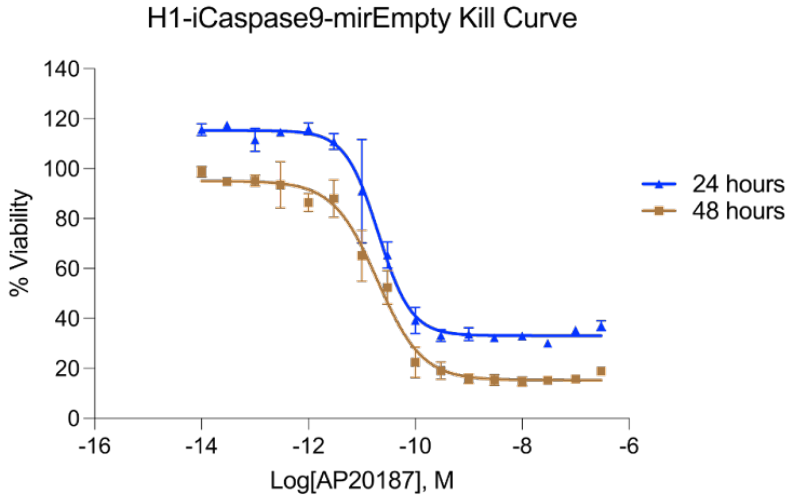
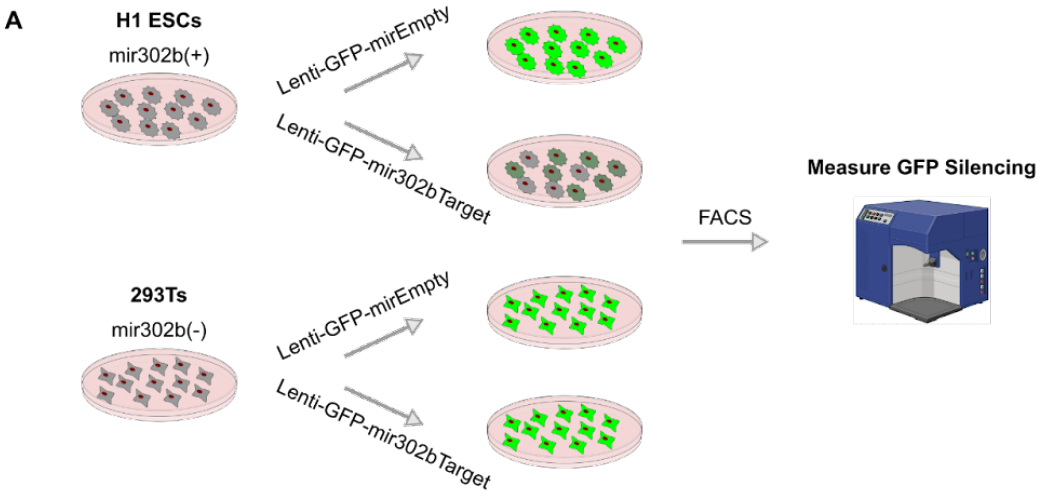
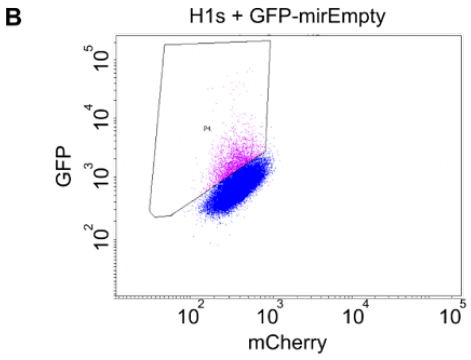
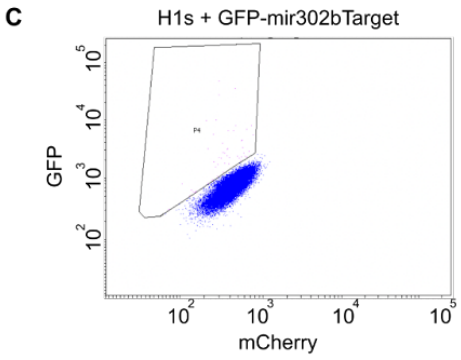

H1-iCaspase9-mirEmpty Kill Curve
% Viability
140
120
100
80
60
40
20
0
-16
-14
-12
-10
-8
-6
Log[AP20187], M
24 hours
48 hours

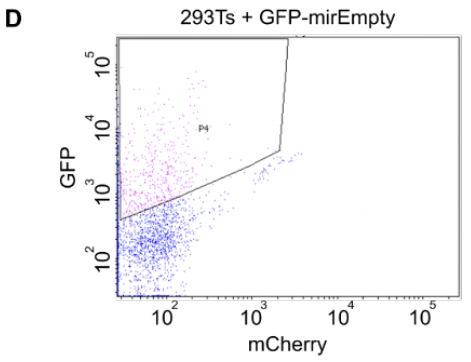
Figure 3.5: mir302b-mediated GFP Silencing in H1 ESCs.

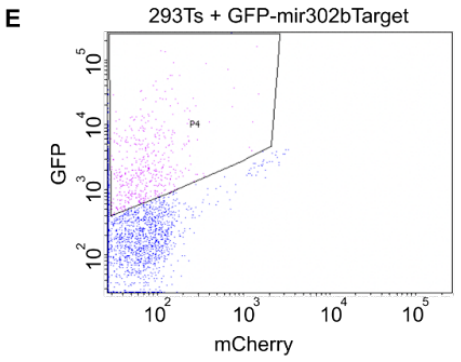
(A) Schematic depicting experimental workflow for assessing functionality of lineage-specific miRNA-mediated gene circuit silencing, in which a pluripotency miRNA - mir302b - target site is inserted in the 3' UTR of GFP. This is compared to a negative control, GFP-mirEmpty, in which there is no miRNA target site in the 3' UTR of GFP. (B-C) FACS plots for H1 ESCs transduced with lentivirus encoding for GFP-mirEmpty (B) and GFP-mir302b (C). (D-E) FACS plots for 293Ts transduced with lentivirus encoding for GFP-mirEmpty (D) and GFP-mir302b (E).


A
H1 ESCs
mir302b(+)
Lenti-GFP-mirEmpty
Lenti-GFP-mir302bTarget
293Ts
mir302b(-)
Lenti-GFP-mirEmpty
Lenti-GFP-mir302bTarget
FACS
Measure GFP Silencing


B
H1s + GFP-mirEmpty
GFP
mCherry


C
H1s + GFP-mir302bTarget
GFP
mCherry


D
293Ts + GFP-mirEmpty
GFP
mCherry


E
293Ts + GFP-mir302bTarget
GFP
mCherry

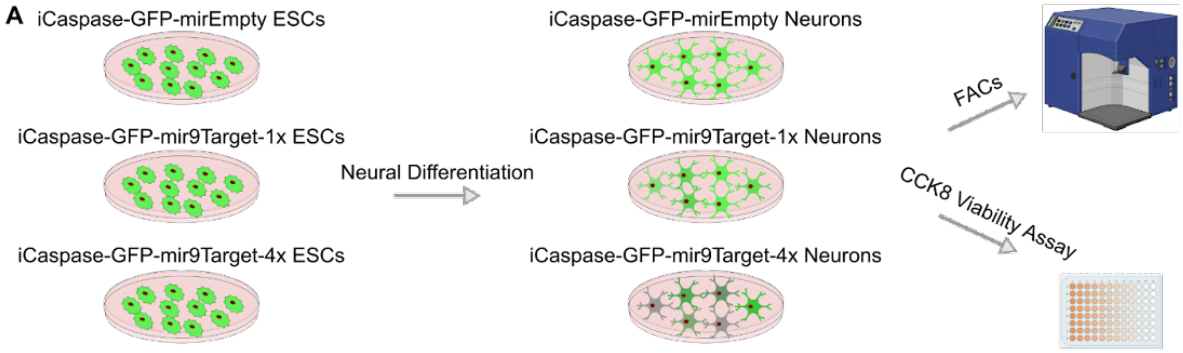
It has been implicated in neural stem cell proliferation, self-renewal, and differentiation [54]. Along with neural expression, mir9 is also known to be expressed in glial tissue, which justifies its use in teratoma enrichment assays, as more of the neural niche will theoretically be protected. Thus, we utilized the mir9 target sequence in a miRNA-regulated suicide gene expression circuit as a proof-of-concept for miRNA-mediated iCaspase9 repression and neural survival.

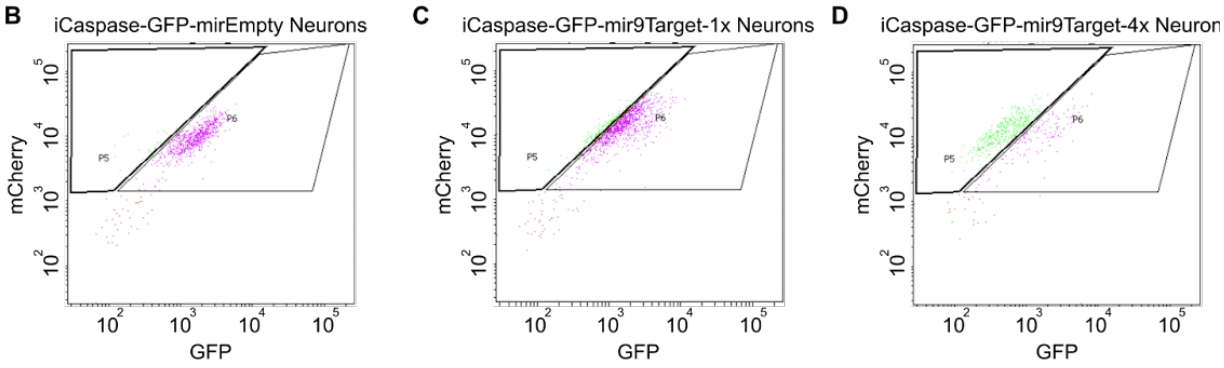
The effectiveness of mir9-mediated iCapsase9-GFP repression was assessed in two contexts: 1) 1X, in which the mir9 target site is inserted into the 3' UTR once; and 2) 4x, in which the mir9 target site is inserted into the 3' UTR four times. BXB1-mediated recombination in attP-mCherry cells was utilized to generated 3 H1 ESC cell lines: 1) iCaspase9-GFP-mirEmpty; 2) iCaspase9-GFP-mir9Target-1x; 3) iCaspase9-GFP-mi9Target-4x; followed by neural differentiation for 9 days, and finally FACS analysis and a CCK8 cell-viability assay (**Figure 3.6A**). FACS analysis demonstrated GFP repression in both mir9Target cell lines, with GFP repression being the most effective in the 4x condition (**Figure 3.6C, D, E**). In parallel, neurons were treated with AP20187 for one day, and then tested for viability. In accordance with the observed GFP repression, iCaspase9-GFP-mi9Target-4x neurons were significantly more viable than iCaspase9-GFP-mi9Target-1x neurons, while iCaspase9-GFP-mi9Empty neurons were completely dead after one day of AP20187 treatment (**Figure 3.6F**). These results demonstrate the efficacy of mir9-mediated iCaspase9 transcript knockdown in the context of PSC-derived neurons.

Motivated by the ability to enrich for neural cell-types in vitro using the mir9-mediated apoptosis circuit (**Figure 3.6**), we next attempted to achieved neural enrichment in the context of the teratoma. Because of the superiority of the iCaspase9-GFP-mi9Target-4x cell line, in regards to mir9-mediated transgene repression, we injected these cells into immunodeficient mice in order to form teratomas (**Figure 3.7A**). We grew 8 teratoma mice, injected half daily with AP20187 for

Figure 3.6: mir9-mediated iCaspase9 Silencing in H1-derived Neurons.

(A) Schematic depicting experimental workflow for assessing functionality of lineage-specific miRNA-mediated gene circuit silencing, in which a neural miRNA - mir9 - target site is inserted in the 3' UTR of iCaspase-GFP. This target site is inserted once (1x) or 4 times in tandem (4x). These are compared to a negative control, GFP-mirEmpty, in which there is no miRNA target site in the 3' UTR of iCaspase-GFP. (B) FACS plots for iCaspase-GFP-mirEmpty neurons. (C) FACS plots for iCaspase-GFP-mir9Target-1x neurons. (D) FACS plots for iCaspase-GFP-mir9Target-4x neurons. (E) CCK8 viability of neurons after one day of iCaspase induction via 10 nm of AP20187.


A
iCaspase-GFP-mirEmpty ESCs
iCaspase-GFP-mir9Target-1x ESCs
iCaspase-GFP-mir9Target-4x ESCs
Neural Differentiation
iCaspase-GFP-mirEmpty Neurons
iCaspase-GFP-mir9Target-1x Neurons
iCaspase-GFP-mir9Target-4x Neurons
FACs
CCK8 Viability Assay


B
iCaspase-GFP-mirEmpty Neurons
C
iCaspase-GFP-mir9Target-1x Neurons
D
iCaspase-GFP-mir9Target-4x Neuron
mCherry
GFP
P5
P6

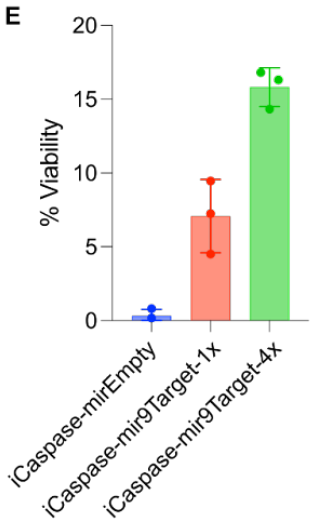
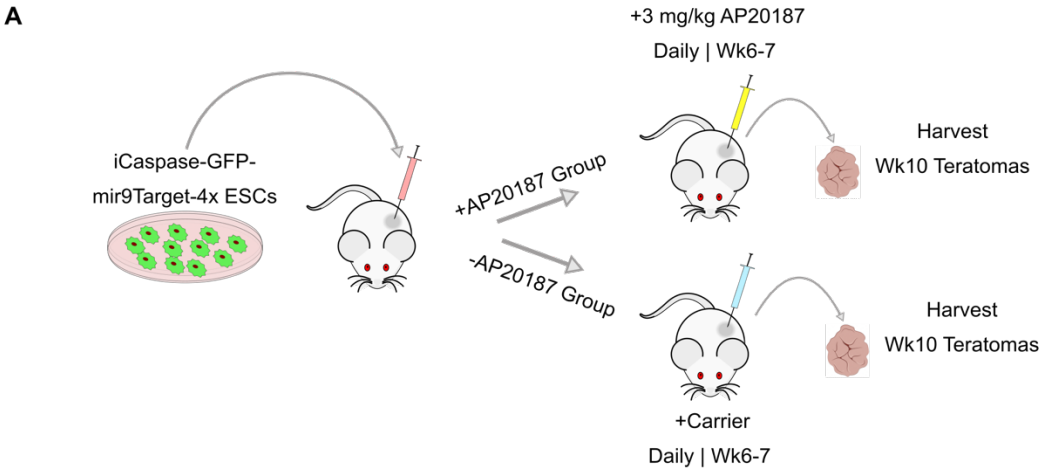
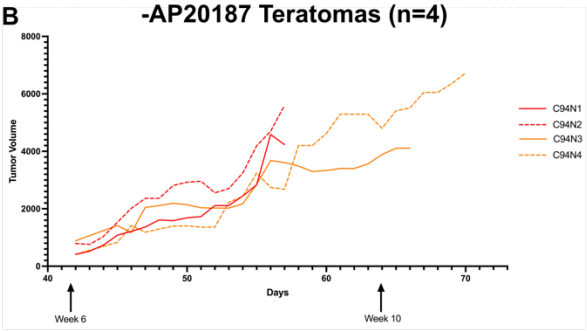
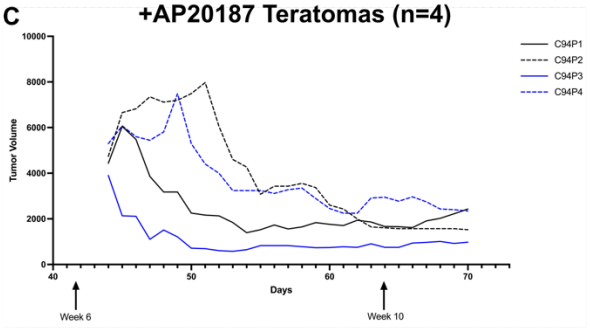

E
% Viability
20
15
10
5
0
iCaspase-mirEmpty
iCaspase-mir9Target-1x
iCaspase-mir9Target-4x

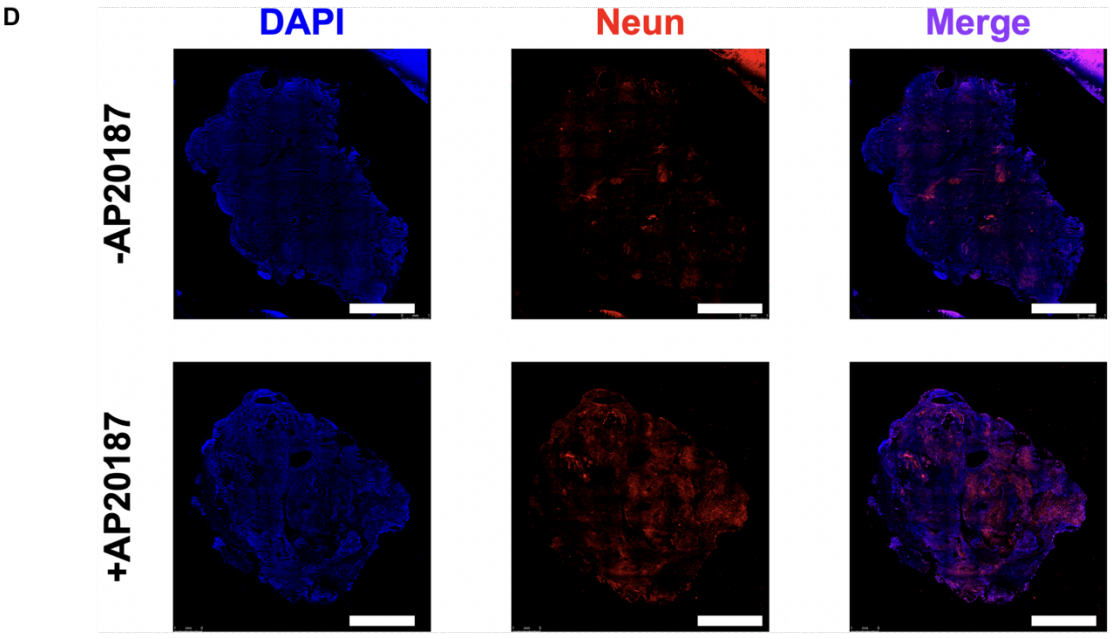
Figure 3.7: mir9-mediated Neural Enrichment in Teratomas.

(A) Schematic depicting experimental workflow for neural enrichment in iCaspase-mir9Target-4x Teratomas. (B) Growth curves for -AP20187 teratomas (n=4). (C) Growth curves for +AP20187 teratomas (n=4). (D) Representative immunostains stains from -AP20187 and +AP20187 teratomas. Teratoma were stained with DAPI and NEUN to visualize nuclei and neural cells, respectively. (Scale bar = 2 mm)


A
+3 mg/kg AP20187
Daily | Wk6-7
iCaspase-GFP-
mir9Target-4x ESCs
+AP20187 Group
-AP20187 Group
Harvest
Wk10 Teratomas
Harvest
Wk10 Teratomas
+Carrier
Daily | Wk6-7


B
-AP20187 Teratomas (n=4)
Tumor Volume
Days
Week 6
Week 10
C94N1
C94N2
C94N3
C94N4


C
+AP20187 Teratomas (n=4)
Tumor Volume
Days
Week 6
Week 10
C94P1
C94P2
C94P3
C94P4


D
DAPI
Neun
Merge
-AP20187
+AP20187

one week starting at week 6, and finally harvested all teratomas at week 10 (**Figure 3.7A**). Non-AP20187 mice received a week of daily carrier injections starting at week 6, like AP20187 mice. While non-AP20187 teratomas show continuous growth through all the 10 weeks of teratoma development (**Figure 3.7B**), growth curves from AP20187 teratomas show significant tumor volume reduction following injection of the chemical dimerizer (**Figure 3.7C**). Finally, teratomas were harvested after 10 weeks of in vivo growth, sectioned and immunostained to check for neural enrichment. Compared to -AP20187 teratomas, NEUN antibody staining showed significant enrichment of neural cells in the +AP20187 teratomas (**Figure 3.7D**), as seen by the higher percentage of NEUN+ cells.

Methods

Cell Culture

The H1 hPSC cell line was maintained under feeder-free conditions in mTeSR medium (Stem Cell Technologies). Prior to passaging, tissue-culture plates were coated with growth factor-reduced Matrigel (Corning) diluted in DMEM/F-12/Glutamax medium (Thermo Fisher Scientific), and incubated for 30 minutes at 37C, 5% CO2. Cells were dissociated and passaged using the dissociation reagent Versene (Thermo Fisher Scientific). Cells were passaged a maximum of 4 times for proper expansion prior to injection.

Attp-mCherry Cell Line Generation

The sgRNA targeting AAVS1 locus of the human genome (spacer sequence GGGCCACTAGGGACAGGAT) was cloned into the lentiCRISPR v2 plasmid (Addgene #52961). To generate the knockin donor plasmid, we cloned the CAG promoter followed by a cassette of attP-mCherry into the pCR4-Blunt-TOPO vector (Thermo Fisher Scientific). Two homology arms were amplified from upstream (804 bp) and downstream (837 bp) of the sgRNA targeting site in AAVS1 genomic locus and constructed into the donor plasmid flanking the CAG-

mCherry cassette. Between the upstream homology arm and the CAG promoter, we inserted a splice acceptor sequence followed by a T2A linked blasticidin resistance gene. H1 ESCs were electroporated using a 4D-Nucleofector system and P3 Primary Cell X kit (Lonza) according to the manufacturer's instructions. Briefly, the H1s were dissociated into single cells. 1E6 cells were mixed with 100 uL nucleofection reagents and 10 ug DNA (5 ug attP-mCherry donor + 5 ug lentiCRISPR v2) and electroporated. The cells were recovered with pre-warmed medium and then cultured on inactivated MEF feeders in 10 cm dishes with mTeSR medium supplemented with 0.5 mM ROCK-inhibitor. Afterward, the mTeSR medium without ROCK-inhibitor was refreshed daily. 2 mg/ml blasticidin were added into the culture medium 7 days after electroporation. The cells were cultured without passage until clones emerged on the plate. The clones were checked under the microscope and those with mCherry expression were picked up and expanded individually. To detect genomic integration, the genomic DNA from cultured cells was extracted using DNeasy Blood & Tissue Kits (QIAGEN). Approximately 500 ng of genomic DNA was used for each PCR reaction using KAPA HiFi HotStart Ready Mix (Kapa Biosystems). The PCR amplification of the left and right arm utilized primers that amplified regions spanning both the PGP1 AAVS1 endogenous locus and the engineered cassette.

The primer sequences are listed below:

Table 3.1: Genotyping Primers

Left Arm Forward Primer	ACTTCCCCTCTTCCGATGTTG
Left Arm Reverse Primer	ATTGTAGCCGTTGCTCTTTCA
Right Arm Forward Primer	GACTACACCATCGTGGAACAGTACGA

Table 3.1: Genotyping Primers

Right Arm Reverse Primer	CTGCCTGGAGAAGGATGCAGGA

Attb Plasmid and BXB1-mediated Cell Line Generation

Promotor-less attB plasmid constructs were generated by digesting pCAG-GFP (Addgene #11150) with SalI and EagI restriction enzymes, and cloning in attB-iCaspase-P2A-GFP cassettes via Gibson Assembly (NEB). The iCaspase insert was amplified from pMSCV-F-del Casp9.IRES.GFP (Addgene #15567). H1 ESCs were electroporated using 4D-Nucleofector system and P3 Primary Cell X kit (Lonza) according to the manufacturer's instruction. Briefly, the H1s were dissociated into single cells. 1E6 cells were mixed with 100 uL nucleofection reagents and 10 ug DNA (5 ug pCAG-NLS-HA-Bxb1 (Addgene #51271) + 5 ug attB-iCaspase-P2A-GFP) and electroporated. Cells were expanded, single-cell cloned, and genotyped for AAVS1 recombination insertion.

Neural Differentiation

Inducible NeuroD1 cell lines were generated and differentiated as previously described [19].

AP20187 Kill Curve and Dosing

AP20187 (Tocris) was reconstituted in DMSO. Concentrations ranging from 10 uM - 0.1 pM were deposited on single-cell cloned H1-attB-iCaspase cells. Subsequently, a CCK8 growth assay was performed to assess cell viability.

AP20187 *in vivo* Administration

For in vivo experiments, AP20187 was injected at 3 mg/kg via intraperitineal injections.

Teratoma Generation

A subcutaneous injection of 5-10 million PSCs in a slurry of Matrigel and mTeSR medium (1:1) was made in the right flank of anesthetized Rag2/;gc/ immunodeficient mice. Weekly

monitoring of teratoma growth was made by quantifying approximate elliptical area with the use of calipers measuring outward width and height.

Chapter 4 - Charting and Probing the Activity of ADARs in Human Development and Cell-fate Specification.

Abstract

Adenosine deaminases acting on RNA (ADARs) impact diverse cellular processes and pathological conditions. While we possess important insights into their roles in adult tissues, their functions in early cell fate specification remain less understood. To address this, we devised a comprehensive framework to investigate ADARs in human development. We began by charting time-course RNA editing profiles in human organs from fetal to adult stages, enabling broad insights into RNA editing trends across diverse tissues. Next, we utilized hPSC differentiation to experimentally probe ADARs across specific tissue-types, harnessing brain organoids as neural specific, and teratomas as pan-tissue developmental models. We found that time-series teratomas faithfully recapitulated developmental trends observed in fetal tissue. Motivated by this, we conducted pan-tissue, single-cell CRISPR-KO screens of ADARs in teratomas to assess their role in cell-fate specification. Knocking out *ADAR1* led to a global decrease in RNA editing across all germ-layers. Intriguingly, we observed a significant fitness defect in mesodermal tissues, and an enrichment of adipogenic cells, revealing a novel role for *ADAR1* in mesenchymal differentiation. Collectively, we developed a multi-pronged framework charting time-resolved RNA editing profiles in fetal and fetal-like organ tissues, thereby shedding light on the role of ADARs in development and cell fate-specification.

Introduction

Adenosine-to-inosine (A-to-I) editing is a critical post-transcriptional modification that alters RNA nucleotides, is catalyzed by adenosine deaminases acting on RNA (ADAR) enzymes, and is one of the most prominent epitranscriptomics modification in metazoans [55–61]. There are 3 human ADAR enzymes: *ADAR1*, *ADARB1*, and *ADARB2*. *ADAR1* is constitutively expressed

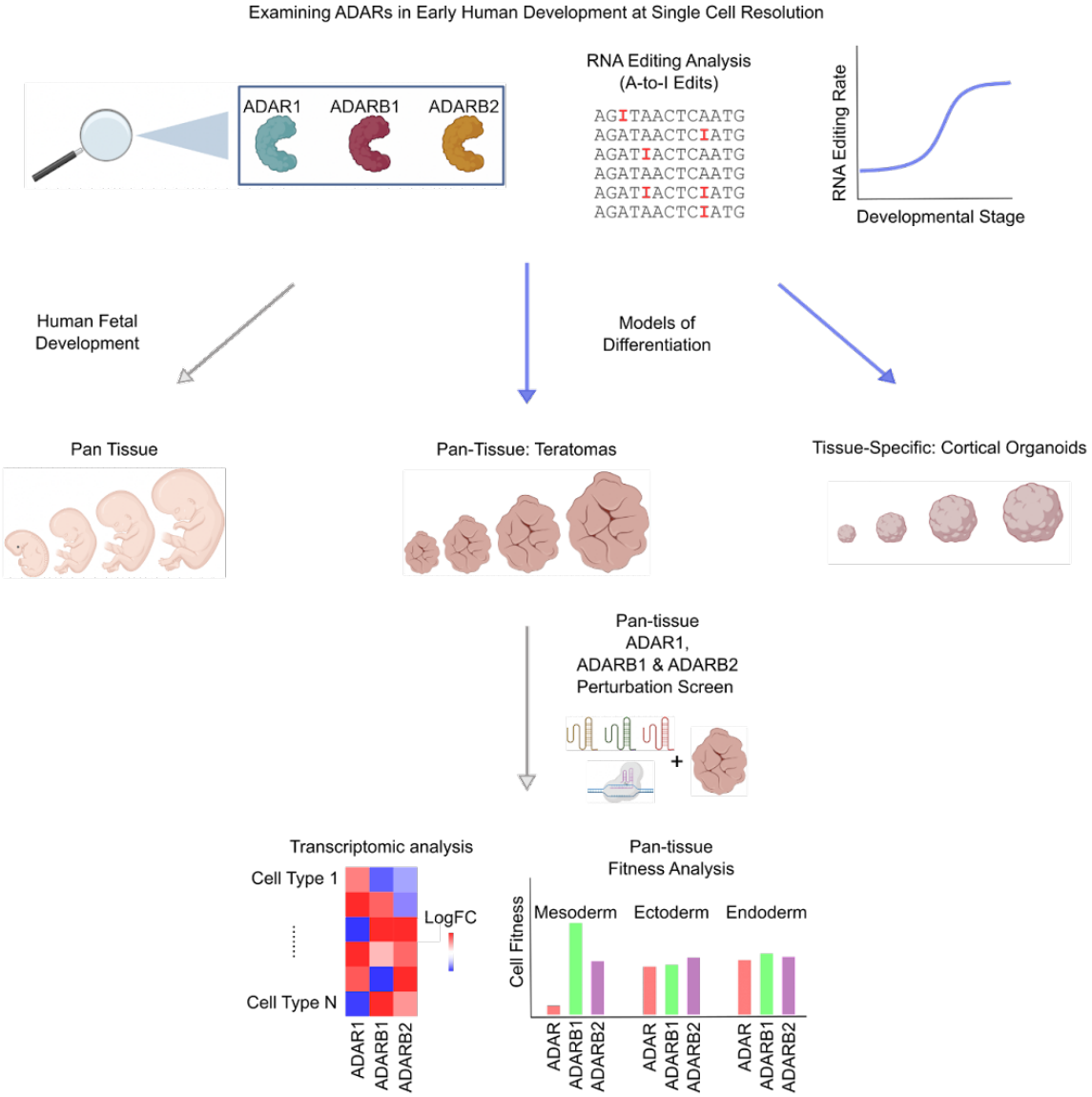
across most tissues and is mostly known to edit repetitive RNA [62]. *ADARB1* is also expressed in most tissues, with the highest expression in the CNS, and known to mostly edit coding regions [62,63]. *ADARB2* is exclusively expressed in the CNS and contains a catalytically inactive deaminase domain; thus, it does not edit RNA and is mostly known to play a regulatory role in RNA editing [62,64]. A-to-I RNA editing leads to the incorporation of inosines, which are interpreted as guanosines by cellular machinery. Consequently, A-to-I editing can yield unique protein isoforms [63,65,66] and also influence miRNA [67] and mRNA stability [68], splicing [69], and localization[70]. Furthermore, the dysregulation of ADAR proteins, both via their RNA editing-dependent and -independent [71,72] functions, have been linked to various diseases, such as cancer [73], neurological disorders [74,75], viral infection [76], metabolic disorders [77], and auto-immune malignancies [78].

Systematic investigations into RNA editing profiles in adult human tissue have documented editing sites across tissues and uncovered key regulators in RNA editing [66,79,80]. In addition to their roles in adult tissues, ADAR enzymes are pivotal regulators of post-transcriptional RNA processing during various developmental stages [80–84]. In mice, *Adar1*-KO [71,85,86] and *Adarb1*-KO [87,88] result in embryonic and perinatal lethality, respectively, highlighting their putative role in developmental processes. However, due to these lethality phenotypes, mouse models are limited in their capacity to systematically study ADAR's role in cell fate specification. Lastly, given the prevalence of A-to-I editing in primate-specific *Alu* elements [89], mouse models may not capture the human-specific roles that ADARs have in human development.

Taken together, while significant insights exist regarding ADARs' roles in mouse models and adult human tissues, a knowledge gap remains concerning their impact on early human organogenesis. Towards this, we established here a multi-pronged framework (**Figure 4.1**) to systematically map and functionally interrogate their role in human development. We first charted

Figure 4.1: Chapter 4 Graphical Abstract

Our graphical abstract illustrates this study's workflow. We begin by analyzing RNA editing and ADAR expression dynamics across diverse tissue types, utilizing both human organ data and hPSC-derived developmental models. Subsequently, we employed the teratoma for a pan-tissue ADAR perturbation screen, probing the functional roles of ADAR proteins across all three germ layers.


Examining ADARs in Early Human Development at Single Cell Resolution
ADAR1
ADARB1
ADARB2
RNA Editing Analysis
(A-to-I Edits)
AGITAACTCAATG
AGATAACTCIATG
AGATIACTCAATG
AGATAACTCAATG
AGATIACTCIATG
AGATAACTCIATG
RNA Editing Rate
Developmental Stage
Human Fetal
Development
Models of
Differentiation
Pan Tissue
Pan-Tissue: Teratomas
Tissue-Specific: Cortical Organoids
Pan-tissue
ADAR1,
ADARB1 & ADARB2
Perturbation Screen
+
Transcriptomic analysis
Cell Type 1
Cell Type N
LogFC
ADAR1
ADARB1
ADARB2
Pan-tissue
Fitness Analysis
Cell Fitness
Mesoderm
Ectoderm
Endoderm
ADAR
ADARB1
ADARB2

time-course RNA editing profiles in five different human organ tissues, from fetal to adult stages. Next, to experimentally probe ADARs, we utilized two human pluripotent stem cell (hPSC) derived models of development: cerebral organoids, *in vitro* derived brain tissue-specific constructs; and teratoma tissues, *in vivo* derived pan-tissue constructs [90]. The multilineage nature of the teratoma allowed the analysis of RNA editing across all germ-layers in a single experiment, motivating a deeper investigation into teratoma tissue as a developmental model for RNA editing. By generating time-series teratomas, we observed that neural tissue in teratomas replicates transcriptomic and epitranscriptomic developmental patterns evident in fetal cerebral tissue. Moreover, a pan-tissue, single-cell CRISPR-KO screen in teratomas unveiled *ADAR1*'s integral role across all germ layers, with a distinct emphasis on adipogenic cell-fate determination, suggesting *ADAR1*'s potential implication in obesity-related phenotypes.

This study provides a pan-tissue, single-cell-level investigation into the spatiotemporal profiles and functional roles of ADARs during human fetal development. By investigating a range of fetal and fetal-like tissues, we elucidate the dynamic RNA editing landscape, casting light on its implications in human development and disease.

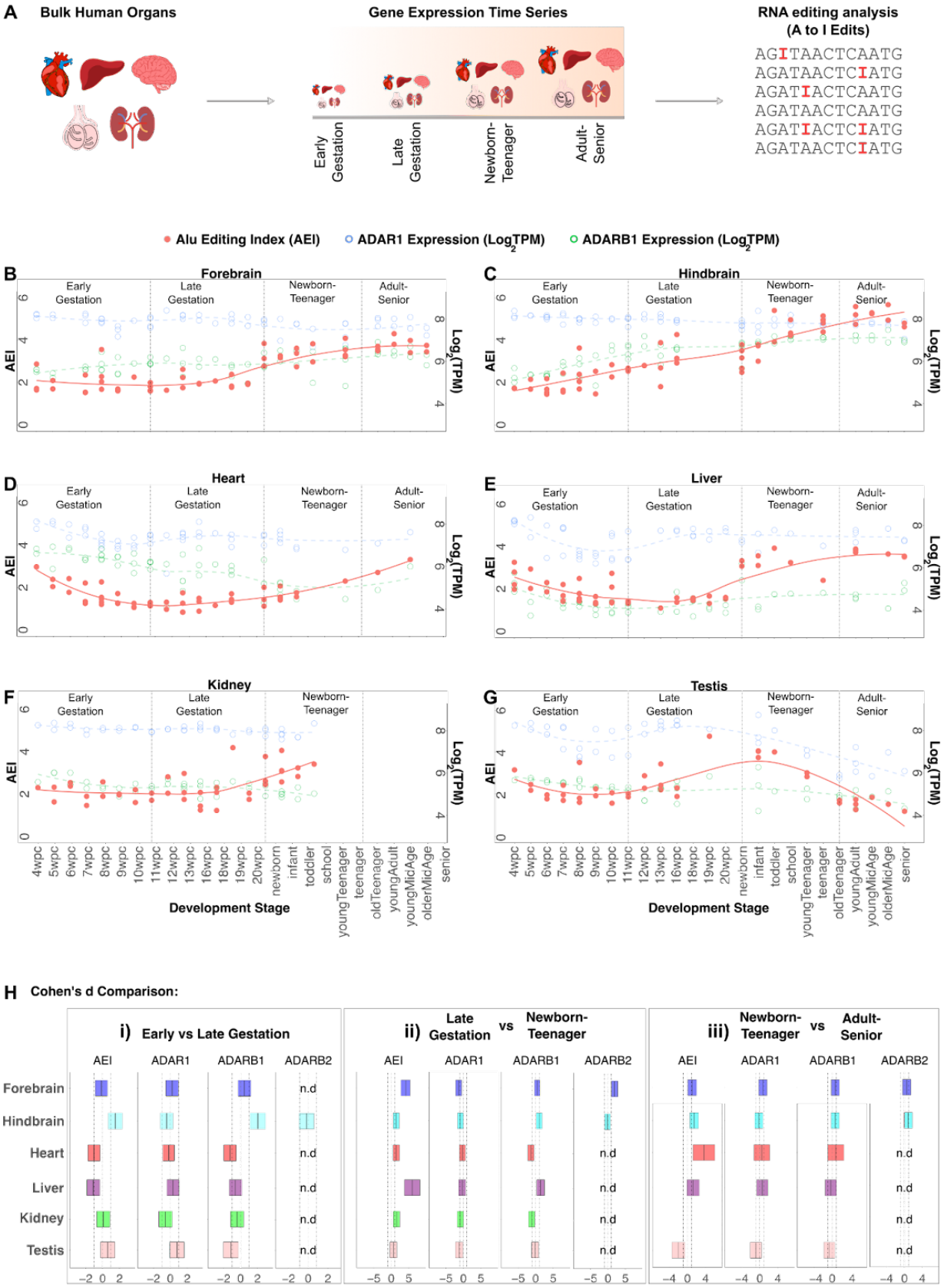
Results

RNA Editing Analysis in Bulk Human Organs

We first employed a time-resolved mammalian gene expression database [91] to systematically investigate temporal RNA editing levels in multiple human organs, including the forebrain, hindbrain, heart, liver, kidney and testis. We categorized the data into four developmental stages: early gestation (4-10 weeks post conception, wpc), late gestation (11-20 wpc), newborn-teenager (0-19 years old), and adult-senior (25-63 years old) (**Figure 4.2A)**.

Figure 4.2: Time series Bulk RNA Editing Analysis in Human Fetal Tissues.

(A) Schematic depicting the workflow for this investigation into RNA editing and ADAR expression dynamics in bulk organ tissues across lifespan. Organ time-series data is binned into 4 developmental groups - early gestation (4-10 wpc), late gestation (11-20 wpc), newborn-to-teenager (0-19 years of age), and adult-to-senior (25-63 years of age) - for conducting comparisons across sequential developmental periods. (B-G) AEI levels (solid red circles), as well as *ADAR1* (open blue circles) and *ADARB1* (open green circles) expression values, are charted throughout lifespan for the forebrain (B), hindbrain (C), heart (D), liver (E), kidney (F), and testis (G). (H) Cohen's d comparison between sequential time groups for AEI, *ADAR1*, *ADARB1*, and *ADARB2*, across all organs (not detected = n.d).


A
Bulk Human Organs
Gene Expression Time Series
RNA editing analysis
(A to I Edits)
AGITAACTCAATG
AGATAACTCIATG
AGATIACTCAATG
AGATAACTCAATG
AGATIACTCIATG
AGATAACTCIATG
Early Gestation
Late Gestation
Newborn-Teenager
Adult-Senior
Alu Editing Index (AEI)
ADAR1 Expression (Log2TPM)
ADARB1 Expression (Log2TPM)
B Forebrain
C Hindbrain
D Heart
E Liver
F Kidney
G Testis
AEI
Log2(TPM)
4wpc
5wpc
6wpc
7wpc
8wpc
9wpc
10wpc
11wpc
12wpc
13wpc
16wpc
18wpc
19wpc
20wpc
newborn
infant
toddler
school
youngTeenager
teenager
oldTeenager
youngAdult
youngMidAge
olderMidAge
senior
Development Stage
H Cohen's d Comparison:
i) Early vs Late Gestation
ii) Late Gestation vs Newborn-Teenager
iii) Newborn-Teenager vs Adult-Senior
AEI
ADAR1
ADARB1
ADARB2
Forebrain
Hindbrain
Heart
Liver
Kidney
Testis
n.d

A-to-I editing events are predominantly observed in *Alu* elements – repetitive DNA sequences exclusive to the primate genome [89]. Given its robustness, and especially as it accounts for the editing activity in low covered regions, while avoiding the need to quantify the editing level per-site, we assessed the global editing levels in bulk human organs via the recently described Alu editing index (AEI) [92] . This is obtained by calculating the ratio of A-to-G mismatches to the total adenosine coverage within Alu regions. We then measured the AEI shifts and ADAR enzyme expression dynamics across successive developmental stages.

In the forebrain, we observed a significant rise in AEI (p-value = 1.24E-10, Cohen's D = 4.12) during the late gestation to the newborn-teenager transition, accompanied by a decrease in *ADAR1* (p-value = 2.54E-3, Cohen's D = -1.35) and an increase in *ADARB2* (p-value = 1.86E-05, Cohen's D = 1.91) (**Figure 4.2B**). Furthermore, in the hindbrain we observe an upward AEI trend throughout all of lifespan (**Figure 4.2C**). We observe concordant significant increases between both the AEI and *ADARB1* expression in the early-to-late gestation transition – AEI (p-value = 7.72E-04, Cohen's D = 1.50) and *ADARB1* (p-value = 1.17E-06, Cohen's D = 2.04) – and from late gestation to the newborn-teenager transition – AEI (p-value = 4.57E-04, Cohen's D = 1.42) and *ADARB1* (p-value = 1.50E-02, Cohen's D = 0.99) – which suggests *ADARB1* is driving the increase in AEI during hindbrain development (**Figure 4.2C**). These findings are consistent with a recent report tracking RNA editing in developing human brain tissue [80].

In the heart, we observe pronounced AEI fluctuations throughout development (**Figure 4.2D**). Initially, in the early-to-late gestation transition, there is a notable drop in AEI (p-value = 2.73E-03, Cohen's D = -1.03), followed by an increase in AEI during the late gestation to newborn-teenager transition (p-value = 4.56E-04, Cohen's D = 1.43). (**Figure 4.2D**). This initial decrease and subsequent increase in AEI are concordant with drops in *ADARB1* expression during both the

early-to-late gestation transition (p-value = 2.15E-03, Cohen's D = -1.35) and the late gestation to newborn-teenager transition (p-value = 4.56E-04, Cohen's D = 1.43) (**Figure 4.2D**).

The liver exhibited very dynamic shifts in AEI across lifespan (**Figure 4.2E**), with an initial decrease in AEI during the early-to-late gestation transition (p-value = 3.21E-04, Cohen's D = -1.11), a robust increase in AEI during the late gestation to the newborn-teenager transition (p-value = 3.28E-05, Cohen's D = 6.00), and finally another significant increase in AEI during the newborn-teenager to adult-senior transition (p-value = 4.40E-02, Cohen's D = 1.26). Again, the initial drop and subsequent rise in AEI were concordant with decreases (p-value = 4.83E-02, Cohen's D = -0.60) and subsequent increases (p-value = 4.36E-02, Cohen's D = 1.43) in *ADARB1* expression, respectively.

In the kidney, we observed a significant increase in AEI during the late gestation to newborn-teenager transition (p-value = 4.77E-04, Cohen's D = 1.51), coinciding with a significant drop in *ADARB1* (p-value = 1.25E-02, Cohen's D = -1.10) (**Figure 4.2F**).

In the testis we observed a robust AEI reduction (p-value = 9.94E-05, Cohen's D = -5.21) during the newborn-teenager to adult transition, which is concordant with a robust and significant decrease in *ADAR1* expression (p-value = 3.33E-03, Cohen's D = -2.19) (**Figure 4.2G**).

Intrigued by a notable drop in heart and liver AEI during early gestation (**Figure 4.2D, E**), we further investigated this trend by dividing the early gestation period into two phases: 4-7 wpc and 8-11 wpc. During this developmental shift, we observed a significant reduction in AEI in heart tissue (p-value = 4.35E-03, Cohen's D = -1.80), which is concordant with a significant decrease in *ADAR1* (p-value = 1.81E-05, Cohen's D = -2.64). In the liver, we also noted a substantial decrease in AEI (p-value = 2.13E-02, Cohen's D = -1.00), accompanied by significant reductions

in both *ADAR1* (p-value = 1.79E-04, Cohen's D = -1.82) and *ADARB1* (p-value = 9.75E-3, Cohen's D = -1.24) expression.

In summary, there are few convergent trends with respect to AEI and ADAR expression dynamics across tissues. *ADARB1* expression dynamics closely follows the progressive increase in AEI throughout all of hindbrain development, and a drop in AEI during very early gestation is accompanied by a drop in *ADAR1* expression in both heart and liver tissue. After examining RNA editing trends throughout human development in bulk tissues, we next shifted our attention to a single-cell investigation of RNA editing at prenatal stages to uncover cell-type specific trends that may govern early organogenesis. To explore this, we employed lab-grown, hPSC-derived model systems, complemented with human fetal tissue data [93].

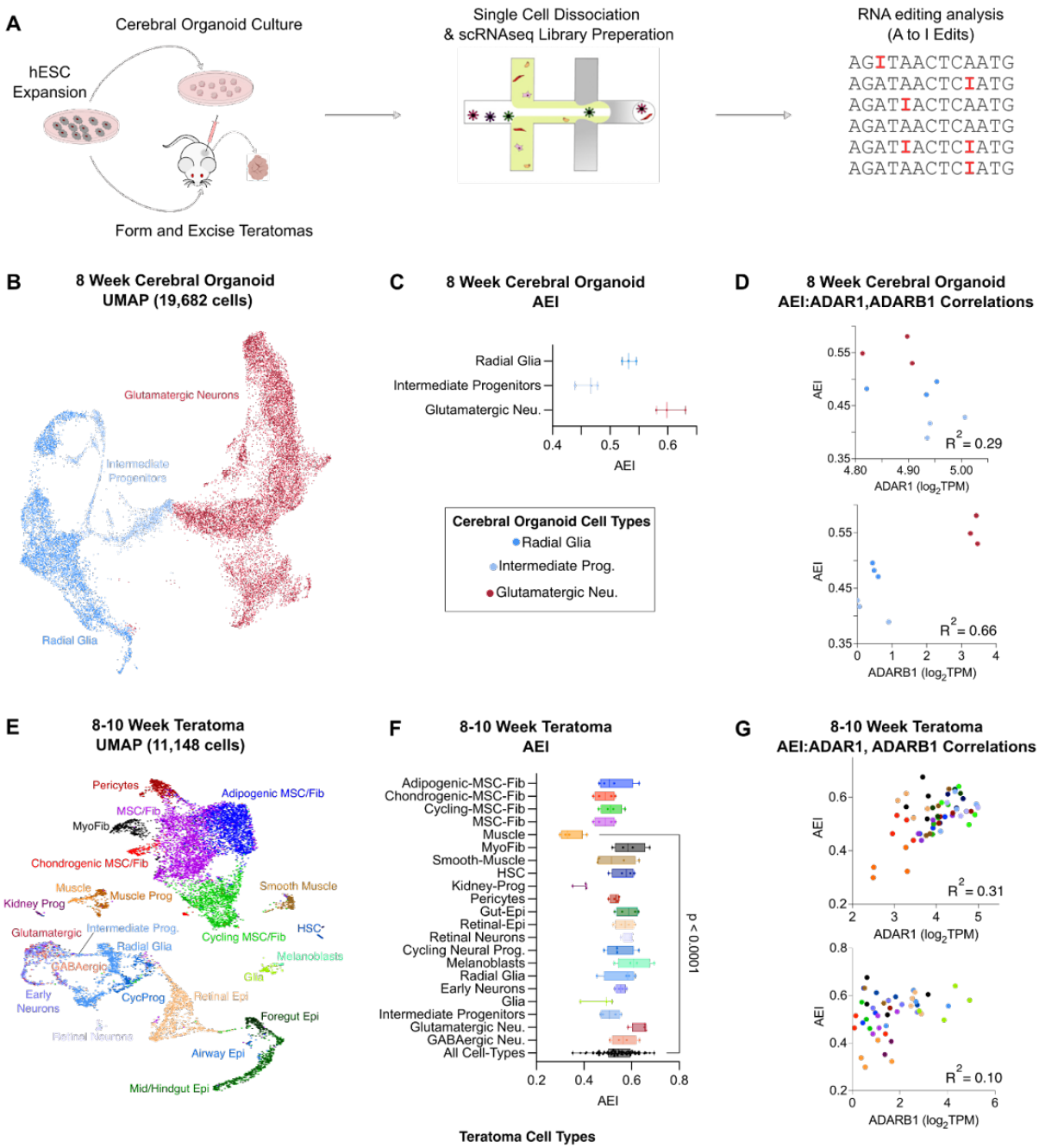
Single-Cell RNA Editing Analysis in hPSC-derived Model Systems

To gain deeper insights into the significance of A-to-I editing during early human tissue development, we conducted RNA editing analysis on 8-week-old cerebral organoids and 8-10 week-old teratomas, both derived from human pluripotent stem cells (**Figure 4.3A**). As compared to bulk RNA sequencing, single-cell RNA sequencing data is characterized by more limited coverage and a 3' bias due to the poly-A based capture probes utilized via the 10X Chromium platform[94]. Thus, we decided to utilize the robust AEI index to more accurately estimate global editing levels.

We started our investigation with cerebral organoids and observed that 8-week-old cerebral organoids were composed of 3 main cell types: Radial Glia, Intermediate Progenitors, and Glutamatergic Neurons (**Figure 3.3B**). We measured AEI trends across these cell-types (**Figure 3.3C**). To gain insight into the pan-tissue influence of ADARs, we next considered the teratoma as an additional model system. Among neural lineages, we observed 8–10-week-old teratomas

Figure 4.3: Single Cell RNA Editing Analysis in hESC-derived Model Systems.

(A) Workflow schematic for conducting single cell RNA editing analysis on hESC-derived cerebral organoids and teratomas. (B) UMAP plot from 8-week cerebral organoids processed through the single cell RNA sequencing pipeline. (C) AEI values for all major 8-week cerebral organoid cell-types, with each data point calculated as a pseudo-bulk value from each cerebral organoid cell-type (n=3 cerebral organoids). (D) Correlation between AEI to *ADAR1* expression (top) and to *ADARB1* expression (bottom) for all 8-week cerebral organoid cell-types. (E) Aggregated UMAP plot from 4 H1 teratoma processed through the single cell RNA sequencing pipeline. (F) Box-and-whiskers plot showing the AEI values for all major teratoma cell-types, with each data point calculated as a pseudo-bulk value (n=4 WT teratomas). (G) Correlation between AEI value to *ADAR1* expression (top) and to *ADARB1* expression (bottom) for all teratoma cell-types.


A
Cerebral Organoid Culture
hESC Expansion
Form and Excise Teratomas
Single Cell Dissociation & scRNAseq Library Preperation
RNA editing analysis (A to I Edits)
AGITAACTCAATG
AGATAACTCIATG
AGATIACTCAATG
AGATAACTCAATG
AGATIACTCIATG
AGATAACTCIATG
B
8 Week Cerebral Organoid UMAP (19,682 cells)
Glutamatergic Neurons
Intermediate Progenitors
Radial Glia
C
8 Week Cerebral Organoid AEI
Radial Glia
Intermediate Progenitors
Glutamatergic Neu.
AEI
Cerebral Organoid Cell Types
Radial Glia
Intermediate Prog.
Glutamatergic Neu.
D
8 Week Cerebral Organoid AEI:ADAR1,ADARB1 Correlations
AEI
ADAR1 (log2TPM)
R2 = 0.29
ADARB1 (log2TPM)
R2 = 0.66
E
8-10 Week Teratoma UMAP (11,148 cells)
Pericytes
Adipogenic MSC/Fib
MSC/Fib
MyoFib
Chondrogenic MSC/Fib
Muscle
Muscle Prog
Kidney Prog
Smooth Muscle
Intermediate Prog.
Glutamatergic
Radial Glia
Cycling MSC/Fib
HSC
GABAergic
Melanoblasts
Glia
Early Neurons
CycProg
Retinal Epi
Foregut Epi
Retinal Neurons
Airway Epi
Mid/Hindgut Epi
F
8-10 Week Teratoma AEI
Adipogenic-MSC-Fib
Chondrogenic-MSC-Fib
Cycling-MSC-Fib
MSC-Fib
Muscle
MyoFib
Smooth-Muscle
HSC
Kidney-Prog
Pericytes
Gut-Epi
Retinal-Epi
Retinal Neurons
Cycling Neural Prog.
Melanoblasts
Radial Glia
Early Neurons
Glia
Intermediate Progenitors
Glutamatergic Neu.
GABAergic Neu.
All Cell-Types
p < 0.0001
AEI
G
8-10 Week Teratoma AEI:ADAR1, ADARB1 Correlations
AEI
ADAR1 (log2TPM)
R2 = 0.31
ADARB1 (log2TPM)
R2 = 0.10
Teratoma Cell Types
Mesoderm
Adipogenic MSC/Fib
MSC/Fib
Muscle
Kidney Prog
Chondrogenic MSC/Fib
MyoFib
Pericytes
Immune
Cycling MSC/Fib
Muscle Prog
Smooth Muscle
Endoderm
Foregut Epi
Mid/Hindgut Epi
Airway Epi
Ectoderm
Retinal Neurons
Early Neurons
Retinal Epi
GABAergic Neu.
Glia
CycProg
Shared Cell Types With Cerebral Organoids
Intermediate Prog.
Glutamatergic Neu.
Radial Glia

included Radial Glia, Intermediate Progenitors, Glutamatergic Neurons and also GABAergic neurons, and across all tissues composed of over 20 distinct cell-types representative of all three germ layers (**Figure 4.3E**)[90]. We confirmed these cell-types exhibit a strong correlation with their corresponding fetal counterparts (**Figure 4.4A**). As expected, *ADAR1* exhibited ubiquitous expression across all teratoma cell-types, while *ADARB2* was neural-specific and *ADARB1* was most highly expressed in neural tissues, and at lower levels elsewhere (**Figure 4.4D**). Moreover, we assessed the reliability of teratoma single-cell RNA sequencing libraries for RNA editing analysis by comparing single-cell editing rates with those from bulk teratoma RNA sequencing libraries, revealing a high correlation between these two library generation and sequencing modalities (**Figure 4.4B**). Similar cell-type diversity and range of AEI levels were observed when AEI analysis was conducted across 3 alternative PSC lines (**Figure 4.5**).

To quantify RNA editing levels in unique teratoma cell-types, we created pseudo-bulk samples by pooling single cell RNA count matrices from matching cell-types together (refer Methods). Our observations revealed varying levels of AEI across teratoma cell types (**Figure 4.3F**). Notably, we observe a significantly lower AEI in muscle cells compared to the average teratoma AEI (**Figure 4.3F**), which is consistent with reports examining RNA editing across adult human bulk tissue samples [62,92]. Furthermore, *ADAR1* and *ADARB1* expression explained 31% and 10% of the variance in teratoma cell-type AEI, respectively (**Figure 4.3G**), which is also consistent with adult human bulk tissue sample reports [62].

In summary, the teratoma presents a promising model for RNA editing analysis due to its rich panoply of cell-types spanning all three germ-layers and its notable AEI-to-ADAR correlation trends among these cell-types. Importantly, generating the teratoma only required 8-10 weeks

Figure 4.4: Single Cell RNA Editing Analysis in WT Teratomas.

(A) Transcriptomic correlation between teratoma cell-types and cells from the Human Fetal Cell Atlas database. (B) Correlation between bulk teratoma and single cell teratoma RNA sequencing editing levels across repetitive editing sites. (C) Marker gene expression for all teratoma cell-types. (D) *ADAR1*, *ADARB1*, and *ADARB2* expression levels for all teratoma cell-types.

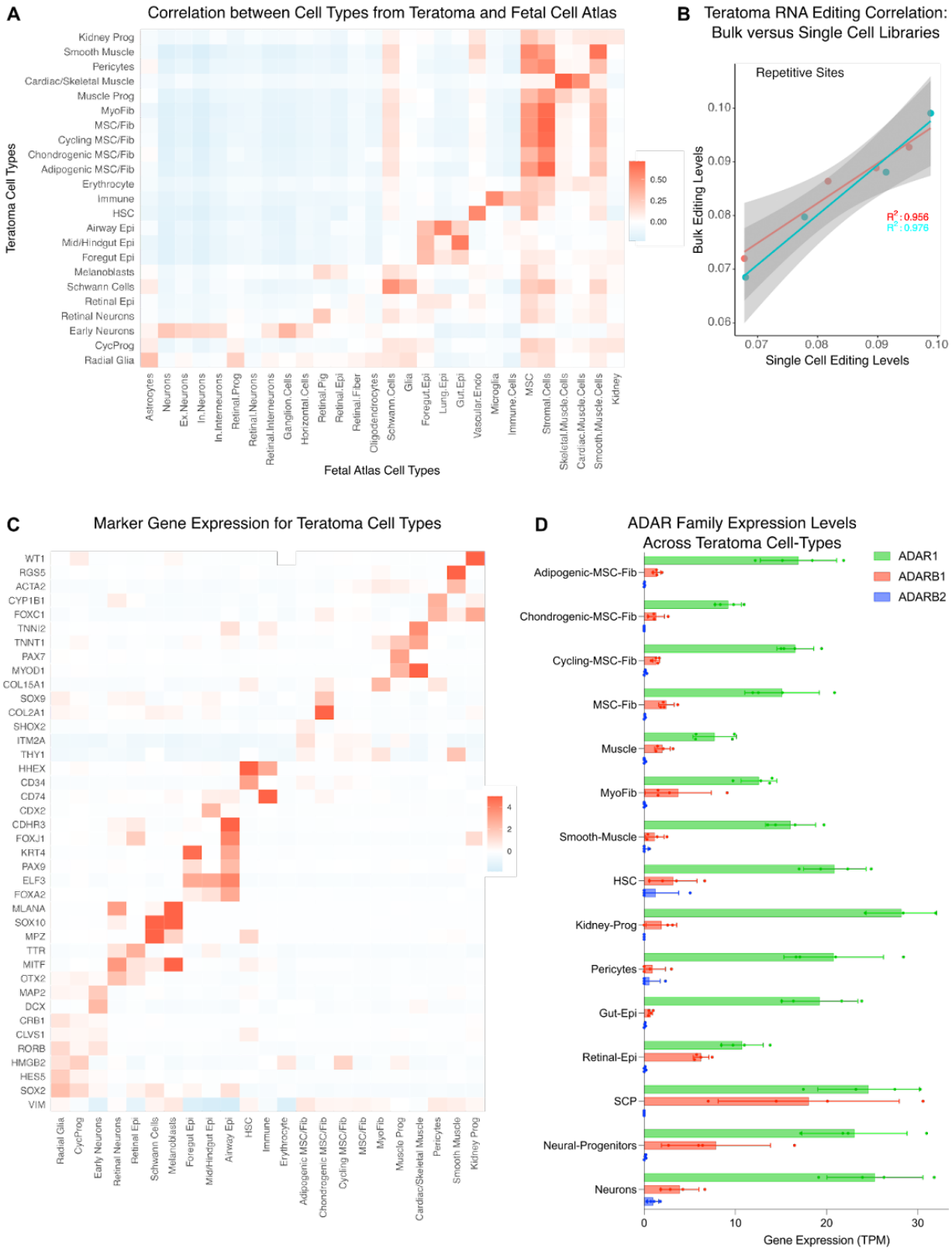
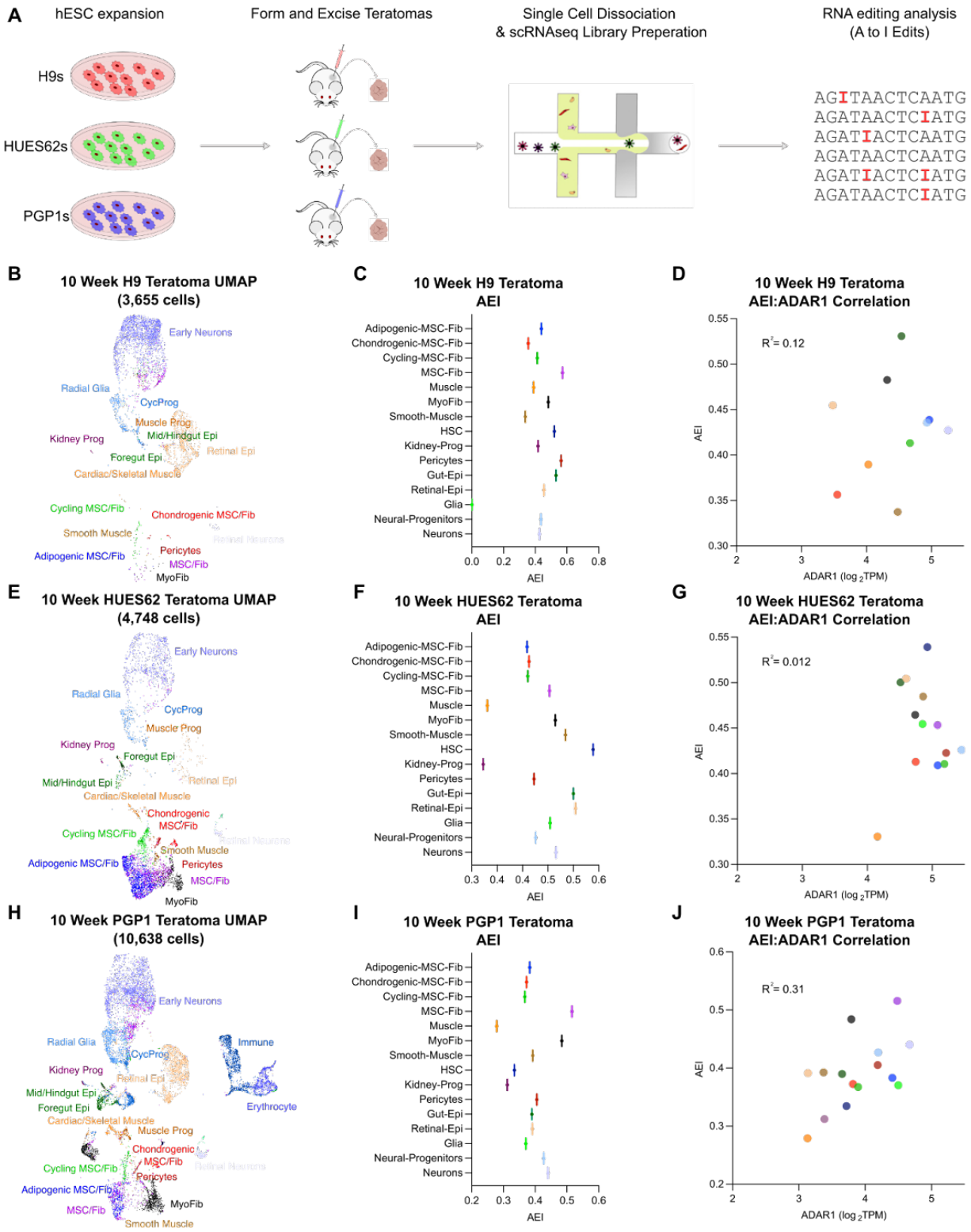

A
Correlation between Cell Types from Teratoma and Fetal Cell Atlas
Teratoma Cell Types
Kidney Prog
Smooth Muscle
Pericytes
Cardiac/Skeletal Muscle
Muscle Prog
MyoFib
MSC/Fib
Cycling MSC/Fib
Chondrogenic MSC/Fib
Adipogenic MSC/Fib
Erythrocyte
Immune
HSC
Airway Epi
Mid/Hindgut Epi
Foregut Epi
Melanoblasts
Schwann Cells
Retinal Epi
Retinal Neurons
Early Neurons
CycProg
Radial Glia
0.50
0.25
0.00
Astrocytes
Neurons
Ex.Neurons
In.Neurons
In.Interneurons
Retinal.Prog
Retinal.Neurons
Retinal.Interneurons
Ganglion.Cells
Horizontal.Cells
Retinal.Pig
Retinal.Epi
Retinal.Fiber
Oligodendrocytes
Schwann.Cells
Glia
Foregut.Epi
Lung.Epi
Gut.Epi
Vascular.Endo
Microglia
Immune.Cells
MSC
Stromal.Cells
Skeletal.Muscle.Cells
Cardiac.Muscle.Cells
Smooth.Muscle.Cells
Kidney
Fetal Atlas Cell Types
B
Teratoma RNA Editing Correlation:
Bulk versus Single Cell Libraries
Repetitive Sites
Bulk Editing Levels
0.10
0.09
0.08
0.07
0.06
R^2: 0.956
R^2: 0.976
0.07
0.08
0.09
0.10
Single Cell Editing Levels
C
Marker Gene Expression for Teratoma Cell Types
WT1
RGS5
ACTA2
CYP1B1
FOXC1
TNNI2
TNNT1
PAX7
MYOD1
COL15A1
SOX9
COL2A1
SHOX2
ITM2A
THY1
HHEX
CD34
CD74
CDX2
CDHR3
FOXJ1
KRT4
PAX9
ELF3
FOXA2
MLANA
SOX10
MPZ
TTR
MITF
OTX2
MAP2
DCX
CRB1
CLVS1
RORB
HMGB2
HES5
SOX2
VIM
4
2
0
Radial Glia
CycProg
Early Neurons
Retinal Neurons
Retinal Epi
Schwann Cells
Melanoblasts
Foregut Epi
Mid/Hindgut Epi
Airway Epi
HSC
Immune
Erythrocyte
Adipogenic MSC/Fib
Chondrogenic MSC/Fib
Cycling MSC/Fib
MSC/Fib
MyoFib
Muscle Prog
Cardiac/Skeletal Muscle
Pericytes
Smooth Muscle
Kidney Prog
D
ADAR Family Expression Levels
Across Teratoma Cell-Types
ADAR1
ADARB1
ADARB2
Adipogenic-MSC-Fib
Chondrogenic-MSC-Fib
Cycling-MSC-Fib
MSC-Fib
Muscle
MyoFib
Smooth-Muscle
HSC
Kidney-Prog
Pericytes
Gut-Epi
Retinal-Epi
SCP
Neural-Progenitors
Neurons
0
10
20
30
Gene Expression (TPM)

Figure 4.5: RNA editing and ADAR expression analysis in alternative PSC lines.

(A) Workflow schematic for generating teratomas from various PSC lines and conducting downstream single-cell RNA editing analysis across alternative lines. (B) UMAP plot from H9 teratoma processed through the single cell RNA sequencing pipeline. (C) AEI values for H9 teratoma cell-types, with each data point calculated as a pseudo-bulk value from each H9 teratoma cell-type (n=1 teratoma). (D) Correlation between AEI value to *ADAR1* expression for all H9 teratoma cell-types. (E) UMAP plot from HUES62 teratoma processed through the single cell RNA sequencing pipeline. (F) AEI values for HUES62 teratoma cell-types, with each data point calculated as a pseudo-bulk value from each HUES62 teratoma cell-type (n=1 teratoma). (G) Correlation between AEI value to *ADAR1* expression for all HUES62 teratoma cell-types. (H) UMAP plot from PGP1 teratoma processed through the single cell RNA sequencing pipeline. (I) AEI values for PGP1 teratoma cell-types, with each data point calculated as a pseudo-bulk value from each PGP1 teratoma cell-type (n=1 teratoma). (J) Correlation between AEI value to *ADAR1* expression for all PGP1 teratoma cell-types.


A
hESC expansion
Form and Excise Teratomas
Single Cell Dissociation
& scRNAseq Library Preperation
RNA editing analysis
(A to I Edits)
H9s
HUES62s
PGP1s
AGITAACTCAATG
AGATAACTCIATG
AGATIACTCAATG
AGATAACTCAATG
AGATIACTCIATG
AGATAACTCIATG
B
10 Week H9 Teratoma UMAP
(3,655 cells)
Early Neurons
Radial Glia
CycProg
Muscle Prog
Kidney Prog
Mid/Hindgut Epi
Foregut Epi
Retinal Epi
Cardiac/Skeletal Muscle
Cycling MSC/Fib
Chondrogenic MSC/Fib
Smooth Muscle
Retinal Neurons
Adipogenic MSC/Fib
Pericytes
MSC/Fib
MyoFib
C
10 Week H9 Teratoma
AEI
Adipogenic-MSC-Fib
Chondrogenic-MSC-Fib
Cycling-MSC-Fib
MSC-Fib
Muscle
MyoFib
Smooth-Muscle
HSC
Kidney-Prog
Pericytes
Gut-Epi
Retinal-Epi
Glia
Neural-Progenitors
Neurons
0.0 0.2 0.4 0.6 0.8
AEI
D
10 Week H9 Teratoma
AEI:ADAR1 Correlation
R²= 0.12
AEI
0.55 0.50 0.45 0.40 0.35 0.30
2 3 4 5
ADAR1 (log 2TPM)
E
10 Week HUES62 Teratoma UMAP
(4,748 cells)
Early Neurons
Radial Glia
CycProg
Muscle Prog
Kidney Prog
Foregut Epi
Retinal Epi
Mid/Hindgut Epi
Cardiac/Skeletal Muscle
Chondrogenic
MSC/Fib
Cycling MSC/Fib
Retinal Neurons
Smooth Muscle
Adipogenic MSC/Fib
Pericytes
MSC/Fib
MyoFib
F
10 Week HUES62 Teratoma
AEI
Adipogenic-MSC-Fib
Chondrogenic-MSC-Fib
Cycling-MSC-Fib
MSC-Fib
Muscle
MyoFib
Smooth-Muscle
HSC
Kidney-Prog
Pericytes
Gut-Epi
Retinal-Epi
Glia
Neural-Progenitors
Neurons
0.3 0.4 0.4 0.5 0.5 0.6
AEI
G
10 Week HUES62 Teratoma
AEI:ADAR1 Correlation
R²= 0.012
AEI
0.55 0.50 0.45 0.40 0.35 0.30
2 3 4 5
ADAR1 (log 2TPM)
H
10 Week PGP1 Teratoma UMAP
(10,638 cells)
Early Neurons
Radial Glia
CycProg
Immune
Kidney Prog
Retinal Epi
Mid/Hindgut Epi
Erythrocyte
Foregut Epi
Cardiac/Skeletal Muscle
Muscle Prog
Chondrogenic
MSC/Fib
Cycling MSC/Fib
Retinal Neurons
Pericytes
Adipogenic MSC/Fib
MyoFib
MSC/Fib
Smooth Muscle
I
10 Week PGP1 Teratoma
AEI
Adipogenic-MSC-Fib
Chondrogenic-MSC-Fib
Cycling-MSC-Fib
MSC-Fib
Muscle
MyoFib
Smooth-Muscle
HSC
Kidney-Prog
Pericytes
Gut-Epi
Retinal-Epi
Glia
Neural-Progenitors
Neurons
0.2 0.3 0.4 0.5 0.6
AEI
J
10 Week PGP1 Teratoma
AEI:ADAR1 Correlation
R²= 0.31
AEI
0.6 0.5 0.4 0.3 0.2
2 3 4 5
ADAR1 (log 2TPM)
Teratoma Cell Types
Mesoderm
Adipogenic MSC/Fib
MSC/Fib
Muscle
Kidney Prog
Chondrogenic MSC/Fib
MyoFib
Pericytes
Immune
Cycling MSC/Fib
Muscle Prog
Smooth Muscle
Endoderm
Foregut Epi
Mid/Hindgut Epi
Airway Epi
Ectoderm
Retinal Neurons
Early Neurons
Glia
Retinal Epi
Radial Glia
CycProg

compared to the 5-months needed to generate cerebral organoids with comparable neural cell-type diversity[95]. Given these favorable characteristics, we decided to proceed with the teratoma for our subsequent investigations.

RNA Editing Analysis in Time Series Teratomas

To investigate spatiotemporal RNA editing trends in teratoma cell types and determine their correlations with human fetal tissue, we produced teratomas at four distinct developmental stages: 4, 6, 8, and 10 weeks. We then constructed single nucleus RNA sequencing (snRNA-seq) libraries and conducted RNA editing analysis across these samples (**Figure 4.6A**).

First, we confirmed that teratomas from the same developmental stage strongly correlated with each other (**Figure 4.7A**). We then examined how cell-type proportions change during teratoma development. Notably, early-stage teratomas exhibited a higher proportion of germ-layer progenitors, while later-stage teratomas were predominantly composed of their differentiated progeny (**Figure 4.7B-F**). For instance, in the ectoderm, week 4 teratomas were primarily composed of radial glia, constituting about 65% of the ectodermal cell types, while this proportion gradually decreased to approximately 35% in week 10 teratomas (**Figure 4.7C**). Conversely, the relative proportion of early neurons in the ectoderm was less than 10% in week 4 teratomas, gradually increasing to around 40% in week 10 teratomas (**Figure 4.7C**). We observed similar trends in the endoderm, where precursor definitive endoderm cells in early-stage teratomas gave rise to mid/hindgut epithelium in the later stage teratomas (**Figure 4.7D**); and also in the mesoderm, where COL15A positive MSC/Fibroblast progenitors [96] (termed here as MyoFib cells) exhibit a high initial relative mesodermal percentage that progressively decreases during teratoma development

Figure 4.6: Single Nuclei RNA Editing Analysis Across Teratoma Development.

(A) Workflow schematic for running single nuclei RNA editing analysis across teratoma development. (B) Aggregated and down-sampled UMAP plots from H1 teratomas across weeks 4-10 of development, processed through the single cell RNA sequencing pipeline. 2-5 teratoma samples were processed for each time point. UMAP plots are down-sampled so the total number of cells plotted is consistent across all time points. (C) AEI values for all teratoma cell-types across development. (D) Pan germ-layer correlation between teratoma cell-types and corresponding cell-types from the Human Fetal Cell Atlas database. (E) Transcriptomic correlation between time series teratomas and time resolved human fetal cortex data between Gestational Weeks (GW) 14-19. (F) RNA editing rate, via AEI, correlation between time series teratomas and time resolved human fetal cortex data between GW14-19.

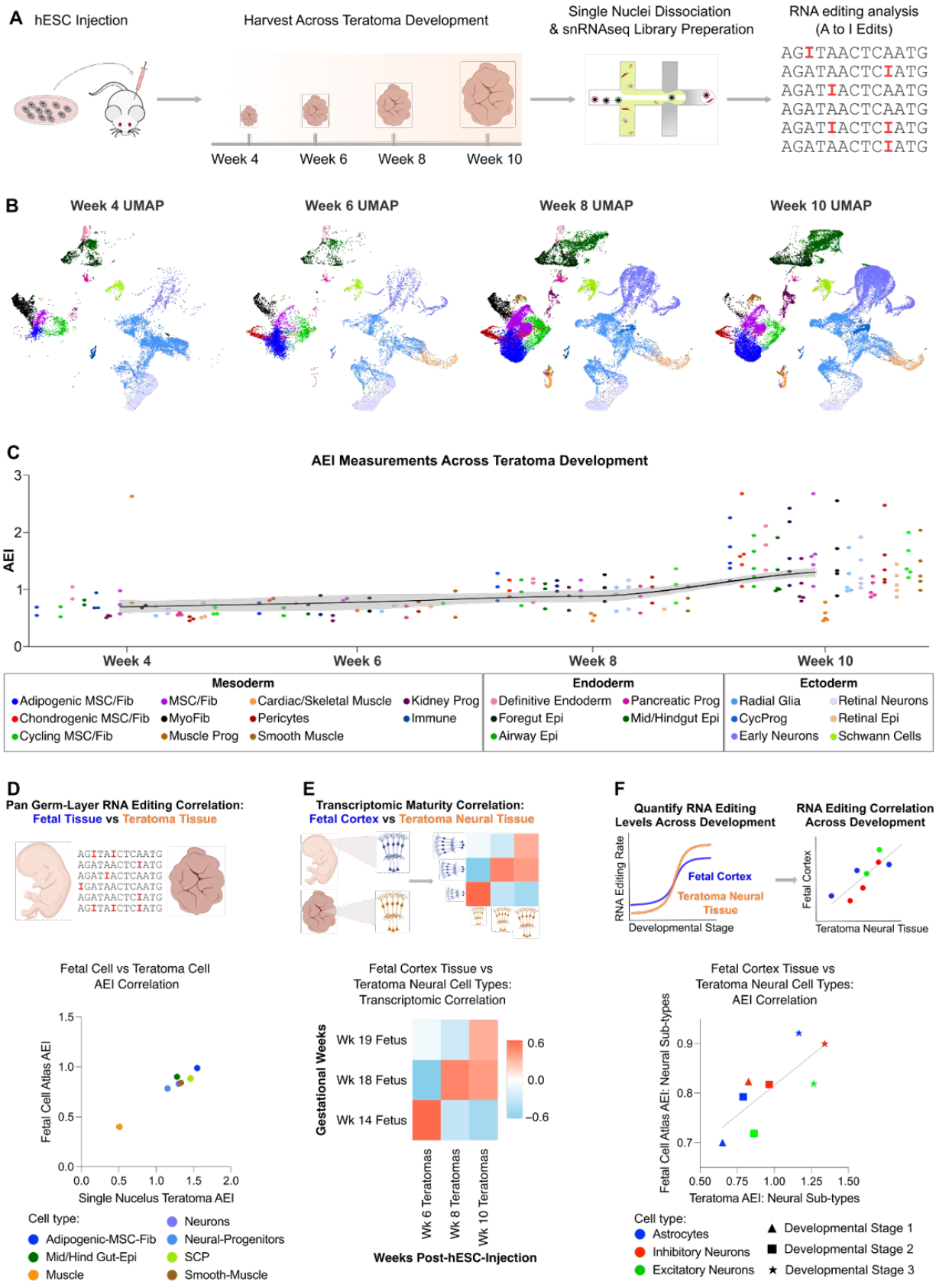
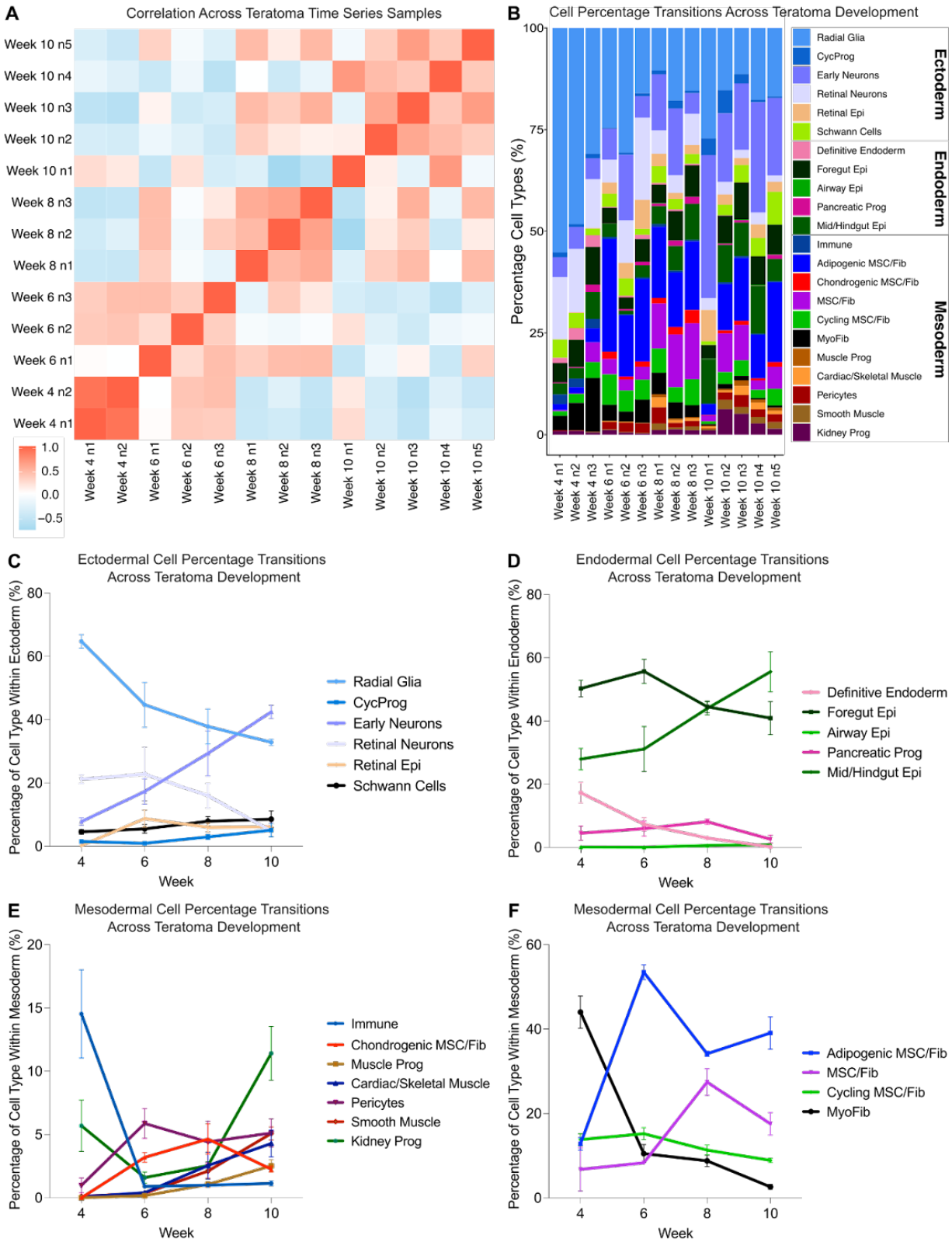

A
hESC Injection
Harvest Across Teratoma Development
Single Nuclei Dissociation & snRNAseq Library Preperation
RNA editing analysis (A to I Edits)
AGITAACTCAATG
AGATAACTCIATG
AGATIACTCAATG
AGATAACTCAATG
AGATIACTCIATG
AGATAACTCIATG
Week 4
Week 6
Week 8
Week 10
B
Week 4 UMAP
Week 6 UMAP
Week 8 UMAP
Week 10 UMAP
C
AEI Measurements Across Teratoma Development
AEI
Week 4
Week 6
Week 8
Week 10
Mesoderm
Adipogenic MSC/Fib
MSC/Fib
Cardiac/Skeletal Muscle
Kidney Prog
Chondrogenic MSC/Fib
MyoFib
Pericytes
Immune
Cycling MSC/Fib
Muscle Prog
Smooth Muscle
Endoderm
Definitive Endoderm
Pancreatic Prog
Foregut Epi
Mid/Hindgut Epi
Airway Epi
Ectoderm
Radial Glia
Retinal Neurons
CycProg
Retinal Epi
Early Neurons
Schwann Cells
D
Pan Germ-Layer RNA Editing Correlation:
Fetal Tissue vs Teratoma Tissue
Fetal Cell vs Teratoma Cell AEI Correlation
Fetal Cell Atlas AEI
Single Nucelus Teratoma AEI
Cell type:
Adipogenic-MSC-Fib
Mid/Hind Gut-Epi
Muscle
Neurons
Neural-Progenitors
SCP
Smooth-Muscle
E
Transcriptomic Maturity Correlation:
Fetal Cortex vs Teratoma Neural Tissue
Fetal Cortex Tissue vs Teratoma Neural Cell Types: Transcriptomic Correlation
Gestational Weeks
Wk 19 Fetus
Wk 18 Fetus
Wk 14 Fetus
Wk 6 Teratomas
Wk 8 Teratomas
Wk 10 Teratomas
Weeks Post-hESC-Injection
F
Quantify RNA Editing Levels Across Development
RNA Editing Rate
Fetal Cortex
Teratoma Neural Tissue
Developmental Stage
RNA Editing Correlation Across Development
Fetal Cortex
Teratoma Neural Tissue
Fetal Cortex Tissue vs Teratoma Neural Cell Types: AEI Correlation
Fetal Cell Atlas AEI: Neural Sub-types
Teratoma AEI: Neural Sub-types
Cell type:
Astrocytes
Inhibitory Neurons
Excitatory Neurons
Developmental Stage 1
Developmental Stage 2
Developmental Stage 3

Figure 4.7: Percent cell type variation across teratoma development.

(A) Transcriptomic correlation across teratoma time series samples, depicting higher levels in comparisons correlation from teratomas samples of the same developmental time point. (B) Percentage of all teratoma cell-types across development, for each teratoma sample. (C) Ectodermal cell-types proportions across teratoma development. (D) Endodermal cell-types proportions across teratoma development. (E) Sub-set of mesodermal cell-types proportions across teratoma development. (F) Sub-set of remaining mesodermal cell-types proportions across teratoma development. Cell-type percentages are normalized within each germ-layer.


A
Correlation Across Teratoma Time Series Samples
Week 10 n5
Week 10 n4
Week 10 n3
Week 10 n2
Week 10 n1
Week 8 n3
Week 8 n2
Week 8 n1
Week 6 n3
Week 6 n2
Week 6 n1
Week 4 n2
Week 4 n1
1.0
0.5
0.0
−0.5
B
Cell Percentage Transitions Across Teratoma Development
Percentage Cell Types (%)
Radial Glia
CycProg
Early Neurons
Retinal Neurons
Retinal Epi
Schwann Cells
Ectoderm
Definitive Endoderm
Foregut Epi
Airway Epi
Pancreatic Prog
Mid/Hindgut Epi
Endoderm
Immune
Adipogenic MSC/Fib
Chondrogenic MSC/Fib
MSC/Fib
Cycling MSC/Fib
MyoFib
Muscle Prog
Cardiac/Skeletal Muscle
Pericytes
Smooth Muscle
Kidney Prog
Mesoderm
C
Ectodermal Cell Percentage Transitions
Across Teratoma Development
Percentage of Cell Type Within Ectoderm (%)
Week
D
Endodermal Cell Percentage Transitions
Across Teratoma Development
Percentage of Cell Type Within Endoderm (%)
Week
E
Mesodermal Cell Percentage Transitions
Across Teratoma Development
Percentage of Cell Type Within Mesoderm (%)
Week
F
Mesodermal Cell Percentage Transitions
Across Teratoma Development
Percentage of Cell Type Within Mesoderm (%)
Week

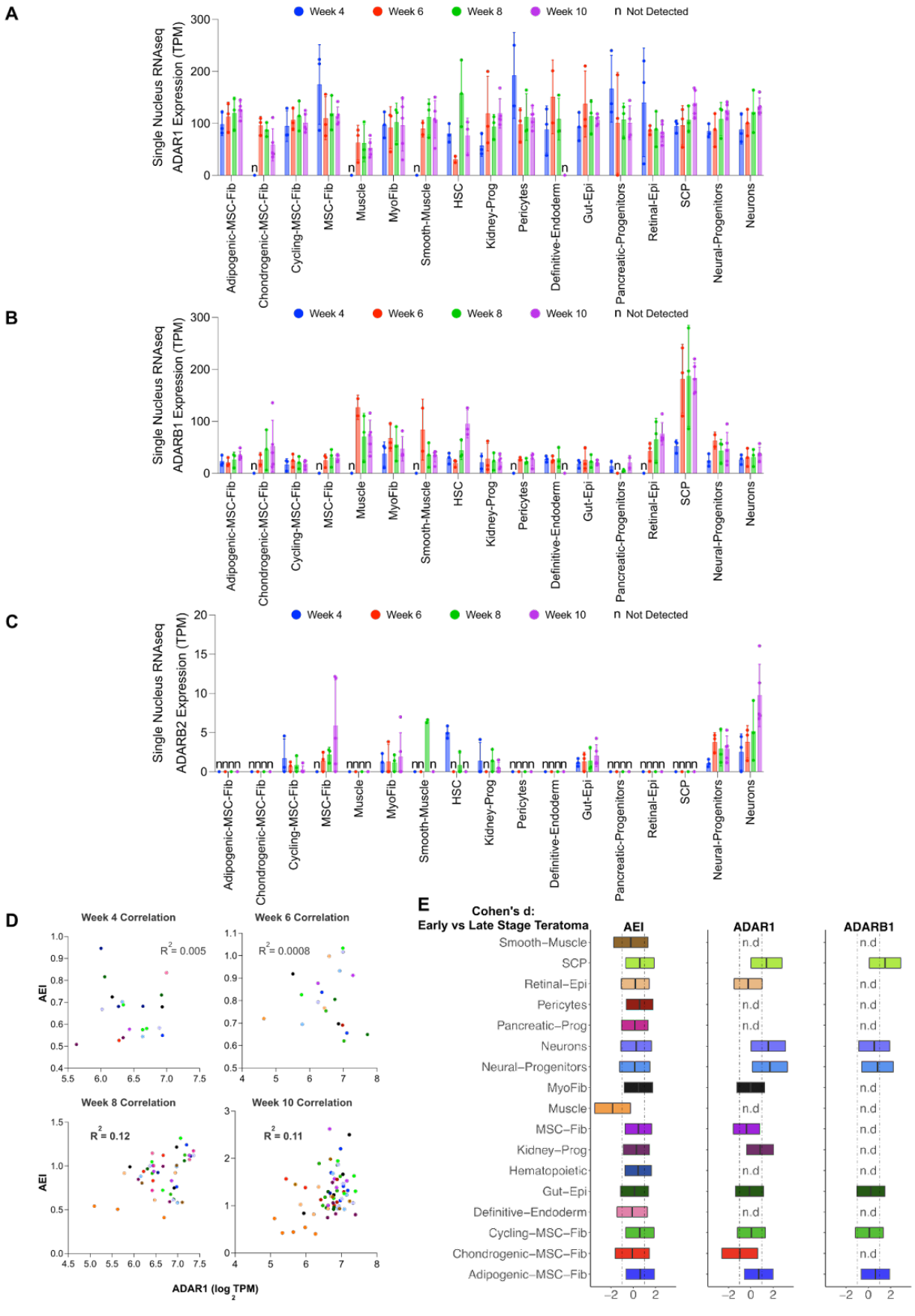
(**Figure 4.7F**). These results demonstrate that teratoma cell types reproducibly mature into their differentiated progeny in a progressive manner.

Similar to week 10 teratomas (**Figure 4.4D**), we observed ubiquitous expression of *ADAR1* across all stages of teratoma development, while *ADARB2* exhibited neural-specific expression, and *ADARB1* showed strongest expression in neural tissues but also was expressed at low levels elsewhere (**Figure 4.8A-C**). Interestingly, our analysis revealed that the contribution of *ADAR1* to AEI variation in week 4 and 6 teratomas was less than 1% but increased to approximately 10% by weeks 8 and 10 (**Figure 4.8D**). Furthermore, we observed a 2-fold increase in average AEI throughout teratoma development (pooled Cohen's d = 1.21) (**Figure 4.6C**), which was largely driven by, increases in *ADAR1* (pooled Cohen's d = 0.25) (**Figure 4.8E**). Notably, within Schwann Cell Progenitors (SCPs), *ADARB1* (p-value = 4.13E-02, Cohen's D = 1.73), in comparison to *ADAR1* (p-value = 9.26E-02, Cohen's D = 1.21), seemed to mostly contribute to the significant rise in SCP AEI (p-value = 1.56E-03, Cohen's D = 2.01) throughout teratoma development (**Figure 3.8E**). These results motivated a deeper investigation into the reliability teratoma RNA editing dynamics.

To examine if the teratoma could serve as a reliable model for spatiotemporal RNA editing in human development, we correlated AEI values from teratomas with those from corresponding fetal cell types and found a strong correlation across cell-types from all germ-layers (**Figure 4.6D**). Given the dynamic RNA editing profiles evident in fetal brain tissue (**Figure 4.2B, 4.2C**) [80], our attention next shifted to teratoma neural tissue for a more in-depth analysis of time-resolved RNA editing in the teratoma.

Figure 4.8: ADAR enzyme family expression in teratoma cell types across time.

(A) *ADAR1* expression values for all teratoma cell-types, across teratoma developmental time-points. (B) *ADARB1* expression values for all teratoma cell-types, across teratoma developmental time-points. (C) *ADARB2* expression values for all teratoma cell-types, across teratoma developmental time-points. (D) AEI to *ADAR1* expression correlation for all teratoma cell-types, across teratoma developmental time-points. (E) Cohen's d comparison between Early Teratoma (weeks 4+6) samples versus Late Teratoma (weeks 8+10) samples for AEI, *ADAR1*, *ADARB1*, and *ADARB2*, across all teratoma cell-types (not detected = n.d).


A
Week 4
Week 6
Week 8
Week 10
n Not Detected
Single Nucleus RNAseq
ADAR1 Expression (TPM)
300
200
100
0
Adipogenic-MSC-Fib
Chondrogenic-MSC-Fib
Cycling-MSC-Fib
MSC-Fib
Muscle
MyoFib
Smooth-Muscle
HSC
Kidney-Prog
Pericytes
Definitive-Endoderm
Gut-Epi
Pancreatic-Progenitors
Retinal-Epi
SCP
Neural-Progenitors
Neurons
B
Single Nucleus RNAseq
ADARB1 Expression (TPM)
C
Single Nucleus RNAseq
ADARB2 Expression (TPM)
20
15
10
5
0
D
Week 4 Correlation
R2 = 0.005
Week 6 Correlation
R2 = 0.0008
Week 8 Correlation
R2 = 0.12
Week 10 Correlation
R2 = 0.11
AEI
ADAR1 (log2 TPM)
E
Cohen's d:
Early vs Late Stage Teratoma
AEI
ADAR1
ADARB1
Smooth−Muscle
SCP
Retinal−Epi
Pericytes
Pancreatic−Prog
Neurons
Neural−Progenitors
MyoFib
Muscle
MSC−Fib
Kidney−Prog
Hematopoietic
Gut−Epi
Definitive−Endoderm
Cycling−MSC−Fib
Chondrogenic−MSC−Fib
Adipogenic−MSC−Fib
n.d
−2
0
2

Transcriptomic correlations across fetal cortex and teratoma neural cell-types revealed that neural cell-types within the teratoma progressively mature as the teratoma develops, as measured via the progressive increase in cosine similarity index with successive developmental time-points (**Figure 4.6E**). This progressive transcriptomic maturation prompted us to compare RNA editing in teratoma neural tissue with that in fetal cerebrum tissue over time (**Figure 4.6F**). Consistent with our maturity findings, the AEI measurements between teratoma and fetal samples showed a strong epitranscriptomic alignment throughout development (**Figure 4.6F**).

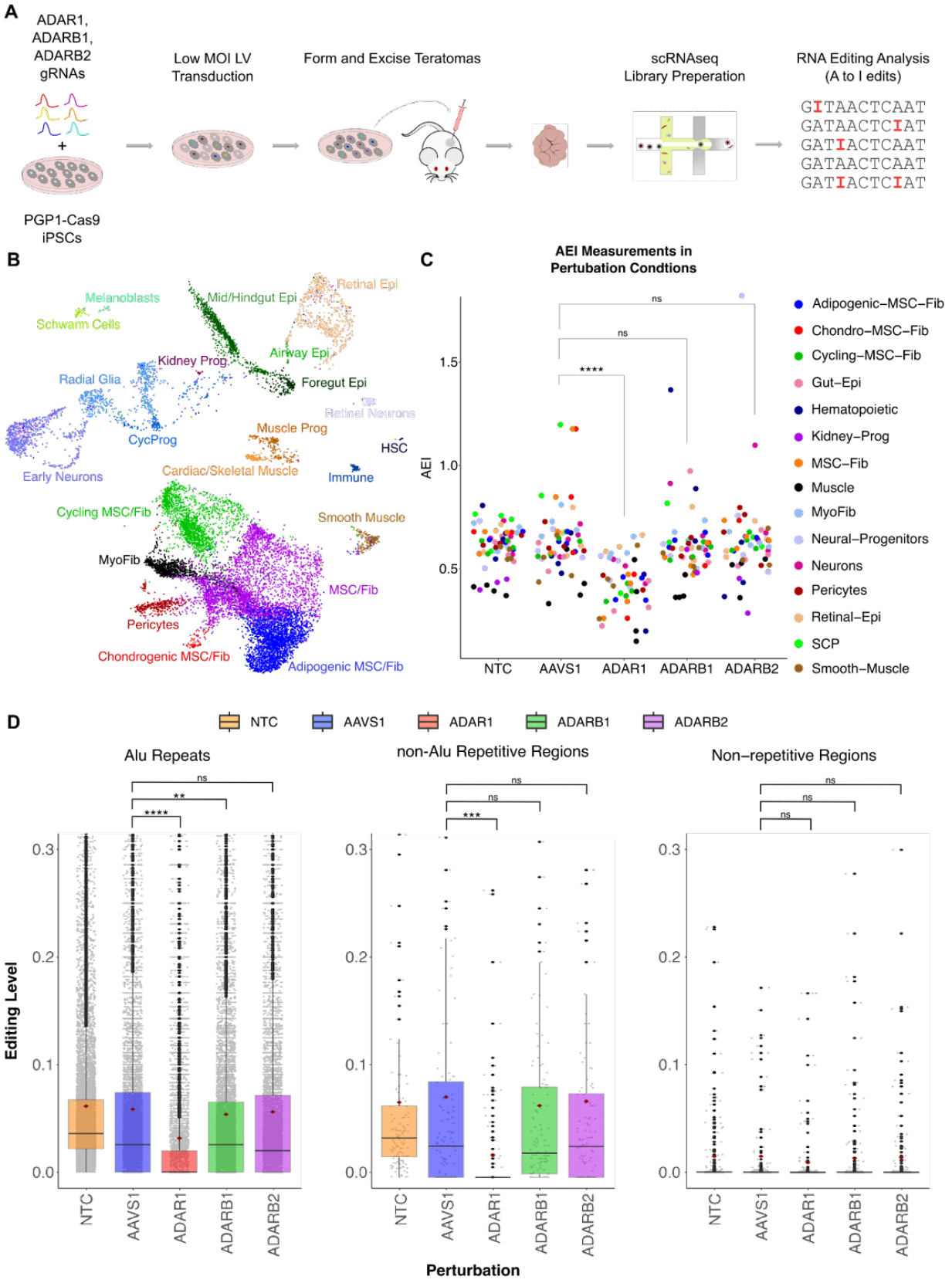
In summary, our correlations of transcriptomic maturity and global editing with human fetal tissue corroborate the teratoma as an effective tool to model temporal RNA editing patterns during early fetal development. This motivated the use of the teratomas as a platform to probe and interrogate the functions of ADARs in human development via pooled CRISPR-KO screens.

RNA Editing Analysis in ADAR-KO teratomas

To directly assess the functions of ADAR family enzymes during cell-fate specification, we adopted a previously generated iPSC line, in which the Cas9 protein is knocked-in to the *AAVS1* locus of the PGP1 iPSCs (PGP1-Cas9s) [90]. PGP1-Cas9s were transduced with a gRNA library targeting *ADAR1*, *ADARB1*, and *ADARB2*, injected into immunodeficient mice, and formed teratomas (ADAR-KO teratomas) (**Figure 4.9**). After a 10-week growth period, ADAR-KO teratomas were excised, and single-cell RNA sequencing libraries were generated. We confirmed that previously identified teratoma cell-types were present in ADAR-KO teratomas (**Figure 4.9B**), and investigated RNA editing signatures across all cell-types utilizing both the AEI and a supervised site-specific analysis of A-to-I editing sites cataloged in existing RNA editing database [97].

Figure 4.9: RNA Editing Analysis in ADAR-KO teratomas.

(A) Workflow schematic for ADAR-KO screen in PGP1-Cas9 iPSCs, followed by teratoma formation and downstream single cell RNA editing analysis. (B) Aggregated UMAP plot from 4 PGP1-Cas9 + ADAR-KO teratomas. (C) AEI values for all teratoma cell-types in ADAR-KO teratomas. (D) Pan-teratoma site-specific RNA editing analysis across KO conditions.


A
ADAR1, ADARB1, ADARB2 gRNAs
+
PGP1-Cas9 iPSCs
Low MOI LV Transduction
Form and Excise Teratomas
scRNAseq Library Preperation
RNA Editing Analysis (A to I edits)
GITAACTCAAT
GATAACTCIAT
GATIACTCAAT
GATAACTCAAT
GATIACTCIAT
B
Melanoblasts
Schwann Cells
Mid/Hindgut Epi
Retinal Epi
Airway Epi
Kidney Prog
Radial Glia
Foregut Epi
Retinal Neurons
Muscle Prog
CycProg
HSC
Early Neurons
Cardiac/Skeletal Muscle
Immune
Cycling MSC/Fib
Smooth Muscle
MyoFib
MSC/Fib
Pericytes
Chondrogenic MSC/Fib
Adipogenic MSC/Fib
C
AEI Measurements in Pertubation Condtions
AEI
NTC
AAVS1
ADAR1
ADARB1
ADARB2
Adipogenic–MSC–Fib
Chondro–MSC–Fib
Cycling–MSC–Fib
Gut–Epi
Hematopoietic
Kidney–Prog
MSC–Fib
Muscle
MyoFib
Neural–Progenitors
Neurons
Pericytes
Retinal–Epi
SCP
Smooth–Muscle
D
Alu Repeats
non-Alu Repetitive Regions
Non–repetitive Regions
Editing Level
Perturbation

In comparison to *AAVS1* double-strand break induction control cells (*AAVS1*-KO), we observed a statistically significant drop in AEI upon *ADAR1* knockout (p-value = 9.6E-9, Cohen's d = -1.39) (**Figure 4.9C**). In contrast, the disruption of *ADARB1* and *ADARB2* did not lead to discernible effects on global editing patterns. Furthermore, we examined site-specific RNA editing levels across bonafide editing sites cataloged in the RediPortal database[97] (**Figure 4.9D**). Consistent with the above AEI analysis results, editing sites within *Alu* repeats (Cohen's d = -0.29) and non-*Alu* repetitive regions (Cohen's d = -0.45) exhibited a statistically significant drop in editing levels upon *ADAR1* knockout (**Figure 4.9D**). We also noticed a significant decrease in editing of *Alu* repeats following *ADARB1* knockout, although the impact is relatively minor compared to the effect of *ADAR1* knockout (Cohen's d = -0.05) (**Figure 4.9D**). In contrast, *ADARB2* knockout did not result in any significant changes in Alu editing (**Figure 4.9D**). These observations highlight the primary role of *ADAR1* in A-to-I editing within *Alu* elements.

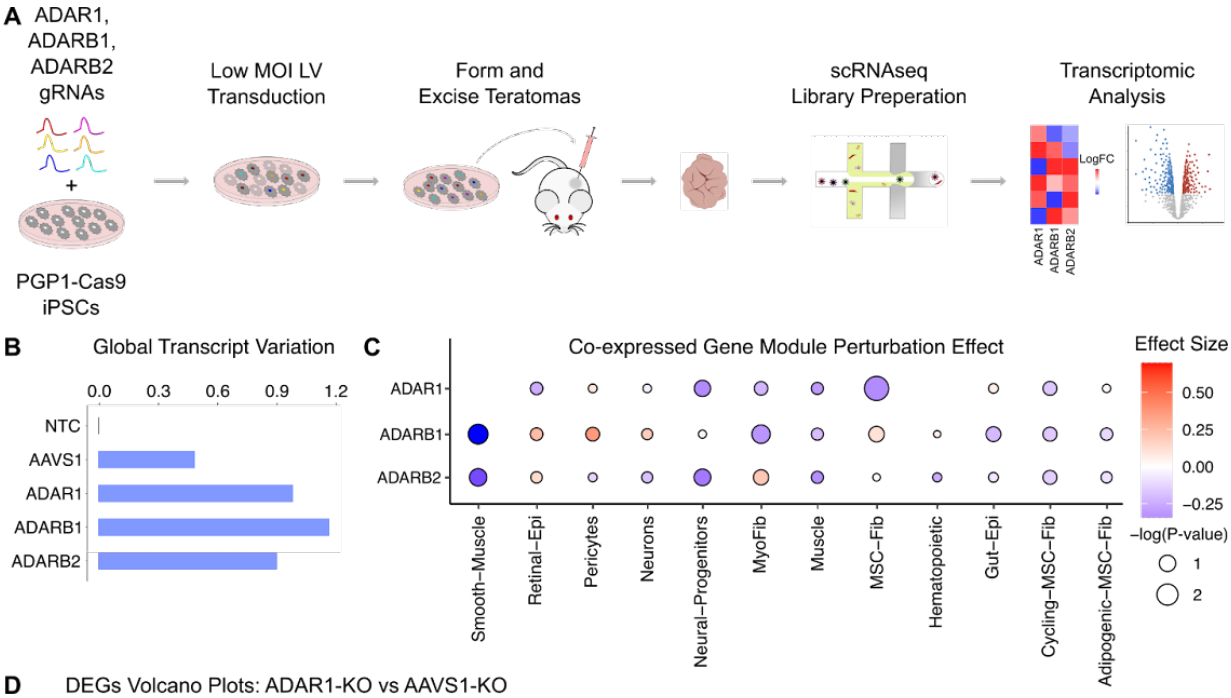
Transcriptomic and Fitness Analysis in ADAR-KO teratomas

Motivated by the observed decline in RNA editing upon *ADAR1*-KO, we extended our investigation to the transcriptomic level within ADAR-KO teratoma cell types (**Figure 4.10A**). Employing the Earth's Mover Distance (EMD), a measure of dissimilarity between KO and non-targeting control (NTC) samples (refer Methods), we observed heightened global transcriptomic variation in *ADAR1*-KO, *ADARB1*-KO, and *ADARB2*-KO samples, compared to *AAVS1*-KO samples (**Figure 4.10B**).

To further investigate this transcriptomic variation upon *ADAR1*-, *ADARB1*-, and *ADARB2*-KO, we conducted weighted correlation network analysis (refer Methods), revealing a reduced expression (negative effect size, **Figure 4.10C**) of co-expressed genes across various cell types upon *ADAR1* knockout, with the most pronounced impact observed in mesodermal cell-

Figure 4.10: Transcriptomic Analysis in ADAR-KO Teratomas Reveals Germ-Layer Specific Fitness Defects.

(A) Workflow schematic for ADAR-KO screen in PGP1-Cas9 iPSCs, followed by teratoma formation and downstream transcriptomic analysis. (B) Pan-teratoma global transcriptomic variation across perturbation conditions. (C) Cell-type specific gene-module effect across perturbation conditions. (D) Differential expressed genes (DEGs) for *ADAR1*-KO versus *AAVS1*-KO across the three germ layers. Teratoma cell-types from the same germ-layer are transcriptomicaly grouped together. (E) Multi-lineage fitness analysis assessing the fitness defect across perturbation conditions. Cell counts from the same germ-layer are summed together.


A
ADAR1, ADARB1, ADARB2 gRNAs
+
PGP1-Cas9 iPSCs
Low MOI LV Transduction
Form and Excise Teratomas
scRNAseq Library Preperation
Transcriptomic Analysis
LogFC
ADAR1
ADARB1
ADARB2
B
Global Transcript Variation
0.0
0.3
0.6
0.9
1.2
NTC
AAVS1
ADAR1
ADARB1
ADARB2
C
Co-expressed Gene Module Perturbation Effect
Effect Size
0.50
0.25
0.00
−0.25
−log(P-value)
1
2
Smooth−Muscle
Retinal−Epi
Pericytes
Neurons
Neural−Progenitors
MyoFib
Muscle
MSC−Fib
Hematopoietic
Gut−Epi
Cycling−MSC−Fib
Adipogenic−MSC−Fib
D
DEGs Volcano Plots: ADAR1-KO vs AAVS1-KO
Ectoderm
Mesoderm
Endoderm
−log(P-value)
LogFC
GADD45B
CCN2
GADD45G
TRAF4
GADD45A
BECN1
TRAF3
TNFRSF19
E
Multi-lineage Fitness Analysis
Ectoderm
Mesoderm
Endoderm
p = 0.5777
p = 0.2066
p = 0.0091
p = 0.0294
p = 0.1948
p = 0.0953
p = 0.0025
p = 0.1176
p = 0.9957
p = 0.9511
p = 0.1826
p = 0.4568
Log2(Post-Harvest Frequency / Pre-Injection Frequency)
NTC
AAVS1
ADAR1
ADARB1
ADARB2
Perturbation

types, such as in the MyoFib, MSC/Fib, and muscle, as well as in ectodermal Neural-Progenitors (**Figure 4.10C**). Next, we conducted differential expression analysis across knockout samples to examine these global transcriptomic impacts upon knockout.

First, DEGs were discovered at the germ-layer level across perturbations, and calculated relative to *AAVS1*-KOs (refer Methods). Discovered DEGs were used to conduct Gene Ontology (GO) term analysis in order to identify upregulated and downregulated gene families (refer Methods; **Figures 4.11A-E**). GO term analysis revealed the upregulation of the JNK pathway activation gene family in the Mesoderm *ADAR1*-KO cells (**Figure 4.11B**). Interestingly, we observed that the JNK pathway genes were significantly upregulated solely in the mesoderm, and not in other germ layers (**Figure 4.10D**). Considering the JNK pathway is a stress and apoptosis induction pathway[98], we next conducted a fitness analysis across perturbations. Initially, we observed a pan-teratoma fitness defect in *ADAR1*-KO cells, while no significant fitness impact was observed in *ADARB1*-KO or *ADARB2*-KO cells (**Figure 4.12B**). Germ-layer level analysis unveiled a significant fitness defect in both the mesoderm *ADAR1*-KO and ectoderm *ADAR1*-KO (**Figure 4.10E**), compared to NTC. However, the effect size of this fitness defect of *ADAR1*-KO, compared to NTC, is most pronounced in the mesoderm (Cohen's d = 3.12) compared to the ectoderm (Cohen's d = 2.16). These results mirror the significant fitness defect - which ultimately leads to embryonic lethality - seen in *Adar1*-KO mice[86], further corroborating the teratoma as a faithful model which recapitulates phenotypes seen in bonafide mammalian development.

Cell-Type Enrichment in ADAR-KO Teratomas Suggests ADAR1 as an Inhibitor of Human Adipogenesis

Figure 4.11: Gene Ontology Discovery Pipeline for ADAR-KO Teratomas.

(A) Inflammatory and immune gene panel expression for germ-layer ADAR1-KO samples, normalized to AAVS1-KO expression. Values represented as Log2(Fold Change). (A) Computational pipeline for DEG discovery. (B) Top GO terms identified from Mesoderm *ADAR1*-KO versus *AAVS1*-KO DEGs discovery pipeline. (C) Top GO terms identified from Mesoderm *ADARB1*-KO versus *AAVS1*-KO DEGs discovery pipeline. (D) Top GO terms identified from Ectoderm *ADAR1*-KO versus *AAVS1*-KO DEGs discovery pipeline. (E) Top GO terms identified from Ectoderm *ADARB1*-KO versus *AAVS1*-KO DEGs discovery pipeline.

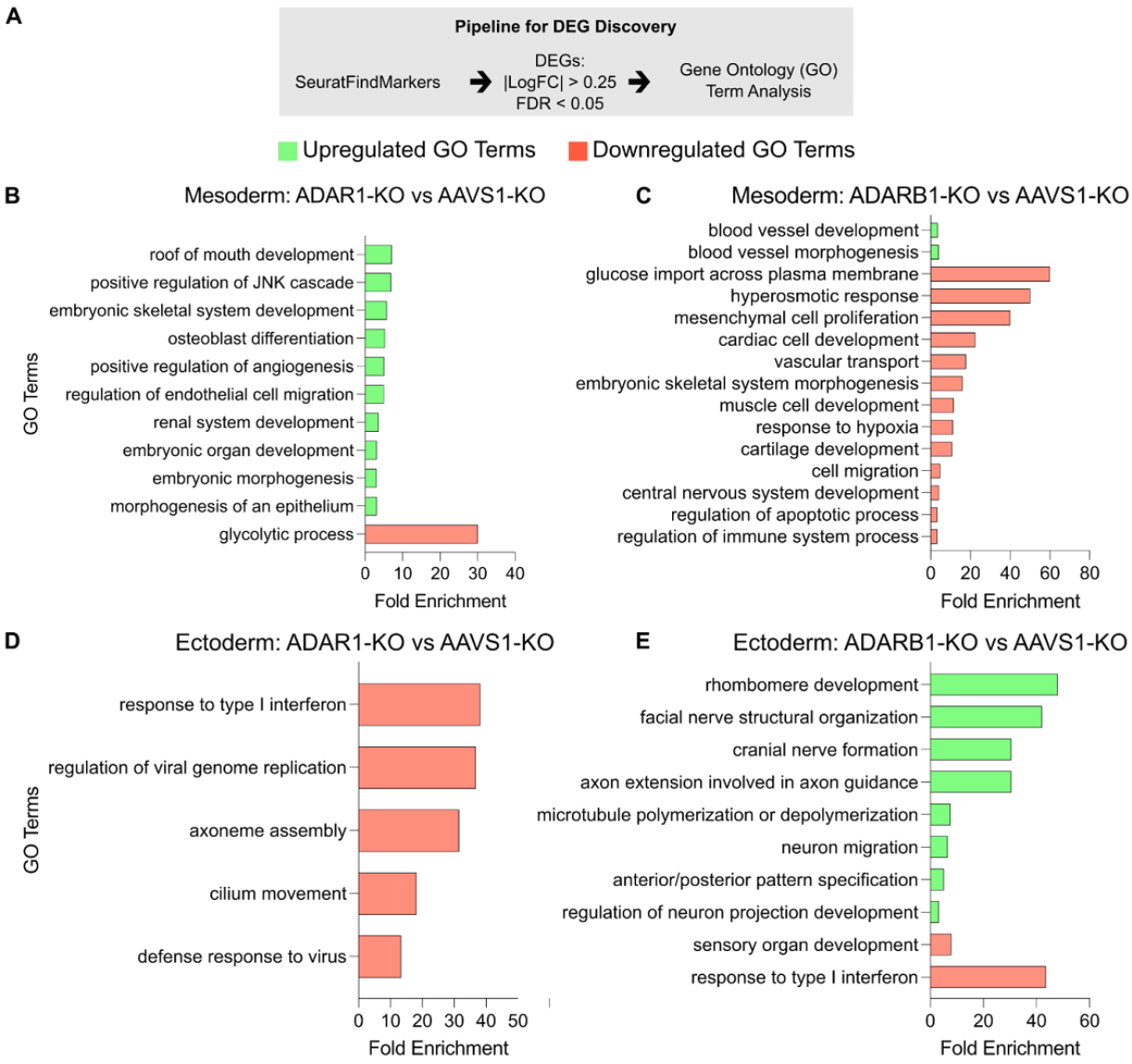
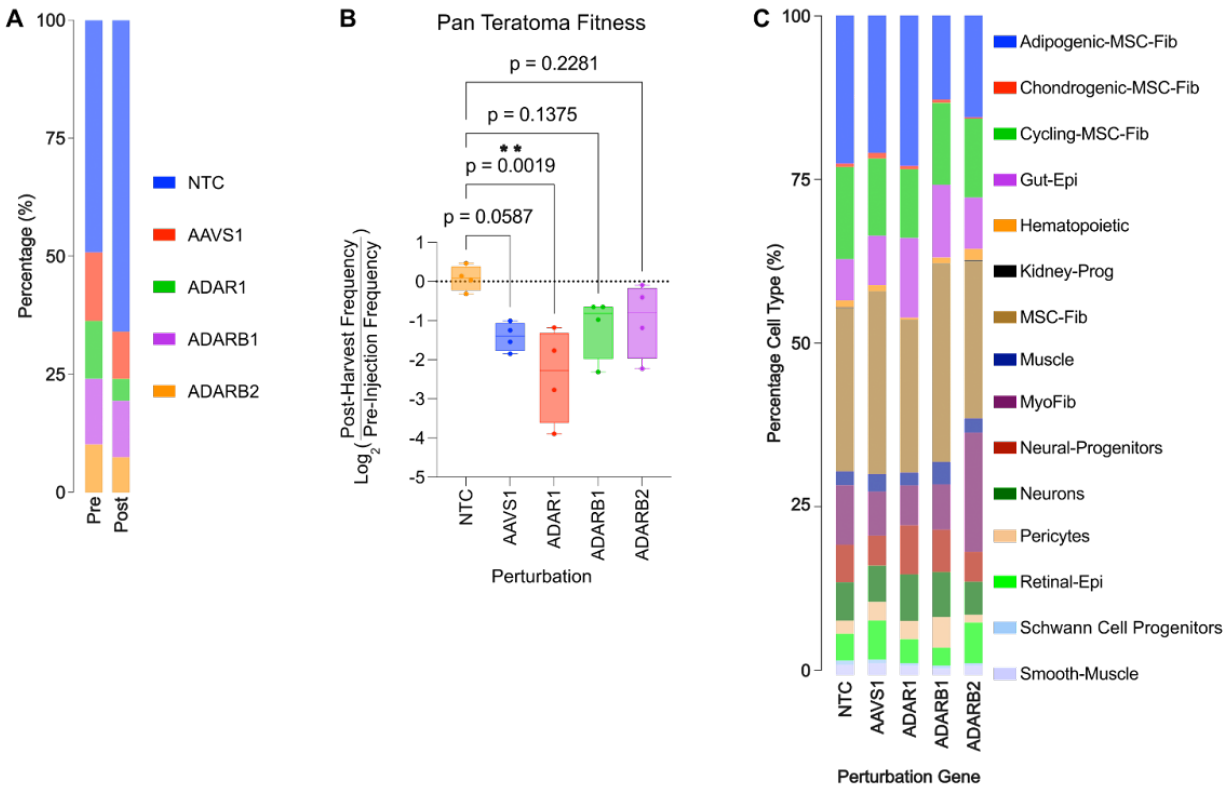

A
Pipeline for DEG Discovery
SeuratFindMarkers
DEGs:
|LogFC| > 0.25
FDR < 0.05
Gene Ontology (GO)
Term Analysis
Upregulated GO Terms
Downregulated GO Terms
B
Mesoderm: ADAR1-KO vs AAVS1-KO
GO Terms
roof of mouth development
positive regulation of JNK cascade
embryonic skeletal system development
osteoblast differentiation
positive regulation of angiogenesis
regulation of endothelial cell migration
renal system development
embryonic organ development
embryonic morphogenesis
morphogenesis of an epithelium
glycolytic process
0 10 20 30 40
Fold Enrichment
C
Mesoderm: ADARB1-KO vs AAVS1-KO
blood vessel development
blood vessel morphogenesis
glucose import across plasma membrane
hyperosmotic response
mesenchymal cell proliferation
cardiac cell development
vascular transport
embryonic skeletal system morphogenesis
muscle cell development
response to hypoxia
cartilage development
cell migration
central nervous system development
regulation of apoptotic process
regulation of immune system process
0 20 40 60 80
Fold Enrichment
D
Ectoderm: ADAR1-KO vs AAVS1-KO
GO Terms
response to type I interferon
regulation of viral genome replication
axoneme assembly
cilium movement
defense response to virus
0 10 20 30 40 50
Fold Enrichment
E
Ectoderm: ADARB1-KO vs AAVS1-KO
rhombomere development
facial nerve structural organization
cranial nerve formation
axon extension involved in axon guidance
microtubule polymerization or depolymerization
neuron migration
anterior/posterior pattern specification
regulation of neuron projection development
sensory organ development
response to type I interferon
0 20 40 60
Fold Enrichment

Figure 4.12: Pan-Teratoma View of ADAR-KO Impact on Cell Fitness.

(A) Pan-teratoma percentage for each perturbation in the pre-injection population and post-teratoma harvest. (B) Pan-teratoma fitness analysis across perturbation conditions in ADAR-KO teratomas. (C) Percentage of all teratoma cell-types across all perturbations.


A
100
75
50
25
0
Percentage (%)
Pre
Post
NTC
AAVS1
ADAR1
ADARB1
ADARB2
B
Pan Teratoma Fitness
p = 0.2281
p = 0.1375
**
p = 0.0019
p = 0.0587
1
0
-1
-2
-3
-4
-5
Log2(Post-Harvest Frequency / Pre-Injection Frequency)
NTC
AAVS1
ADAR1
ADARB1
ADARB2
Perturbation
C
100
75
50
25
0
Percentage Cell Type (%)
NTC
AAVS1
ADAR1
ADARB1
ADARB2
Perturbation Gene
Adipogenic-MSC-Fib
Chondrogenic-MSC-Fib
Cycling-MSC-Fib
Gut-Epi
Hematopoietic
Kidney-Prog
MSC-Fib
Muscle
MyoFib
Neural-Progenitors
Neurons
Pericytes
Retinal-Epi
Schwann Cell Progenitors
Smooth-Muscle

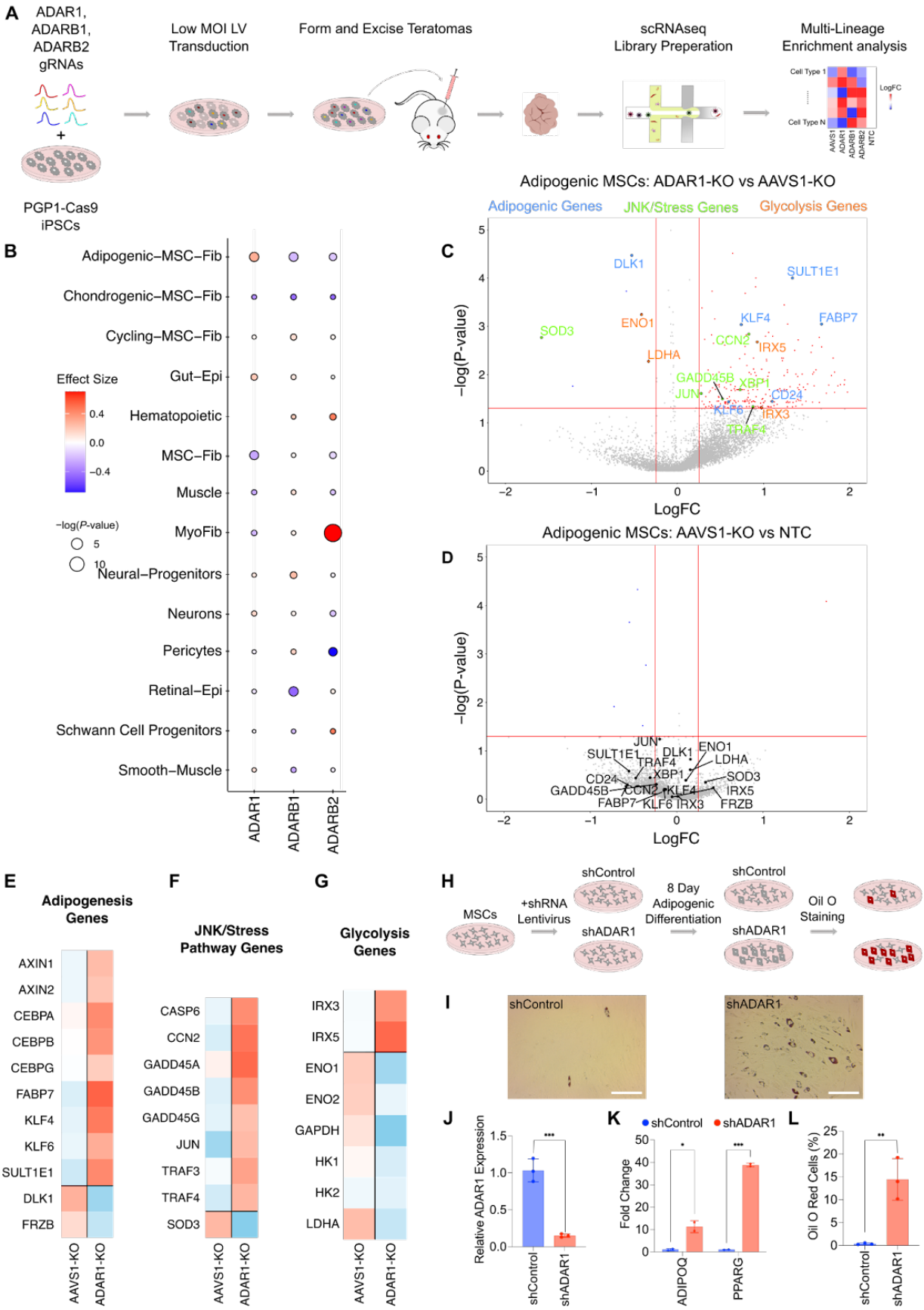
Motivated by the teratoma's unique multi-lineage nature and ability to faithfully recapitulate developmental phenotypes, we investigated the ADAR proteins' significance in human embryonic cell-fate specification, which remains insufficiently investigated and poorly understood. Leveraging ADAR-KO teratomas as an experimental model, our objective was to uncover the extent of the ADAR family of proteins' influence on cell-fate determination across all germ-layers (**Figure 4.13A**).

Our investigation commenced with an enrichment analysis of ADAR-KO teratomas, unveiling a pan-teratoma reduction in *ADAR1*-expressing cells (**Figure 4.12A**) and discernible variations in cell type compositions across perturbations (**Figure 4.12C**). Employing a prior developed pipeline to calculate cell-type compositions following a pooled *in vivo* CRISPR-Cas9 knockout screen (refer Methods)[99], we sought to further analyze these observed effects via cell-type enrichment analysis. Enrichment analysis unveiled a significant increase of Adipogenic-MSCs in *ADAR1*-KO cells compared to *AAVS1*-KO controls (**Figure 4.13B**). This finding was unexpected, as it is contrary to the previously observed fitness defect in mesoderm *ADAR1*-KO cells (**Figure 4.10E**). DEG analysis in *ADAR1*-KO Adipogenic-MSCs vs AAVS1-KO Adipogenic-MSCs revealed many differentially expressed adipogenic, JNK/Stress pathway, and glycolysis genes (**Figure 4.13C**). In contrast, we found no significant DEGs when comparing AAVS1-KO Adipogenic-MSCs to NTC Adipogenic-MSC controls (**Figure 4.13D**). Further analysis into this JNK-mediated stress response in *ADAR1*-KO Adipogenic-MSCs revealed a strong stress response gene expression signature (**Figure 4.13E**). In line with enrichment results (**Figure 4.13B**), we correspondingly observed a strong pro-adipogenic gene expression signature (**Figure 4.13F**).

Inspired by recent literature that links viral RNA infection, and the subsequent stress response, to downregulated glycolysis via MAVS signaling[100,101], we measured the expression of

Figure 4.13: Enrichment Analysis in ADAR-KO Teratomas Reveals ADAR1 as a Potential Inhibitor of Human Adipogenesis.

(A) Workflow schematic for ADAR-KO screen in PGP1-Cas9 iPSCs, followed by teratoma formation and downstream cell-type enrichment analysis. (B) Enrichment analysis depicting cell-type specific enrichment or depletion across perturbation conditions. (C) DEGs from comparing Adipogenic MSC *ADAR1*-KO cells versus Adipogenic MSC AAVS1-KO cells. Genes related to adipogenesis (blue), JNK/stress pathways (green), and glycolysis (red) are highlighted. (D) DEG analysis comparing Adipogenic MSC *AAVS1*-KO cells versus Adipogenic MSC NTC cells. Genes related to adipogenesis, JNK/stress pathways, and glycolysis are denoted. (E) Gene expression analysis across a subset of adipogenic pathway genes in Adipogenic MSC *ADAR1*-KO cells versus Adipogenic MSC AAVS1-KO cells. (F) Gene expression analysis across a subset of JNK/stress pathway genes in Adipogenic MSC *ADAR1*-KO cells versus Adipogenic MSC *AAVS1*-KO cells. (G) Gene expression analysis across a subset of glycolysis pathway genes in Adipogenic MSC *ADAR1*-KO cells versus Adipogenic MSC *AAVS1*-KO cells. (H) Schematic depicting experimental validation of *ADAR1*'s role in adipogenic differentiation from MSCs. (I) Oil O Red staining of 8 day differentiated Adipogenic-MSCs (Scale bar = 250 μm). (J) *ADAR1* transcript quantification in MSCs post shRNA transduction, at day 0 of adipogenic differentiation (n=3). (K) *ADIPOQ* and PPARG transcript quantification in MSCs post shRNA transduction, at day 8 of adipogenic differentiation (n=2). (L) Quantification of Oil O red percent positive Adipogenic-MSCs. Full tilescan images of a 48-well are taken and quantified with respect to total DAPI+ cells (n=3).


A
ADAR1, ADARB1, ADARB2 gRNAs
PGP1-Cas9 iPSCs
Low MOI LV Transduction
Form and Excise Teratomas
scRNAseq Library Preperation
Multi-Lineage Enrichment analysis
Cell Type 1
Cell Type N
LogFC
AAVS1
ADAR1
ADARB1
ADARB2
NTC
B
Adipogenic–MSC–Fib
Chondrogenic–MSC–Fib
Cycling–MSC–Fib
Gut–Epi
Hematopoietic
MSC–Fib
Muscle
MyoFib
Neural–Progenitors
Neurons
Pericytes
Retinal–Epi
Schwann Cell Progenitors
Smooth–Muscle
Effect Size
0.4
0.0
−0.4
−log(P-value)
5
10
ADAR1
ADARB1
ADARB2
C
Adipogenic MSCs: ADAR1-KO vs AAVS1-KO
Adipogenic Genes
JNK/Stress Genes
Glycolysis Genes
−log(P-value)
LogFC
DLK1
SULT1E1
SOD3
ENO1
KLF4
FABP7
CCN2
IRX5
LDHA
GADD45B
XBP1
JUN
CD24
KLF6
IRX3
TRAF4
D
Adipogenic MSCs: AAVS1-KO vs NTC
−log(P-value)
LogFC
JUN
SULT1E1
DLK1
ENO1
TRAF4
LDHA
XBP1
CD24
SOD3
GADD45B
CCN2
KLF4
IRX5
FABP7
KLF6
IRX3
FRZB
E
Adipogenesis Genes
AXIN1
AXIN2
CEBPA
CEBPB
CEBPG
FABP7
KLF4
KLF6
SULT1E1
DLK1
FRZB
AAVS1-KO
ADAR1-KO
F
JNK/Stress Pathway Genes
CASP6
CCN2
GADD45A
GADD45B
GADD45G
JUN
TRAF3
TRAF4
SOD3
AAVS1-KO
ADAR1-KO
G
Glycolysis Genes
IRX3
IRX5
ENO1
ENO2
GAPDH
HK1
HK2
LDHA
AAVS1-KO
ADAR1-KO
H
MSCs
+shRNA Lentivirus
shControl
shADAR1
8 Day Adipogenic Differentiation
shControl
shADAR1
Oil O Staining
I
shControl
shADAR1
J
Relative ADAR1 Expression
***
shControl
shADAR1
K
shControl
shADAR1
Fold Change
*
***
ADIPOQ
PPARG
L
Oil O Red Cells (%)
**
shControl
shADAR1

glycolysis-related genes to see if the *ADAR1*-KO perturbation recapitulated this phenotype. We observed a decreased expression in many glycolytic genes (**Figure 4.13G**). Due to a recent report demonstrating that IRX proteins trigger adipogenesis from hMSCs via repressed glycolysis[102], we measured IRX3 and IRX5 expression levels and observed an increase in their expression levels (**Figure 4.13G**).

To validate the observed increase in Adipogenic-MSCs within teratoma cells following *ADAR1*-KO, we conducted experiments using primary human MSCs *in vitro*. These cells were transduced with lentivirus carrying *ADAR1*-targeting shRNA (shADAR1) or a non-targeting shRNA control (shControl) (**Figure 4.13H**). We first verified the effectiveness of *ADAR1* knockdown after shADAR1 transduction at the onset of adipogenic differentiation (day 0) (**Figure 4.13J**) Furthermore, we observed enhanced adipogenesis at day 8 of adipogenic differentiation in shADAR1 samples, as indicated by an upregulation of adipogenic markers *ADIPOQ* and *PPARG* (**Figure 4.13K**) and a significant increase in Oil Red O positive percent cells compared to the shControl (**Figure 4.13I, L**). These findings validate the observed adipogenic enrichment in *ADAR1*-KO teratomas.

We hypothesize that an increase in dsRNA following *ADAR1*-KO[103] triggers a similar MAVS-mediated stress response following viral infection. Furthermore, we suggest that the stress induced by JNK pathway activation triggers adipogenesis, and synergistically interacts with the IRX-regulated reduction in glycolysis - a known driver of adipogenesis from hMSCs[102]. This intricate interplay potentially suggests putative mechanistic underpinnings behind the observed enrichment of Adipogenic-MSCs following *ADAR1* knockdown.

Discussion

Beyond their known role in editing dsRNA to prevent unwanted innate immune responses, the diverse functions of ADAR proteins, especially in embryonic development, are understudied. Here, we characterized spatiotemporal RNA editing patterns across various human fetal tissues, extended our investigations to lab-grown developmental model systems, and probed the functional roles of ADARs in cell-fate determination across all three germ layers via CRISPR-Cas9 KO screening in the teratoma.

Our time-course analysis of human fetal-to-adult organ data identified novel trends that motivate future investigations into the spatiotemporal role of ADARs beyond brain tissue [80]. For instance, we noted sharp drops in global editing levels during early gestation (4-9 wpc) in both heart and liver tissue. Investigating whether this conserved phenomena is coincidental or driven by convergent underlying mechanisms will be an intriguing avenue for future analysis. Furthermore, the sharp decline in AEI and *ADAR1* expression between years 25-63 of testis development suggests a senescence- or aging-related phenotype, as diminished *ADAR1* expression is associated with senescence [104]. This observation raises the possibility that the testis may exhibit an accelerated-aging phenotype, in relation to the other organs we've studied, given that the testis are the only examined organ tissue in our analysis which exhibits a decrease in *ADAR1* expression in the later time-points. Considering our dataset encompasses individuals up to 63 years old, it would be intriguing to explore time-series transcriptomic profiles in more advanced stages of aging (beyond 63 years of age) to determine whether this decline in *ADAR1* expression is a convergent trend across other organ tissues as well.

Prenatal RNA editing dynamics particularly intrigued us, as they may offer glimpses into the elusive black-box that is mammalian development. Towards this, we characterized RNA editing trends in cerebral organoids and teratomas, as both provide a lab-grown source of prenatal-

to-perinatal tissue that can be experimentally manipulated. We observed that the teratoma is a powerful model system as it allows the exploration of RNA editing profiles across all three germ-layers simultaneously, and faithfully recapitulates RNA editing trends seen in human fetal development. Lastly, through our unbiased, pan-tissue screening in the teratoma, we observed a fitness defect in *ADAR1*-KO cells, that is most strongly seen in mesodermal cells, and uncovered a novel *ADAR1*-mediated role in inhibiting adipogenesis from human MSCs.

ADAR1's role in murine cell adipogenesis has been appreciated in recent years [105,106], but has yet to be explored in a human context. We propose this adipogenic effect is in part driven by a decrease in *ADAR1* expression, putatively leading to an increase in cellular dsRNA. The presence of dsRNA triggers a stress response[107], repressing glycolysis [100,101] and potentially driving adipogenic programs downstream[102]. Lastly, although there is substantial research linking inflammation and obesity [108,109], the link to decreased RNA editing has yet to be explored. Our findings indicate that inflammation triggered via *ADAR1* downregulation, potentially from cellular senescence and physiological aging [104], may be a driver for obesity. Future studies could compare RNA editing levels in adipogenic tissue from obese and non-obese individuals to investigate if these trends hold in an adult organism.

Our findings further corroborate the importance of ADARs in human development and extend previous spatiotemporal editing investigations. We also demonstrated that the teratoma could be used to model prenatal RNA editing dynamics and uncovered the role of ADARs in mesenchymal stem cell fate specification. The link between *ADAR1* dysregulation and adipogenesis could also motivate the development of therapeutics for obesity.

A limitation of this work is the use of 3' biased RNA sequencing libraries, which limited the scope of the examined editome. This leads to a loss of editing information across the RNA

transcript, most notably in re-coding regions. Future studies could involve generating single-cell full-length RNA sequencing libraries, which will enable the tracking of re-coding regions, at the single-cell level.

The epitranscriptome is a powerful regulatory system within the cell that is understudied and holds great potential for discoveries in gene expression dynamics. This study is an extensive investigation of RNA editing and ADAR expression dynamics across various human tissues, providing a unique opportunity to uncover RNA editing trends and the diverse functional roles of ADAR proteins in embryonic development and organogenesis.

Methods

Cell Culture

H1, H9, HUES62, and PGP1 PSC cell lines were maintained under feeder-free conditions in mTeSR1 medium (Stem Cell Technologies 85850). Prior to passaging, tissue-culture plates were coated with growth factor-reduced Matrigel (Corning 356230) diluted in DMEM/F-12 medium (Thermo Fisher Scientific 11320033) and incubated for 30 minutes at 37 ºC, 5% CO2. Cells were dissociated and passaged using the dissociation reagent Versene (Thermo Fisher Scientific 15040066).

Primary human MSCs (Lonza PT-2501) were cultured in MSCGM (Lonza PT-3001) and passaged with Trypsin-EDTA 0.05% (Thermo Fisher Scientific 25300054). To induce adipogenic fate, MSC media was replaced with Stempro Adipogenic Induction Media (Thermo Fisher Scientific A1007001) for 8 days.

PGP1-Cas9 Cell Line Generation

The PGP1-Cas9 line was generated and comprehensively characterized in a previous publication [90].

ADAR Targeting Guide Library Generation

To select the ADAR knockout sgRNA guides sequences, guides that targeted the *ADAR*, *ADARB1* and *ADARB2* genes were taken from the Brunello knockout library [110] and from the Achilles knockout library (https://depmap.org/portal/download/all/), giving a total of eight guides per gene. The guides were then compared using the CRISPR-Cas9 guide RNA design checker from IDT (https://www.idtdna.com/site/order/designtool/index/CRISPR_SEQUENCE), and four guides per gene were selected that had maximum on target editing, and minimal off target editing predictions.

The Gibson assembly reactions were set up as follows: 100 ng digested backbone, 3:10 molar ratio of insert, 2X Gibson assembly master mix (New England Biolabs), H20 up to 20 µl. After incubation at 50 °C for 1 h, the product was transformed into One Shot Stbl3 chemically competent Escherichia coli (Invitrogen). A fraction (150 µL) of cultures was spread on carbenicillin (50 µg/ml) LB plates and incubated overnight at 37 ºC. Individual colonies were picked, introduced into 5 ml of carbenicillin (50 µg/ml) LB medium and incubated overnight in a shaker at 37 ºC. The plasmid DNA was then extracted with a QIAprep Spin Miniprep Kit (Qiagen), and Sanger sequenced to verify correct assembly of the vector. Following verification of the vector, larger amounts of plasmid were obtained by seeding 150 µl of bacterial stock into 150 ml of LB medium containing carbenicillin (50 µg/ml) and incubating overnight in a shaker at 37 ºC for 16-18 h. The plasmid DNA was then extracted using a Plasmid Maxi Kit (Qiagen).

ADAR-KO Library Lentiviral Production

HEK 293T cells were maintained in high glucose DMEM supplemented with 10% fetal bovine serum (FBS). Cells were seeded in a 15 cm plate 1 day prior to transfection, such that they were 60–70% confluent at the time of transfection. For each plate, 36 µl of Lipofectamine 2000

(Thermo Fisher Scientific 11668027) was added to 2 mL of Opti-MEM (Thermo Fisher Scientific 31985062). Separately 3 μg of pMD2.G (Addgene #12259), 12 μg of pCMV delta R8.2 (Addgene #12263) and 9 μg of the pooled gRNA vector library was added to 2 mL of Opti-MEM. After 5 minutes of incubation at room temperature, the Lipofectamine 2000 and DNA solutions were mixed and incubated at room temperature for 30 minutes. During the incubation period, medium in each plate was replaced with 20 ml of fresh, pre-warmed medium per well. After the incubation period, the mixture was added dropwise to each plate of HEK 293T cells. Supernatant containing the viral particles was harvested after 48 and 72 hours, filtered with 0.45 μm filters (Steriflip, Millipore), and further concentrated using Amicon Ultra-15 centrifugal ultrafilters with a 100,000 NMWL cutoff (Millipore) down to a final volume of 600–800 μl, for each 15 cm plate. Finally, the concentrated supernatant was divided into 200 μl aliquots and frozen at −80°C. Two 15 cm plates worth of lentivirus was adequate.

ADAR-KO Library Lentiviral Transduction

For viral transduction, on day −1, PGP1-Cas9 cells were dissociated to a single cell suspension using Accutase and seeded into Matrigel-coated 6-well plates at a density of 300,000 cells per well in mTeSR containing ROCK inhibitor, Y27632 (10 μM, Tocris 1254). The next day, day 0, cells were approximately 20% confluent. Medium containing Y27632 was replaced with mTeSR1 within 16 hours after plating and cells were allowed to recover for at least 8 hours prior to addition of virus.

Recovered cells were then transduced with lentivirus added to fresh mTeSR containing polybrene (5 μg/ml, Millipore TR1003G). On day 1, medium was replaced with fresh mTeSR1. Puromycin (Thermo Fisher Scientific J67236-XF) selection was started from day 2 onward at a

selection dose of 1 μg/ml. Medium containing puromycin was replaced daily, and cells were harvested for teratoma formation after 5 days of selection.

Teratoma Formation

A subcutaneous injection of 6–8 million PSCs in a slurry of growth factor reduced Matrigel (Corning 356230) and mTeSR medium (Stem Cell Technologies 85850) (1:1) was made in the right flank of anesthetized, 6–8 week old, male Rag2−/−;γc−/− immunodeficient mice. Weekly monitoring of teratoma growth was made by quantifying approximate elliptical area (mm2) with the use of calipers measuring outward width and height. When a tumor size of at least 10 mm in any direction, daily monitoring is conducted. Teratomas are grown for up to 10 weeks prior to harvesting.

Teratoma Single Cell Dissociation

Mice were euthanized by slow release of CO_2 followed by secondary means via cervical dislocation. Tumor area was shaved, sprayed with 70% ethanol, and then extracted via surgical excision using scissors and forceps. Tumor was rinsed with PBS, weighed, and photographed. Tumor was then cut into small 1 inch pieces in a semi-random fashion and frozen in OCT for sectioning and H&E staining courtesy of the Moore's Cancer Center Histology Core. Remaining tumor was cut into small pieces 1–2 mm in diameter and subjected to standard GentleMACS (Miltenyi) protocols: Human Tumor Dissociation Kit (medium tumor settings) and Red Blood Cell Lysis Kit. For single cell RNA-seq, samples were also processed with the Dead Cell Removal Kit (Miltenyi). Single cells were then resuspended in PBS + 0.04% BSA for processing on the 10X Genomics Chromium [111] platform and downstream sequencing on an Illumina NovaSeq platform. For bulk processing, after red blood cell removal, cells were divided into multiple tubes and pelleted for 5 minutes at 300g. Supernatant was then removed and cells were either directly frozen

at −80°C or resuspended in RNALater (Thermo Fisher Scientific AM7021), incubated overnight at 4°C, RNALater removed and then frozen at −80°C.

Single Cell Sequencing Library Generation

Cells were passed through a 40 μm filter, spun down at 300 rcf for 5 minutes, and resuspended in 0.04% BSA solution. Cells were then loaded into the Chromium Chip B (10x Genomics) and single cell libraries were made using Chromium Single Cell 3′ Reagent Kits v3 workflow (10x Genomics). Fastq files were aligned to a hg38 reference and expression matrices were generated using the count command in cellranger v3.0.1 (10X Genomics).

ADAR-KO gRNA Amplification

gRNA barcodes were amplified from ADAR-KO teratoma cDNA generated by the 10x single cell platform, and gDNA from the same cells harvested pre-teratoma injection. Amplicons were prepared for deep sequencing through a two-step PCR process.

For amplification of barcodes from cDNA, the first step was performed as four separate 50 µl reactions for each sample. 2.5 µl of the cDNA was input per reaction with Kapa Hifi Hotstart ReadyMix (Kapa Biosystems KK2602). The PCR primers used were, CROPseq_sgRNA_barcodeAmp_Fprimer: GACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTTGTGGAAAGGACGAAACAC and NEBNext Universal PCR Primer for Illumina (New England Biolabs). The thermocycling parameters were 95°C for 3 min; 20–26 cycles of (98°C for 20 s; 65°C for 15 s; and 72°C for 30 s); and a final extension of 72°C for 5 min. The numbers of cycles were tested to ensure that they fell within the linear phase of amplification. Amplicons (~450 bp) of 4 reactions for each sample were pooled, size-selected and purified with Agencourt AMPure XP beads (Beckman Coulter, Inc.) at a 0.8 ratio. The second step of PCR was performed with two separate 50 µl reactions with

50 ng of first step purified PCR product per reaction. NEBNext Multiplex Oligos for Illumina (Dual Index Primers) were used to attach Illumina adapters and indices to the samples. The thermocycling parameters were: 95°C for 3 min; 6 cycles of (98°C for 20 s; 65°C for 15 s; 72°C for 30 s); and 72°C for 5 min. The amplicons from these two reactions for each sample were pooled, size-selected and purified with Agencourt AMPure XP beads at an 0.8 ratio. The purified second-step PCR library was quantified by Qubit dsDNA HS assay (Thermo Fisher Scientific) and used for downstream sequencing on an Illumina HiSeq platform.

For amplification of barcodes from genomic DNA, genomic DNA was extracted from stored cell pellets with a DNeasy Blood and Tissue Kit (Qiagen). The first step PCR was performed as two separate 50 μl reactions for each sample. 500 ng of genomic DNA was input per reaction with Kapa Hifi Hotstart ReadyMix. The PCR primers used were, sgRNAsequenceAmp_gDNA_F: ACACTCTTTCCCTACACGACGCTCTTCCGATCTTATATATCTTGTGGAAGGACGAAA CACCG and sgRNAsequenceAmp_gDNA_R: GACTGGAGTTCAGACGTGTGCTCTTCCGATCTCCTTATTTTAACTTGCTATTTCTAGC TCTA. The thermocycling parameters were: 95°C for 3 min; 24–32 cycles of (98°C for 20 s; 60°C for 15 s; and 72°C for 30 s); and a final extension of 72°C for 5 min. The numbers of cycles were tested to ensure that they fell within the linear phase of amplification. Amplicons (~150 bp) of the two reactions for each sample were pooled, size-selected with Agencourt AMPure XP beads at a ratio of 1.6. The second step of PCR was performed as two separate 50 μl reactions with 25 ng of first step purified PCR product per reaction. NEBNext Multiplex Oligos for Illumina (Dual Index Primers) were used to attach Illumina adapters and indices to the samples. The thermocycling parameters were: 95°C for 3 min; 6–8 cycles of (98°C for 20 s; 65°C for 20 s; 72°C for 30 s); and 72°C for 2 min. The amplicons from these two reactions for each sample were pooled, size-selected

with Agencourt AMPure XP beads at a ratio of 1.6. The purified second-step PCR library was quantified by Qubit dsDNA HS assay (Thermo Fisher Scientific) and used for downstream sequencing on an Illumina NovaSeq platform.

Single Nuclei Sequencing Library Generation

To extract nuclei, teratoma tissues frozen in OCT were processed as previously described (dx.doi.org/10.17504/protocols.io.5qpvoby69l4o/v2). Extracted nuclei were passed through a 50 μm filter, followed by a 30 μm filter and pelleted by spinning at 900 rcf at 4 ºC for 10 minutes. Resuspended nuclei in PBS + 1% BSA solution were stained for DAPI and viable nuclei were collected via FACS. Sorted nuclei were pelleted under the same condition and resuspended in PBS + RNase Inhibitors for generating single nucleus sequencing libraries using Chromium Single Cell 3′ Reagent Kits v3 (10x Genomics).

Cell Type Annotation

Annotations of cell types were carried out by transferring previous annotations on the wild type teratoma using scANVI [112] and Seurat label transfer [113]. Cells with predictive scores greater than 0.8 were annotated and marker gene expression were checked for each annotated cell type and unannotated cell cluster.

Differentially Expressed Gene and Gene Ontology Analysis

The FindMarkers command in the Seurat R package was used to identify differentially expressed marker genes. Wilcoxon rank-sum test was used with other arguments set as the following: min.pct = 0.1, logfc.threshold = 0.25. At least 3 cells were required for comparison. Adjusted p-values were calculated using Benjamin-Hochberg procedure. All differential genes that were statistically significant were inputted into the PANTHER classification system to identify enriched gene ontology terms [114].

AEI RNA Editing Analysis

Raw sequencing reads are parsed to create pseudobulk reads for each cell state, such as cell types, batches of experiment and knockout conditions and the *dedup* function from umitools package is used to remove PCR duplicate reads. *Alu* editing index (AEI) was calculated for deduplicated bam files for each experiment condition using the RNAEditingIndexer v1.0. AEI is an aggregated index representing the ratio of edited reads (A-to-G mismatches) over the total coverage of adenosine bases in *Alu* elements and has been demonstrated as a robust index with respect to coverage, read lengths as well as batch effects introduced by sample preparation. *Alu* repeats outlined by UCSC genome browser were supplied to designate regions where AEI was calculated over and SNP sites included in hg38CommonGenomicSNPs150 were discarded.

Site-Specific RNA Editing Analysis

Editing levels at known sites were calculated similar to existing computational pipelines [80]. First, coordinate-sorted BAM files of mapped, parsed and deduplicated reads were generated with STAR. Then, nucleotide coordinates of known editing sites previously identified in REDIportal were examined with customized scripts with samtools mpileup function [97]. Filters were applied to remove SNP sites in common genomic variation in dbSNP (v150) and sites within 5 base pairs to read ends and splice sites. At least 5 reads covering the site and at least 3 edited reads observed were required to call an editing event for single cell / single nucleus sequencing read pseudobulk samples.

RNA Editing Analysis in Bulk Human Organs

RNA sequencing data for all bulk human organs was retrieved on Array Express via accession code E-MTAB-6814 [91]. Time series data was split into 4 groups, based on their developmental stage annotations: Group 1) 4 weeks post conception (wpc) to 10 wpc were binned

into “early gestation”; Group 2) 11 wpc to 20 wpc were binned into “late gestation”; Group 3) 0 to 20 years of age were binned into “newborn-teeager”; and Group 4) 25-63 years of age were binned into “adult-senior”. Comparisons were made between sequential time series groups - Group 1 versus 2, Group 2 versus 3, and Group 4 versus 4 - in order to quantify the shift in RNA editing and ADAR expression between temporal stage transitions. A two-tailed Student’s t-test was done to determine significance between groups and the Cohen’s d test comparison was run to determine effect size.

AEI measurements were performed as mentioned above.

RNA Editing Analysis in Wild Type Teratomas

RNA sequencing data for teratomas was retrieved via NCBI GEO accession code GSE: 156170. AEI measurements were performed as mentioned above.

RNA Editing Analysis in Human Fetal Organs

RNA sequencing data for fetal organs was retrieved via raw data provided at dbGaP (accession number phs002003.v1.p1). AEI measurements were performed as mentioned above.

RNA Editing Analysis in Cerebral Organoids

RNA sequencing data for cerebral organoids was retrieved via Synapse: syn26346373 [115]. AEI measurements were performed as mentioned above.

Teratoma Screen Guide Calling

To assign the guide identity to each cell in the teratoma screen, the guide count matrix was first obtained from CellRanger v3.0.1. The total number of guides for each cell was calculated and the guide identity for each cell was called for a cell with at least 5 guide counts by assigning the identity of the guide with the maximum number of guide counts expressed. For cells having more

than 1 maximum guide count, the cells were annotated as 'multiplet' and removed from downstream analysis.

Germ-Layer and Pan-Teratoma Fitness Analysis

For germ-layer fitness analysis, the pre-injection frequency values for each unique guide were quantified from pre-injection ADAR-KO cell gDNA counts. After teratoma harvest, cell counts for each unique guide were added up for each germ layer, and then divided by the total number of cells within that germ layer. This frequency (post-harvest frequency) was divided by the pre-injection guide frequency (pre-injection frequency) of each unique guide, and then transformed with log2-scaling, resulting in the log2(Fold-Change) value. Within each teratoma, log2(Fold-Change) values for unique guides targeting the same gene were averaged, and n=4 log2(Fold-Change) values were plotted using 4 unique ADAR-KO teratomas.

For pan-teratoma fitness analysis, the total number of cells for each unique guide were added up, and then divided by the total number of cells in the teratoma. Subsequent calculations were conducted as described above.

WGCNA Gene Set Analysis

To contrast the effects on the transcriptome in the knock-out screen, co-expressed genes for each cell type were identified and the gene expression were gauged for each perturbation using the WGCNA gene set testing pipeline [99]. First, gene modules of the co-expressed genes within each cell type were identified using WGCNA R package [116]. We then calculated the gene-set score for each WGCNA module and used the R function lm to construct a linear regression model for examining how the gene-set scores vary in each perturbation. Teratoma batch differences and sequencing depths variances were corrected. FDR corrected p-values and effect sizes from the linear regression test were calculated.

Teratoma Cell-Type Enrichment Analysis

To compare the cell type composition in the knock-out screen, we adopted a previously established pipeline [99]. Specifically, a Poisson regression model was used to measure the effect size of cell type composition changes with corrected experimental batch differences and extracted p-values were corrected using Benjamin-Hochberg procedure.

shRNA Lentiviral Production and Transduction of MSCs

shControl (Sigma Aldrich SHC002) and shADAR1 (Sigma Aldrich TRCN0000050788) plasmids were packaged into lenvirus as described above.

For viral transduction, on day −5, passage 4 MSCs were dissociated to a single cell suspension and seeded into 48 well plates at a density of 5,000 cells per well in MSCGM. On day -4 cells were then transduced with lentivirus added to fresh MSCGM containing polybrene (5 μg/ml). Media was replaced with fresh MSCGM on day -3, and then again on day -2 with MSCGM + Puromycin (10 μg/ml, Thermo Fisher Scientific). After 2 days of selection, media is replaced with Stempro Adipogenic Induction Media + Puromycin (10 μg/ml) to induce adipogenic fate, and media is replaced every 2-3 days.

Oil O Red and DAPI Staining

3.5 mg/mL Oil Red O Stock Solution (Sigma Aldrich O-0625) is made by dissolving Oil O Red in 100% isopropanol. Oil O Red working solution is made by mixing the Oil O red stock solution with ddH20 at 3:2 ratio. Oil O red working solution is set to sit at room temp for 20 min followed by filtering (0.2 μm).

Cells are fixed with 4% Paraformaldehyde for 1 hour at RT, washed 3x in ddH20, followed by applying Oil O red working solution for 10 minutes at RT. Cells are rinsed with ddH20 3x,

followed by adding 1 μg/mL of DAPI solution (Thermo Fisher Scientific 62248). Cells are finally rinsed with ddH20 3x and subsequently imaged via brightfield and fluorescence microscopy.

Chapter 5 - Conclusion: Future Directions and Closing Remarks

Closing Remarks

In conclusion, this thesis has delved into the multifaceted realm of utilizing pluripotent stem cells for the purpose of engineering multi-lineage and vascularized tissues. Through a comprehensive exploration of in vitro organoid engineering and in vivo teratoma engineering, this research has contributed valuable insights into the potential of pluripotent stem cells in recreating complex tissue architectures, through two orthogonal methodologies: an in vitro and in vivo method. The in vitro method was a novel approach which combined transcription factor overexpression and organoid technology to derive multi-lineage, vascular organoids, while the in vivo method leveraged the multi-lineage and vascularized nature of an in vivo derived teratoma and developed a method for sculpting away undesired teratoma tissue. Although these methods use orthogonal starting points for organ engineering, they both involve the integration of rational engineering onto a self-organized tissue-construct starting material. In my work, I demonstrate the importance of leveraging the powerful self-organizing capabilities of stem cells while applying rational engineering principles to make the difference between naturally derived organoid and teratomas, and real, large-scale human organs.

Future Directions

The application of these methodologies in disease modeling opens new avenues for understanding pathophysiological processes and testing therapeutic interventions. Moreover, the future directions of this research hold immense promise. One crucial aspect lies in the demonstration of long-term maturation of therapeutically relevant tissues. This entails developing strategies to sustain the growth and maturation of engineered tissues over extended periods, addressing a critical challenge in the field. Furthermore, the application of these technologies to

urgently needed transplantable organs, such as the pancreas, liver, and kidney, presents an exciting frontier. The extension of these methodologies to these vital organs holds the potential to alleviate the global organ shortage crisis and significantly impact patient outcomes.

As we step into the future, it is imperative to continue refining and expanding the applications of pluripotent stem cell technologies. Efforts should focus on optimizing culture conditions, enhancing vascularization strategies, and establishing robust systems for the long-term maintenance and maturation of tissues. Collaborative endeavors with multidisciplinary teams, including bioengineers, clinicians, and molecular biologists, will be essential to bridge the gap between bench and bedside. By unlocking the full potential of pluripotent stem cells, we embark on a journey toward realizing the transformative impact of regenerative medicine on human health, offering hope for patients in need of life-saving organ transplants and pushing the boundaries of what is achievable in the realm of tissue engineering.

REFERENCES

1. Lancaster MA, Knoblich JA. Organogenesis in a dish: modeling development and disease using organoid technologies. Science. 2014;345: 1247125.

2. Clevers H. Modeling Development and Disease with Organoids. Cell. 2016;165: 1586–1597.

3. Takebe T, Wells JM. Organoids by design. Science. 2019;364: 956–959.

4. Daniel E, Cleaver O. Vascularizing organogenesis: Lessons from developmental biology and implications for regenerative medicine. Curr Top Dev Biol. 2019;132: 177–220.

5. Petrova TV, Koh GY. Organ-specific lymphatic vasculature: From development to pathophysiology. J Exp Med. 2018;215: 35–49.

6. Grebenyuk S, Ranga A. Engineering Organoid Vascularization. Front Bioeng Biotechnol. 2019;7: 39.

7. Lancaster MA. Brain organoids get vascularized. Nat Biotechnol. 2018;36: 407–408.

8. Guye P, Ebrahimkhani MR, Kipniss N, Velazquez JJ, Schoenfeld E, Kiani S, et al. Genetically engineering self-organization of human pluripotent stem cells into a liver bud-like tissue using Gata6. Nat Commun. 2016;7: 10243.

9. Cakir B, Xiang Y, Tanaka Y, Kural MH, Parent M, Kang Y-J, et al. Engineering of human brain organoids with a functional vascular-like system. Nat Methods. 2019;16: 1169–1175.

10. Garreta E, Prado P, Tarantino C, Oria R, Fanlo L, Martí E, et al. Fine tuning the extracellular environment accelerates the derivation of kidney organoids from human pluripotent stem cells. Nat Mater. 2019;18: 397–405.

11. Homan KA, Gupta N, Kroll KT, Kolesky DB, Skylar-Scott M, Miyoshi T, et al. Flow-enhanced vascularization and maturation of kidney organoids in vitro. Nat Methods. 2019;16: 255–262.

12. Low JH, Li P, Chew EGY, Zhou B, Suzuki K, Zhang T, et al. Generation of Human PSC-Derived Kidney Organoids with Patterned Nephron Segments and a De Novo Vascular Network. Cell Stem Cell. 2019;25: 373–387.e9.

13. Mansour AA, Gonçalves JT, Bloyd CW, Li H, Fernandes S, Quang D, et al. An in vivo model of functional and vascularized human brain organoids. Nat Biotechnol. 2018;36: 432–441.

14. Pham MT, Pollock KM, Rose MD, Cary WA, Stewart HR, Zhou P, et al. Generation of human vascularized brain organoids. Neuroreport. 2018;29: 588–593.

15. Lindgren AG, Veldman MB, Lin S. ETV2 expression increases the efficiency of primitive endothelial cell derivation from human embryonic stem cells. Cell Regen. 2015;4: 1.

16. Parekh U, Wu Y, Zhao D, Worlikar A, Shah N, Zhang K, et al. Mapping Cellular Reprogramming via Pooled Overexpression Screens with Paired Fitness and Single-Cell RNA-Sequencing Readout. Cell Syst. 2018;7: 548–555.e8.

17. Ferland-McCollough D, Slater S, Richard J, Reni C, Mangialardi G. Pericytes, an overlooked player in vascular pathobiology. Pharmacol Ther. 2017;171: 30–42.

18. Wimmer RA, Leopoldi A, Aichinger M, Wick N, Hantusch B, Novatchkova M, et al. Human blood vessel organoids as a model of diabetic vasculopathy. Nature. 2019;565: 505–510.

19. Zhang Y, Pak C, Han Y, Ahlenius H, Zhang Z, Chanda S, et al. Rapid single-step induction of functional neurons from human pluripotent stem cells. Neuron. 2013;78: 785–798.

20. Yang N, Chanda S, Marro S, Ng Y-H, Janas JA, Haag D, et al. Generation of pure GABAergic neurons by transcription factor programming. Nat Methods. 2017;14: 621–628.

21. Albini S, Coutinho P, Malecova B, Giordani L, Savchenko A, Forcales SV, et al. Epigenetic reprogramming of human embryonic stem cells into skeletal muscle cells and generation of contractile myospheres. Cell Rep. 2013;3: 661–670.

22. Yusa K, Zhou L, Li MA, Bradley A, Craig NL. A hyperactive piggyBac transposase for mammalian applications. Proc Natl Acad Sci U S A. 2011;108: 1531–1536.

23. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol. 2018;36: 411–420.

24. Schörnig M, Ju X, Fast L, Ebert S, Weigert A, Kanton S, et al. Comparison of induced neurons reveals slower structural and functional maturation in humans than in apes. Elife. 2021;10. doi:10.7554/eLife.59323

25. Rao L, Qian Y, Khodabukus A, Ribar T, Bursac N. Engineering human pluripotent stem cells into a functional skeletal muscle tissue. Nat Commun. 2018;9: 126.

26. Khodabukus A, Madden L, Prabhu NK, Koves TR, Jackman CP, Muoio DM, et al. Electrical stimulation increases hypertrophy and metabolic flux in tissue-engineered human skeletal muscle. Biomaterials. 2019;198: 259–269.

27. Qian H, Kang X, Hu J, Zhang D, Liang Z, Meng F, et al. Reversing a model of Parkinson's disease with in situ converted nigral neurons. Nature. 2020;582: 550–556.

28. Polstein LR, Gersbach CA. A light-inducible CRISPR-Cas9 system for control of endogenous gene activation. Nat Chem Biol. 2015;11: 198–200.

29. Nihongaki Y, Furuhata Y, Otabe T, Hasegawa S, Yoshimoto K, Sato M. CRISPR-Cas9-based photoactivatable transcription systems to induce neuronal differentiation. Nat Methods. 2017;14: 963–966.

30. Nihongaki Y, Yamamoto S, Kawano F, Suzuki H, Sato M. CRISPR-Cas9-based photoactivatable transcription system. Chem Biol. 2015;22: 169–174.

31. Toda S, Blauch LR, Tang SKY, Morsut L, Lim WA. Programming self-organizing multicellular structures with synthetic cell-cell signaling. Science. 2018;361: 156–162.

32. Liu XS, Wu H, Ji X, Stelzer Y, Wu X, Czauderna S, et al. Editing DNA Methylation in the Mammalian Genome. Cell. 2016;167: 233–247.e17.

33. Victor MB, Richner M, Hermanstyne TO, Ransdell JL, Sobieski C, Deng P-Y, et al. Generation of human striatal neurons by microRNA-dependent direct conversion of fibroblasts. Neuron. 2014;84: 311–323.

34. Wang W, Xue Y, Zhou S, Kuo A, Cairns BR, Crabtree GR. Diversity and specialization of mammalian SWI/SNF complexes. Genes Dev. 1996;10: 2117–2130.

35. McGinnis CS, Murrow LM, Gartner ZJ. DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. Cell Syst. 2019;8: 329–337.e4.

36. Han X, Wang R, Zhou Y, Fei L, Sun H, Lai S, et al. Mapping the Mouse Cell Atlas by Microwell-Seq. Cell. 2018;172: 1091–1107.e17.

37. Gaertner B, Carrano AC, Sander M. Human stem cell models: lessons for pancreatic development and disease. Genes Dev. 2019;33: 1475–1490.

38. Yao Z, Mich JK, Ku S, Menon V, Krostag A-R, Martinez RA, et al. A Single-Cell Roadmap of Lineage Bifurcation in Human ESC Models of Embryonic Brain Development. Cell Stem Cell. 2017;20: 120–134.

39. Wang J, Jenjaroenpun P, Bhinge A, Angarica VE, Del Sol A, Nookaew I, et al. Single-cell gene expression analysis reveals regulators of distinct cell subpopulations among developing human neurons. Genome Res. 2017;27: 1783–1794.

40. Ader M, Tanaka EM. Modeling human development in 3D culture. Curr Opin Cell Biol. 2014;31: 23–28.

41. Faustino Martins J-M, Fischer C, Urzi A, Vidal R, Kunz S, Ruffault P-L, et al. Self-Organizing 3D Human Trunk Neuromuscular Organoids. Cell Stem Cell. 2020;26: 172–186.e6.

42. Völkner M, Zschätzsch M, Rostovskaya M, Overall RW, Busskamp V, Anastassiadis K, et al. Retinal Organoids from Pluripotent Stem Cells Efficiently Recapitulate Retinogenesis. Stem Cell Reports. 2016;6: 525–538.

43. Prior N, Inacio P, Huch M. Liver organoids: from basic research to therapeutic applications. Gut. 2019;68: 2228–2237.

44. Weatherbee BAT, Cui T, Zernicka-Goetz M. Modeling human embryo development with embryonic and extra-embryonic stem cells. Dev Biol. 2021;474: 91–99.

45. Tzouanacou E, Wegener A, Wymeersch FJ, Wilson V, Nicolas J-F. Redefining the progression of lineage segregations during mammalian embryogenesis by clonal analysis. Dev Cell. 2009;17: 365–376.

46. Zhu M, Zernicka-Goetz M. Principles of Self-Organization of the Mammalian Embryo. Cell. 2020;183: 1467–1478.

47. An Y, Sekinaka T, Tando Y, Okamura D, Tanaka K, Ito-Matsuoka Y, et al. Derivation of pluripotent stem cells from nascent undifferentiated teratoma. Dev Biol. 2019;446: 43–55.

48. Alberti C, Cochella L. A framework for understanding the roles of miRNAs in animal development. Development. 2017;144: 2548–2559.

49. Zeng Y. Principles of micro-RNA production and maturation. Oncogene. 2006;25: 6156–6162.

50. Ivey KN, Srivastava D. microRNAs as Developmental Regulators. Cold Spring Harb Perspect Biol. 2015;7: a008144.

51. Wu M-R, Jusiak B, Lu TK. Engineering advanced cancer therapies with synthetic biology. Nat Rev Cancer. 2019;19: 187–195.

52. Straathof KC, Pulè MA, Yotnda P, Dotti G, Vanin EF, Brenner MK, et al. An inducible caspase 9 safety switch for T-cell therapy. Blood. 2005;105: 4247–4254.

53. Coolen M, Katz S, Bally-Cuif L. miR-9: a versatile regulator of neurogenesis. Front Cell Neurosci. 2013;7: 220.

54. Radhakrishnan B, Alwin Prem Anand A. Role of miRNA-9 in Brain Development. J Exp Neurosci. 2016;10: 101–120.

55. Nishikura K. Functions and regulation of RNA editing by ADAR deaminases. Annu Rev Biochem. 2010;79: 321–349.

56. Jain M, Jantsch MF, Licht K. The Editor's I on Disease Development. Trends Genet. 2019;35: 903–913.

57. Samuel CE. Adenosine deaminases acting on RNA (ADARs) are both antiviral and proviral. Virology. 2011;411: 180–193.

58. Reich DP, Tyc KM, Bass BL. C. elegans ADARs antagonize silencing of cellular dsRNAs by the antiviral RNAi pathway. Genes Dev. 2018;32: 271–282.

59. Eisenberg E, Levanon EY. A-to-I RNA editing - immune protector and transcriptome diversifier. Nat Rev Genet. 2018;19: 473–490.

60. Bass BL. RNA editing by adenosine deaminases that act on RNA. Annu Rev Biochem. 2002;71: 817–846.

61. Wang Y, Zheng Y, Beal PA. Adenosine Deaminases That Act on RNA (ADARs). Enzymes. 2017;41: 215–268.

62. Tan MH, Li Q, Shanmugam R, Piskol R, Kohler J, Young AN, et al. Dynamic landscape and regulation of RNA editing in mammals. Nature. 2017;550: 249–254.

63. Rosenthal JJC, Eisenberg E. Extensive Recoding of the Neural Proteome in Cephalopods by RNA Editing. Annu Rev Anim Biosci. 2023;11: 57–75.

64. Wang Y, Chung DH, Monteleone LR, Li J, Chiang Y, Toney MD, et al. RNA binding candidates for human ADAR3 from substrates of a gain of function mutant expressed in neuronal cells. Nucleic Acids Res. 2019;47: 10801–10814.

65. Eisenberg E. Proteome Diversification by RNA Editing. Methods Mol Biol. 2021;2181: 229–251.

66. Gabay O, Shoshan Y, Kopel E, Ben-Zvi U, Mann TD, Bressler N, et al. Landscape of adenosine-to-inosine RNA recoding across human tissues. Nat Commun. 2022;13: 1184.

67. Ohman M. A-to-I editing challenger or ally to the microRNA process. Biochimie. 2007;89: 1171–1176.

68. Wang IX, So E, Devlin JL, Zhao Y, Wu M, Cheung VG. ADAR regulates RNA editing, transcript stability, and gene expression. Cell Rep. 2013;5: 849–860.

69. Kapoor U, Licht K, Amman F, Jakobi T, Martin D, Dieterich C, et al. ADAR-deficiency perturbs the global splicing landscape in mouse tissues. Genome Res. 2020;30: 1107–1118.

70. Valente L, Nishikura K. ADAR gene family and A-to-I RNA editing: diverse roles in posttranscriptional gene regulation. Prog Nucleic Acid Res Mol Biol. 2005;79: 299–338.

71. Hu S-B, Heraud-Farlow J, Sun T, Liang Z, Goradia A, Taylor S, et al. ADAR1p150 prevents MDA5 and PKR activation via distinct mechanisms to avert fatal autoinflammation. Mol Cell. 2023. doi:10.1016/j.molcel.2023.09.018

72. Hubbard NW, Ames JM, Maurano M, Chu LH, Somfleth KY, Gokhale NS, et al. ADAR1 mutation causes ZBP1-dependent immunopathology. Nature. 2022;607: 769–775.

73. Fritzell K, Xu L-D, Lagergren J, Öhman M. ADARs and editing: The role of A-to-I RNA modification in cancer progression. Semin Cell Dev Biol. 2018;79: 123–130.

74. Tan TY, Sedmík J, Fitzgerald MP, Halevy RS, Keegan LP, Helbig I, et al. Bi-allelic ADARB1 Variants Associated with Microcephaly, Intellectual Disability, and Seizures. Am J Hum Genet. 2020;106: 467–483.

75. Singh M. Dysregulated A to I RNA editing and non-coding RNAs in neurodegeneration. Front Genet. 2012;3: 326.

76. Toth AM, Li Z, Cattaneo R, Samuel CE. RNA-specific adenosine deaminase ADAR1 suppresses measles virus-induced apoptosis and activation of protein kinase PKR. J Biol Chem. 2009;284: 29350–29356.

77. Slotkin W, Nishikura K. Adenosine-to-inosine RNA editing and human disease. Genome Med. 2013;5: 105.

78. Rice GI, Kasher PR, Forte GMA, Mannion NM, Greenwood SM, Szynkiewicz M, et al. Mutations in ADAR1 cause Aicardi-Goutières syndrome associated with a type I interferon signature. Nat Genet. 2012;44: 1243–1248.

79. Picardi E, D'Erchia AM, Lo Giudice C, Pesole G. REDIportal: a comprehensive database of A-to-I RNA editing events in humans. Nucleic Acids Res. 2017;45: D750–D757.

80. Cuddleston WH, Fan X, Sloofman L, Liang L, Mossotto E, Moore K, et al. Spatiotemporal and genetic regulation of A-to-I editing throughout human brain development. Cell Rep. 2022;41: 111585.

81. Niescierowicz K, Pryszcz L, Navarrete C, Tralle E, Sulej A, Abu Nahia K, et al. Adar-mediated A-to-I editing is required for embryonic patterning and innate immune response regulation in zebrafish. Nat Commun. 2022;13: 5520.

82. Moore JB 4th, Sadri G, Fischer AG, Weirick T, Militello G, Wysoczynski M, et al. The A-to-I RNA Editing Enzyme Adar1 Is Essential for Normal Embryonic Cardiac Growth and Development. Circ Res. 2020;127: 550–552.

83. Tsivion-Visbord H, Kopel E, Feiglin A, Sofer T, Barzilay R, Ben-Zur T, et al. Increased RNA editing in maternal immune activation model of neurodevelopmental disease. Nat Commun. 2020;11: 5236.

84. Lu D, Lu J, Liu Q, Zhang Q. Emerging role of the RNA-editing enzyme ADAR1 in stem cell fate and function. Biomark Res. 2023;11: 61.

85. Hartner JC, Schmittwolf C, Kispert A, Müller AM, Higuchi M, Seeburg PH. Liver disintegration in the mouse embryo caused by deficiency in the RNA-editing enzyme ADAR1. J Biol Chem. 2004;279: 4894–4902.

86. Liddicoat BJ, Piskol R, Chalk AM, Ramaswami G, Higuchi M, Hartner JC, et al. RNA editing by ADAR1 prevents MDA5 sensing of endogenous dsRNA as nonself. Science. 2015;349: 1115–1120.

87. Chalk AM, Taylor S, Heraud-Farlow JE, Walkley CR. The majority of A-to-I RNA editing is not required for mammalian homeostasis. Genome Biol. 2019;20: 268.

88. Terajima H, Yoshitane H, Ozaki H, Suzuki Y, Shimba S, Kuroda S, et al. ADARB1 catalyzes circadian A-to-I editing and regulates RNA rhythm. Nat Genet. 2017;49: 146–151.

89. Deininger P. Alu elements: know the SINEs. Genome Biol. 2011;12: 236.

90. McDonald D, Wu Y, Dailamy A, Tat J, Parekh U, Zhao D, et al. Defining the Teratoma as a Model for Multi-lineage Human Development. Cell. 2020;183: 1402–1419.e18.

91. Cardoso-Moreira M, Halbert J, Valloton D, Velten B, Chen C, Shao Y, et al. Gene expression across mammalian organ development. Nature. 2019;571: 505–509.

92. Roth SH, Levanon EY, Eisenberg E. Genome-wide quantification of ADAR adenosine-to-inosine RNA editing activity. Nat Methods. 2019;16: 1131–1138.

93. Cao J, O'Day DR, Pliner HA, Kingsley PD, Deng M, Daza RM, et al. A human cell atlas of fetal gene expression. Science. 2020;370. doi:10.1126/science.aba7721

94. Cuddleston WH, Li J, Fan X, Kozenkov A, Lalli M, Khalique S, et al. Cellular and genetic drivers of RNA editing variation in the human brain. Nat Commun. 2022;13: 2997.

95. Uzquiano A, Kedaigle AJ, Pigoni M, Paulsen B, Adiconis X, Kim K, et al. Proper acquisition of cell class identity in organoids allows definition of fate specification programs of the human cerebral cortex. Cell. 2022;185: 3770–3788.e27.

96. Buechler MB, Pradhan RN, Krishnamurty AT, Cox C, Calviello AK, Wang AW, et al. Cross-tissue organization of the fibroblast lineage. Nature. 2021;593: 575–579.

97. Mansi L, Tangaro MA, Lo Giudice C, Flati T, Kopel E, Schaffer AA, et al. REDIportal: millions of novel A-to-I RNA editing events from thousands of RNAseq experiments. Nucleic Acids Res. 2021;49: D1012–D1019.

98. Cui J, Zhang M, Zhang Y-Q, Xu Z-H. JNK pathway: diseases and therapeutic potential. Acta Pharmacol Sin. 2007;28: 601–608.

99. Jin X, Simmons SK, Guo A, Shetty AS, Ko M, Nguyen L, et al. In vivo Perturb-Seq reveals neuronal and glial abnormalities associated with autism risk genes. Science. 2020;370. doi:10.1126/science.aaz6063

100. Zhang W, Wang G, Xu Z-G, Tu H, Hu F, Dai J, et al. Lactate Is a Natural Suppressor of RLR Signaling by Targeting MAVS. Cell. 2019;178: 176–189.e15.

101. He Q-Q, Huang Y, Nie L, Ren S, Xu G, Deng F, et al. MAVS integrates glucose metabolism and RIG-I-like receptor signaling. Nat Commun. 2023;14: 5343.

102. Jiang B, Huang L, Tian T, Wu H, Yao H, Marmo T, et al. IRX5 promotes adipogenesis of hMSCs by repressing glycolysis. Cell Death Discov. 2022;8: 204.

103. Chung H, Calis JJA, Wu X, Sun T, Yu Y, Sarbanes SL, et al. Human ADAR1 Prevents Endogenous RNA from Triggering Translational Shutdown. Cell. 2018;172: 811–824.e14.

104. Hao X, Shiromoto Y, Sakurai M, Towers M, Zhang Q, Wu S, et al. ADAR1 downregulation by autophagy drives senescence independently of RNA editing by enhancing p16INK4a levels. Nat Cell Biol. 2022;24: 1202–1210.

105. Yu Z, Luo R, Li Y, Li X, Yang Z, Peng J, et al. ADAR1 inhibits adipogenesis and obesity by interacting with Dicer to promote the maturation of miR-155-5P. J Cell Sci. 2022;135. doi:10.1242/jcs.259333

106. Cui X-B, Fei J, Chen S, Edwards GL, Chen S-Y. ADAR1 deficiency protects against high-fat diet-induced obesity and insulin resistance in mice. Am J Physiol Endocrinol Metab. 2021;320: E131–E138.

107. George CX, Ramaswami G, Li JB, Samuel CE. Editing of Cellular Self-RNAs by Adenosine Deaminase ADAR1 Suppresses Innate Immune Stress Responses. J Biol Chem. 2016;291: 6158–6168.

108. Ellulu MS, Patimah I, Khaza'ai H, Rahmat A, Abed Y. Obesity and inflammation: the linking mechanism and the complications. Arch Med Sci. 2017;13: 851–863.

109. Saltiel AR, Olefsky JM. Inflammatory mechanisms linking obesity and metabolic disease. J Clin Invest. 2017;127: 1–4.

110. Doench JG, Fusi N, Sullender M, Hegde M, Vaimberg EW, Donovan KF, et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nat Biotechnol. 2016;34: 184–191.

111. Zheng GXY, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, et al. Massively parallel digital transcriptional profiling of single cells. Nat Commun. 2017;8: 14049.

112. Xu C, Lopez R, Mehlman E, Regier J, Jordan MI, Yosef N. Probabilistic harmonization and annotation of single-cell transcriptomics data with deep generative models. Mol Syst Biol. 2021;17: e9620.

113. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM 3rd, et al. Comprehensive Integration of Single-Cell Data. Cell. 2019;177: 1888–1902.e21.

114. Thomas PD, Ebert D, Muruganujan A, Mushayahama T, Albou L-P, Mi H. PANTHER: Making genome-scale phylogenetics accessible to all. Protein Sci. 2022;31: 8–22.

115. Trujillo CA, Gao R, Negraes PD, Gu J, Buchanan J, Preissl S, et al. Complex Oscillatory Waves Emerging from Cortical Organoids Model Early Human Brain Network Development. Cell Stem Cell. 2019;25: 558–569.e7.

116. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9: 559.