UNIVERSITY OF CALIFORNIA SAN DIEGO

Programmable systems to engineer human tissues and electrophysiological sensors

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in

Electrical Engineering (Medical Devices and Systems)

by

Udit Parekh

Committee in charge:

Professor Tse Nga Ng, Chair
Professor Prashant Mali, Co-Chair
Professor Duygu Kuzum
Professor Yu-Hwa Lo
Professor Kun Zhang

2021

The dissertation of Udit Parekh is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2021

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

ACKNOWLEDGEMENTS

Working toward a PhD at UCSD has been equal parts exhilarating and challenging. It has been one of the most intellectually exciting and enriching times I have ever had, and it was enabled by the help, support and kindness of many, throughout my time at UCSD.

First, I cannot thank enough my co-advisers, Tina Ng and Prashant Mali. Without their generosity, mentorship, advice and encouragement this dissertation would not have been possible. Their mentorship has been central to shaping who I am as a scientist, academic and colleague. Their work ethic, curiosity, creativity and energy are a constant inspiration and they have been the best role models for my own research journey. They have both given generously of their time and energy, patiently entertaining my most outlandish ideas and indulging my every scientific fancy.

Tina is the kindest and warmest of mentors, and has been a source of calm, clarity and reassurance throughout. Her excitement about new ideas, combined with her deep knowledge, pragmatism and systematic, insightful analysis are characteristics that I aspire to. I am grateful and immensely privileged to have been able to work with her and to have her ear on everything.

Prashant has been the most engaged and encouraging of mentors, always supporting ambitious goals. His unlimited reservoir of scientific curiosity, readiness to dive into new fields, drive and vision are an inspiration, and a model which I will look to as I set out on my own independent career. His thoughtfulness about mentorship, running a lab and support for independent ideas are all pieces which I will carry with me for the future.

I would like to thank all the members of my committee – Prof. Kun Zhang, Prof. Duygu Kuzum and Prof. Yu-Hwa Lo for their encouragement and support. They have always been ready to help and respond even when I write to them out of the blue. I would like to thank Kun especially,

for his always insightful comments during our collaborations. I hope to be as precise in cutting through details to identify key scientific elements, which have helped to make my projects better.

The work that I have undertaken over the course of my time at UCSD has been enriched, sustained, and sometimes rescued by the help, advice and contributions of many able collaborators across both labs and across UCSD. My deepest thanks to all of them: Amir Dailamy, Yan Wu, Kaiping Wang, Dongxin Zhao, Atharv Worlikar, Neha Shah, Daniella McDonald, Dhruva Katrekar, Ana Moreno, Aditya Kumar, Tejaswy Pailla, Michael Hu, Linda Lei, Ellie Sadatian; Marianna Yusupova; Prof. Brian Eliceiri, Prof. Todd Costantini, Prof. Harinath Garudadri, Prof. Vikash Gilja, Dr Ann Tipps, Thekla Cordes, Prof. Christian Metallo. I would also like to thank my collaborators outside UCSD for their inputs and discussions: Jonathan Ting, Natasha Yamamoto, Juan Zhu and Prof. Ana Arias at UC Berkeley, and Prof Kye-si Kwon at Soonchunhyang University.

The members of both labs have been invaluable allies, friends and partners in this intellectual and emotional journey. I cannot say enough about how fantastic they have been. I am especially indebted to: Amir Dailamy for being an amazing collaborator and friend, the last few years would not have been as much fun without your zany adventures; Yan Wu for many wide-ranging chats and spending time generously helping me learn a tiny bit of computational biology; Kaiping Wang for his steadiness and dedication in driving our projects forward; Daniella McDonald for making even tumor dissections a fun time; Aditya Kumar for many long talks and commiserations about stem cells and reprogramming; Yichen Zhai for his unrivalled enthusiasm and sharp observations about 3D printing; Lulu Yao for always being ready to assist in my struggles with electrochemical analysis; Li Ning for his help with my struggles with the glove box and evaporator; Dongxin Zhao for her friendship, mentorship, persistence and patience; Marianna Yusupova for constant good cheer and always being ready to try new things; and Dhruva Katrekar

and Ana Moreno, who have been the closest of friends from the beginning of my time at UCSD, always present for support, celebration and procrastination.

And finally I would like to thank my family. My parents, Rupal and Nayan Parekh, who have supported everything I have ever wanted to do, have encouraged every step of my education and remain dedicated believers in me, sometimes in the face of evidence. My brother, Rachit, with whom I may disagree often but who will still have my back.

My partner, Deepti Chatti, my closest confidant, best friend and untiring supporter of me and my work. Her presence in my life has filled it with the greatest happiness. No amount of thanks is enough for the many ways in which she has brought joy, meaning and thoughtfulness into our lives. A final and heartiest thanks to our wonderful cat, Chico, who has moved with us from Bangalore and through our various graduate school upheavals with equanimity and love. He brings lightness and joy everywhere and has even approved a few manuscript drafts with a swish of the tail and swipe of the paw. Deepti and Chico, thank you for everything.

Chapters 2-6 are in part reprints of the following materials of which the dissertation author was the lead author or one of the two primary authors:

Chapter 2: Parekh, U.*, Wu, Y.*, Zhao, D., Worlikar, A., Shah, N., Zhang, K. & Mali, P. Mapping Cellular Reprogramming via Pooled Overexpression Screens with Paired Fitness and Single Cell RNA-Sequencing Readout, *Cell Systems*, 7 (5), 548-555.e8 (2018) (* equal contribution).

Chapter 3: Parekh, U., McDonald, D., Dailamy, A., Wu, Y., Cordes, T., Zhang, K., Tipps, A., Metallo, C., Mali, P. Charting oncogenicity of genes and variants across lineages via multiplexed screens in teratomas (2021), currently under peer review.

Chapter 4: Dailamy, A.*, Parekh, U.*, Katrekar, D., McDonald, D., Kumar, A., Moreno, A., Bagheri, P., Ng, T. N. & Mali, P. Programmatic Introduction of Parenchymal Cell Types in Blood Vessel Organoids (2021), currently under peer review.

Chapter 5: Wang, K., Parekh, U., Pailla, T., Garudadri, H., Gilja, V. & Ng, T.N. Stretchable Dry Electrodes with Concentric Ring Geometry for Enhancing Spatial Resolution in Electrophysiology. *Advanced Healthcare Materials* 6 (19), 1700552 (2017)

Chapter 6: Wang, K.*, Parekh, U.*, Ting, J. K., Yamamoto, N. A. D., Zhu, J., Costantini, T., Arias, A. C., Eliceiri, B. P. & Ng, T. N. A Platform to Study the Effects of Electrical Stimulation on Immune Cell Activation During Wound Healing, *Advanced Biosystems* 3 (10), 1900106 (2019) (* equal contribution).

VITA

2005 Bachelor of Technology in Engineering Physics, Indian Institute of Technology Bombay

2007 Master of Electrical Engineering, Rice University

2021 Doctor of Philosophy in Electrical Engineering (Medical Devices and Systems), University of California San Diego

PUBLICATIONS

1. **Parekh, U.,** McDonald, D., Dailamy, A., Wu, Y., Cordes, T., Tipps, A., Zhang, K., Metallo, C. & Mali, P. Charting oncogenicity of genes and variants across lineages via multiplexed screens in teratomas *bioRxiv* (2021), currently under peer review.
2. Dailamy, A.*, **Parekh, U.***, Katrekar, D., McDonald, D., Kumar, A., Moreno, A., Bagheri, P., Ng, T. N. & Mali, P. Programmatic Introduction of Parenchymal Cell Types in Blood Vessel Organoids, currently under peer review. **equal contribution*
3. Moreno, A. M., Aleman, F., Catroli, G. F., Hunt, M., Pla, A., Woller, S. A., Palmer, N., **Parekh, U.**, Hu, M., Goodwill, V., Dryden, I., Hevner, R. F., Yaksh, T. & Mali, P. Long-lasting Analgesia via Targeted in vivo Epigenetic Repression of NaV1.7, *Science Translational Medicine* 13 (584) eaay9056 (2021).
4. McDonald, D*., Wu, Y*., Dailamy, A., Tat, J., **Parekh, U.**, Zhao, D., Hu, M., Tipps, A., Zhang, K. & Mali, P. Defining the Teratoma as a Model for Multi-Lineage Human Development, *Cell* 183 (5), 1402-1419 (2020) **equal contribution*
5. Wang, K.*, **Parekh, U.***, Ting, J. K., Yamamoto, N. A. D., Zhu, J., Costantini, T., Arias, A. C., Eliceiri, B. P. & Ng, T. N. A Platform to Study the Effects of Electrical Stimulation on

Immune Cell Activation During Wound Healing, *Advanced Biosystems* 3 (10), 1900106 (2019) **equal contribution*.

6. **Parekh, U.***, Wu, Y.*, Zhao, D., Worlikar, A., Shah, N., Zhang, K. & Mali, P. Mapping Cellular Reprogramming via Pooled Overexpression Screens with Paired Fitness and Single Cell RNA-Sequencing Readout, *Cell Systems*, 7 (5), 548-555.e8 (2018) **equal contribution*.
7. Hu, M., Dailamy, A., Lei, X. Y., **Parekh, U.**, McDonald, D., Kumar, A. & Mali, P. Facile Engineering of Long-Term Culturable Ex Vivo Vascularized Tissues Using Biologically Derived Matrices, *Advanced Healthcare Materials* 7 (23), 1800845 (2018).
8. Wang, K., **Parekh, U.**, Pailla, T., Garudadri, H., Gilja, V. & Ng, T.N. Stretchable Dry Electrodes with Concentric Ring Geometry for Enhancing Spatial Resolution in Electrophysiology. *Advanced Healthcare Materials* 6 (19), 1700552 (2017).
9. **Parekh, U.***, Yusupova, M.* & Mali, P. Genome engineering in human pluripotent stem cells. *Current Opinion Chemical Engineering.* 15, 56–67 (2017) **equal contribution*.

PATENTS

1. Mali, P., **Parekh, U.**, Dailamy, A., Lei, X., Hu, M. Use of 3D-Printed Freestanding Structures for Ex-Vivo Tissue *WO 2019226710A1* (2019).
2. Zhang, K., Wu, Y., Dailamy, A., Mali, P., McDonald, D., **Parekh, U.**, Hu, M. Novel Method to Engineer Translatable Human Tissue *WO 2020010249A1* (2020).

ABSTRACT OF THE DISSERTATION

Programmable systems to engineer human tissues and electrophysiological sensors

by

Udit Parekh

Doctor of Philosophy in Electrical Engineering (Medical Devices and Systems)

University of California San Diego, 2021

Professor Tse Nga Ng, Chair

Professor Prashant Mali, Co-Chair

To study human development and disease, and develop therapies for regenerative medicine, the ability to create scalable, physiologically relevant human tissues is critical. In this regard, engineering cells and tissues from pluripotent stem cells has led to significant advances in our understanding of human-specific biology and holds promise for therapies. However, several challenges remain, including the ability to differentiate diverse lineages and drive disease states,

vascularization to build tissues at scale, and monitoring capabilities to assess function. In this dissertation, I present a multi-faceted approach to address these limitations in creating and assaying complex tissue. To rapidly discover methods to differentiate pluripotent stem cells, I developed a screening method leveraging single cell RNA-sequencing to study the effects of gene overexpression. Using this approach, I assayed both fitness and transcriptomic responses of transcription factors, mutant gene libraries and whole gene families on pluripotent stem cells in multiple culture conditions. From these responses I built gene regulatory networks and found *ETV2* as a reprogramming factor toward endothelial cells. I further engineered the system in combination with teratoma formation to develop a multiplexed system to assay the potential of genes and variants to drive oncogenic transformation in a tissue-specific manner. I found that *c-MYC* alone or together with myristoylated *AKT1* drives transformation of neural progenitor lineages, while MEK1$^{S218D/S222D}$ drives proliferative advantage in mesenchymal lineages like fibroblasts. I then harnessed these reprogramming approaches to engineer densely vascularized human tissue. I combined reprogramming and chemically directed differentiation by overexpressing lineage-specifying transcription factors in differentiating vascular organoids to introduce neurons and skeletal muscle into the organoids, demonstrating maintenance of molecular and functional characteristics of the parenchymal and vascular lineages. Finally, I developed flexible, printed electrodes to enable the monitoring of electrophysiological signals and electrical perturbation of tissues. I enabled low noise, high spatial resolution measurement of clinically relevant signals using screen-printed, stretchable concentric ring electrodes. I then applied these screen-printed electrodes to study the effects of electrical stimulation on the wound healing response *in vivo*. Lastly, I demonstrated preliminary data on a novel fabrication method to print microelectrodes to map cellular electrical activity.

Chapter 1: Engineering complex human tissue

1.1 Introduction

The recapitulation of complex human tissue is an important and long-standing goal in bioengineering, both to study human development and disease, as well as to create therapies for regenerative medicine.

While animal models[1–4] have been invaluable in the modeling and understanding of development and disease, offering the ability to engineer, perturb and modulate specific cells and tissues in their native settings, these models do not completely elucidate human-specific biology[5,6]. On the other hand, two-dimensional cell culture of immortalized cell lines and primary cells has proven invaluable in studying human biology but here, cells lack the physiological context of the cellular niche[7,8]. Recently, the advent of multi-modal profiling of tissue samples from healthy, post-mortem and patient volunteers is proving to be a key tool in mapping and characterizing human organs in healthy and diseased states[9–16]. But unlike animal models, they do not provide a platform for experimental perturbation and controlled spatiotemporal profiling. Thus, faithful mimics of human organ systems paired with technologies which enable spatiotemporally resolved monitoring and assaying, are critical to gain a deeper understanding of human biology in health, development and disease.

Furthermore, there is an acute clinical need for transplantable organ-specific human cells and tissue[17–20]. Diverse clinical conditions require the transplantation or engraftment of autologous or donor cells and organs. Prominently, treatment for tissue and organ failure, is transplantation of donor cells or organs. However, several constraints limit the number of patients who can receive transplants namely a shortage of donors, limited ability to sustain donor organs *ex vivo*, and difficulty in matching donor and patients. In addition, cell engraftment or organ transplantation are promising therapies for a variety of clinical presentations, such as severe injury with volumetric muscle loss[21], degenerative conditions like Parkinson's disease[22–24], burns[21], ocular conditions

like age related macular degeneration[25,26], or chronic conditions like diabetes[27]. Engineered cell therapies have also gained traction as immunotherapeutics for several diseases including cancer[28], autoimmune conditions[29] and infections[30]. All of these therapies require the extraction, isolation or manufacturing of an array cells and tissues, and developing technologies to address this need as well as technologies for monitoring function and integrity is a major challenge.

In this dissertation I developed a series of technologies to enable the engineering of cells, disease states and vascularized tissue. In tandem I developed printed flexible electrodes at multiple scales to enable the monitoring of electrophysiological signals and electrical stimulation of tissue.

1.2 Pluripotent stem cells

Pluripotent stem cells (PSCs) have an unlimited capacity for proliferation *in vitro* and are capable of differentiating into all cell types in the body, offering an attractive route to providing cells for research and therapeutics.

Human PSCs were first isolated from the inner cell mass of the blastocyst of human embryos[31], and demonstrated the ability to differentiate into cells from all three germ layers via teratoma formation after injection into immunodeficient mice. While technically promising, these embryonic stem cells have had ethical concerns around their use, since deriving them requires the destruction of early-stage human embryos. However, the discovery and rapid adoption of induced pluripotent stem cells (iPSCs)[32,33], where adult somatic cells are converted to a pluripotent stem cell fate via the expression of defined genetic factors, has transformed the use of stem cells in all facets of research and for clinical translation[34].

The advent of iPSCs has opened the possibility of using an autologous cell source for therapy or creating large cell banking facilities[34] which can ease the matching of donor cells to patients. Current cell therapies and organ transplants are typically done with either autologous

cells isolated from the patient, or using matched donor cells or organs. The isolation and engineering of specific autologous cells is time-consuming, expensive and not always feasible, and it is difficult to find matching donors and maintain the viability of excised cells and organs. Thus, there is a pressing need to develop “off-the-shelf” engineered cell and tissue therapies.

In tandem, the development of efficient and versatile genome engineering technologies[35] has also allowed iPSCs to be used effectively for disease modeling, especially the study of variants by generating paired isogenic cell lines, which controls for the variability of differentiation between iPSC lines and controls for the genetic background besides the variant under study. This has led to advances in the study of monogenic diseases and is being applied in large scale screening studies on lineages differentiated from iPSCs[36,37], or to study factors affecting pluripotency and differentiation itself[38].

1.3 Differentiation of pluripotent stem cells

While there has been significant progress towards the development of several cell therapies, there is still a need for methods and technologies to obtain cells across the diversity of lineages present in the human body. Currently, there are three broad, common approaches to differentiate stem cells (**Figure 1.1**). Chemically directed differentiation[39] exposes PSCs to soluble chemical cues informed by knowledge of development, to drive differentiation to desired lineages. While powerful and commonly used, these methods are often lengthy, complex and often result in heterogenous populations with low yields of the desired phenotype. Engineering of the micro-environment is also used for PSC differentiation, where mechanical or chemical properties of the niche are designed to differentiate PSCs to specific lineages, again inspired by knowledge of the *in vivo* niche or developmental dynamics of the extra-cellular environment[40]. A third, and promising strategy, is reprogramming via expression of transcription factors (TFs).

The overexpression of TFs has long been known to drive cell fate conversion whether it is direct conversion between somatic cell lineages[41,42], the induction of pluripotent states from somatic cells[32,33], or for the differentiation of PSCs[43–47]. These methods are efficient, rapid and relatively facile compared to directed differentiation and micro-environment manipulation. The choice of TF to drive these conversions has been made using knowledge of their role in development in combination with systematic trial and error of a set of candidate TFs. This process can be laborious and does not translate easily to scalably testing combinations of TFs, since the search space becomes too large for arrayed approaches. A scalable screening approach would be beneficial to discover new reprogramming recipes.

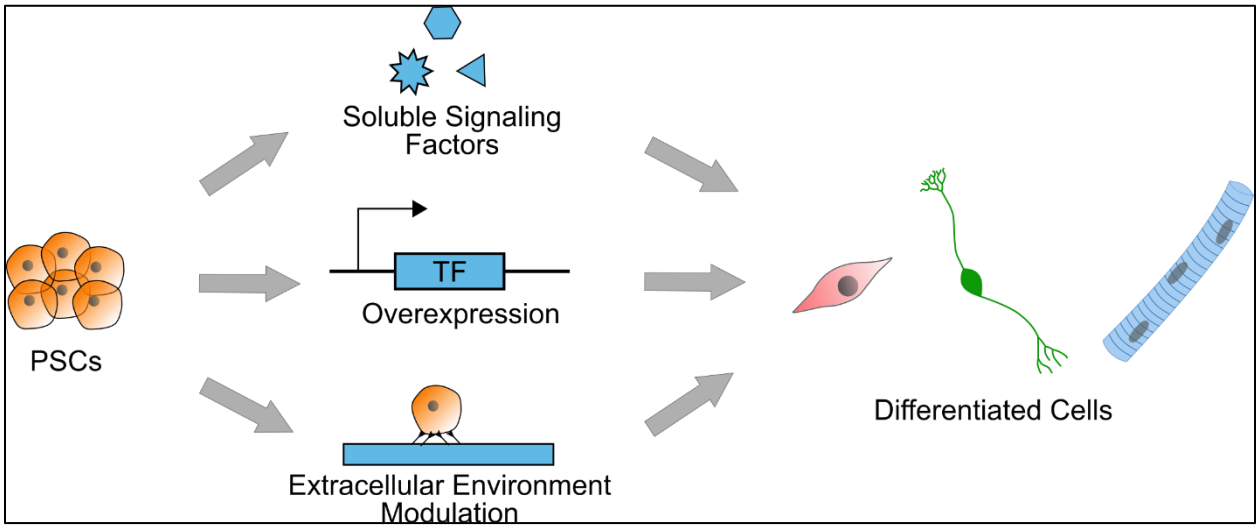


Figure 1.1: Schematic of common methods for pluripotent stem cell differentiation

1.4 Engineering tissue *in vitro*

To fully understand the intricacies of human biology and to eventually obtain organotypic tissue for clinical translation, the ability to engineer complex human tissue is crucial. Three-dimensional tissue in physiological faithful architectures can capture the various cell intrinsic and extrinsic interactions present *in vivo*, can model the development of organs, and functional *ex vivo* tissue may eventually be scaled for translation.

Tissue engineering combines the engineering of cells, material properties of the scaffold matrix, chemical or physiological cues and three-dimensional culture systems to fabricate complex multi-lineage tissues[40]. One promising approach, stem cell-derived organotypic ensembles of cells, termed organoids[48,49], has significantly advanced our ability to model and study the biology of a range of organs including the intestine[50], brain[51] and kidney[52]. These organoids arise by the self-assembly of stem cells which are then guided through the developmental process by temporally appropriate exposure to growth factors, small molecules and morphogens. Remarkably, they recapitulate the complex geometries of cellular organization, they can also recapitulate developmental trajectories and function, and are composed of a panoply of cell types present in the organ of interest, providing powerful platforms for the study of development and the onset of disease[48,53–59]. However, since organoids are formed by guiding the cells through a specific developmental process, they do not contain the entire range of cells which compose an organ since they do not form cross-germ layer cell types and critically many organoid models lack vasculature. Vascularization of organoids is important for two reasons – one, vascular cells play a central role in development and disease[60,61], and secondly scaling engineered tissue without necrosis requires dense vascularization[62,63].

Current methods of vascularizing engineered tissue rely on the angiogenic capacity of primary cells (**Figure 1.2.a**), biofabrication of vasculature, e.g. by bioprinting[64,65] (**Figure 1.2.b**), *in vivo* implantation[66–69] (**Figure 1.2.c**), and recently overexpression of vascular cell-inducing TFs[70] (**Figure 1.2.d**). While all of these are promising techniques, certain limitations must be overcome: Primary cells are a limited cell source which cannot be infinitely renewed; biofabrication methods have so far been unable to generate dense geometries like capillary networks; *in vivo* implantation methods generate chimeric tissue which cannot be used for translation or human-specific biology; and TF overexpression has so far been unable to generate the full set of vascular lineage cells. This paucity of methods for dense vascularization and the

generation of all vascular lineages is a critical bottleneck to advancing organoid and tissue engineered models.

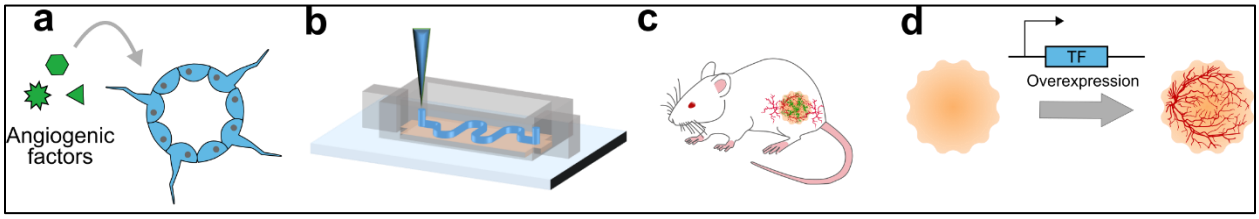


Figure 1.2: Overview of common methods for vascularization

(a) Angiogenic sprouting induced in vessels formed from primary cells. **(b)** Biofabrication with sacrificial materials to generate perfusable channels. **(c)** In vivo implantation for vascularization by host vasculature. **(d)** Overexpression driven differentiation of vascular cells in engineered tissue and organoids.

1.5 Engineering tissue *in vivo*

Organoid models have transformed our ability to model human-specific biology, development and disease but they still face certain challenges. Organoids produce typically immature phenotypes, are not vascularized and typically only contain a subset of lineages present in developing as well as mature organs. To overcome these issues and access difficult to differentiate cell types, one promising strategy is the development of tissues *in vivo* in a xenograft or chimeric format.

Two methods are broadly used for such *in vivo* tissue engineering. In the first, cells or tissues are implanted in an immunodeficient animal model, often in a suitable niche setting, to drive differentiation, maturation and vascularization[68,69]. In the second, animal models are engineered such that they lack the ability to form specific organs. PSCs are then introduced into the developing embryos of these models via blastocyst complementation, such that foreign cells will compensate and form the organ which the engineered cells are incapable of forming[71–73]. Such complementation strategies can be undertaken using PSCs from the same species or from

a different species, to generate inter-species chimeric organs, and have led to the formation of complex organs like the pancreas and kidney.

A third method that we[74] and others[75] have implemented takes advantage of the spontaneous differentiation of human PSCs into teratomas[76] when implanted in immunodeficient animal models. These are typically benign tumors that contain a variety of lineages derived from all three germ layers. This strategy has led to the derivation of difficult to differentiate cell types, like skeletal muscle stem cells[75] and hematopoietic stem cells[76] from teratomas. The PSCs which are implanted can be engineered for differentiation or enrichment of specific lineages, as well as to study development and disease. PSCs undergo a mimic of natural development while forming the teratoma, thus providing a promising platform for the systematic longitudinal study of human biology in health and disease, and the engineering of complex tissue.

1.6 Flexible electronics for functional monitoring and stimulation

While great strides have been made in engineering human tissue, functional monitoring and controlled perturbation of tissues is a significant challenge both *in vivo* and *ex vivo*. Devices and methods to functionally monitor as well as programmably perturb tissues in a spatiotemporally controlled manner are needed[77,78].

The measurement of electrophysiological signals is required at multiple biological scales. Organ-scale signals such as those from the cerebral cortex, heart and skeletal muscle are now commonly deployed in clinical practice. Cellular scale measurements[79] both *in vivo* and *in vitro* are used to study the function and connectivity of electrically active cells and tissue, and for therapy. Additionally, the advent and prevalence of wearable electronic devices[80,81] to assess bodily signals in health and disease has only increased the need for bioelectronic measurement devices. To monitor bioelectronic signals from three-dimensional, soft, deformable and non-uniformly shaped biologically relevant surfaces, flexible and printed electronics are increasingly

being developed[80–85]. Further, electrical stimulation is advancing as a therapy for several conditions, including injury rehabilitation[86], neurological conditions like epilepsy and Parkinson's disease[87,88] and wound healing[89], besides its application in controlling and stimulating prostheses and muscle. Thus electrode devices not only need to be developed for their performance in detection, but for stimulation capabilities as well.

Traditional rigid electrodes for electrophysiology employ a conductive gel to serve as an interface between the electrode and skin. These gel electrodes are liable to motion artifacts and reduced performance over time as the gel dries. On the other hand, flexible electronics can provide superior performance. Conformal electronic devices reduce noise by more intimate contact and lower contact impedance, and enable long-term, continuous recordings with higher spatial resolution by reducing rigid elements and eliminating the need for gels. Flexible electrodes can also be placed at higher density by elimination of the need for gel usage and by accommodation into uneven geometries, thus leading to recordings with higher spatial resolution.

Flexible electronic devices can also address a key remaining challenge in tissue engineering and organoids – that of monitoring and perturbing these engineered tissues in a spatiotemporally controlled manner, as well as potentially after transplantation. Recent demonstrations have shown the implementation of such devices, including porous mesh metal electrode arrays[90], intrinsically stretchable conductive polymers arrays[91,92], and nanowire-based field effect transistor arrays[93]. It has also been shown that devices can be designed to electrically stimulate the tissue, as well as perform controlled chemical release, both of which could be useful for tissue maturation, as has been shown to be the case for cardiac tissue[94].

1.7 Printed electronics

While standard microfabrication methods can be used to produce these devices, printed electronics[82,84] offers further flexibility by providing the ability to rapidly reconfigure geometries,

dimensions and substrates to account for different tissues, locations and individuals. Furthermore, printing techniques have been developed to be compatible with a range of flexible substrates and elastomers, as well as with a range of inorganic and organic materials. There are wide array of methods used in printing functional electronic devices, in this dissertation, I primarily focus on screen printing, inkjet printing and electrohydrodynamic printing.

Screen printing (**Figure 1.3.a**) is a popular, well-established and mature technique with ease of use, adaptability to manufacturing and low cost. In this method, a stencil with the desired pattern cut out is placed on the substrate, and the ink is then pushed through the stencil and deposited on the substrate. Sintering is then required to obtain conductive patterns. This method is suitable for patterns with larger critical dimensions and large scale manufacturing. While reconfigurable, changes in design cannot be immediately implemented, but must first be translated into a stencil.

To fabricate patterns with smaller critical dimensions down to 20-30 μm, and for rapid prototyping and reconfigurability, inkjet printing (**Figure 1.3.b**) is a common method. Here, patterns are fabricated via the deposition of ink through a micron-scale nozzle, with droplet expulsion controlled by a piezoelectric actuator. The printhead motion is programmatically controlled to directly define printed geometries.

Lastly, to print even smaller geometries than those obtainable with inkjet printing, electrohydrodynamic printing (**Figure 1.3.c**) is a promising technique. Here, ink is filled under pressure into a tip with an inner diameter of <10 μm, and an electric field pulse applied between the tip and a conducting substrate to jet droplets of the ink.

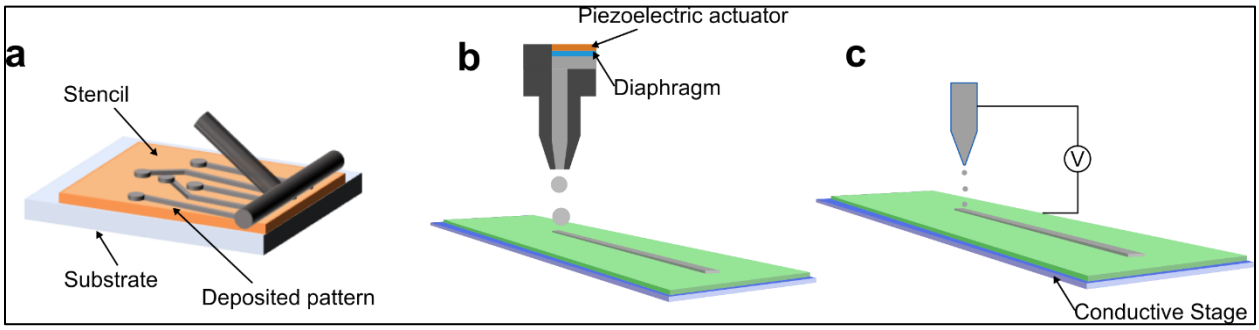


Figure 1.3: Schematic of printing methods

(a) Schematic of screen printing. **(b)** Schematic of inkjet printing. **(c)** Schematic of electrohydrodynamic printing.

1.8 Conducting polymers

The discovery of conducting polymers[95] has introduced a versatile toolbox for organic electronics and recently, the use of conducting polymers has become widespread in devices measuring and stimulating biological samples[96]. Conducting polymers contain sp^2 hybridized carbons in their backbone, which leaves a valence electron in an unhybridized p-orbital. These p-orbitals combine to form a delocalized π bond across the length of the polymer backbone, enabling electronic conduction across the polymer molecule (**Figure 1.4.a**)[97]. To modulate the conductivity of the polymer, it can also be doped in a redox manner by adding dopant molecules to add or remove electrons from the delocalized π bond.

Conducting polymers are attractive materials for biological applications since they have a stiffness comparable to the contacted tissue and importantly, conducting polymers are mixed ionic-electronic conductors which reduces the skin-electrode interface impedance. Additionally, conducting polymers have a high ionic charge injection capacity, which also makes them suitable materials for electrical stimulation. Furthermore, these polymers can also be dispersed and solution processed in a facile manner as compared to fabrication techniques for inorganic materials.

Poly(3,4-ethylenedioxythiophene) (PEDOT)[96] is a conducting polymer with one of the highest known conductivities and is now widely used in organic electronics. In the presence of polystyrene sulfonate (PSS), PEDOT:PSS (**Figure 1.4.b**)[98] can be dispersed in aqueous solution for facile processing, and then cross-linked after deposition to be stable in aqueous environments. In addition, the conductivity of PEDOT:PSS can be further increased by the addition of organic solvents and ionic liquids. Organic solvents like ethylene glycol can both, dope PEDOT as well as change its structural properties to increase conductivity. Ionic liquids on the other hand can screen the interaction between PEDOT and the PSS side chain, allowing PEDOT to form crystalline domains, as well as make the resulting polymer intrinsically stretchable. Finally, PEDOT conductivity can be significantly increased by treatment with sulfuric acid, which removes the insulating PSS side chains and allows PEDOT to form crystalline domains.

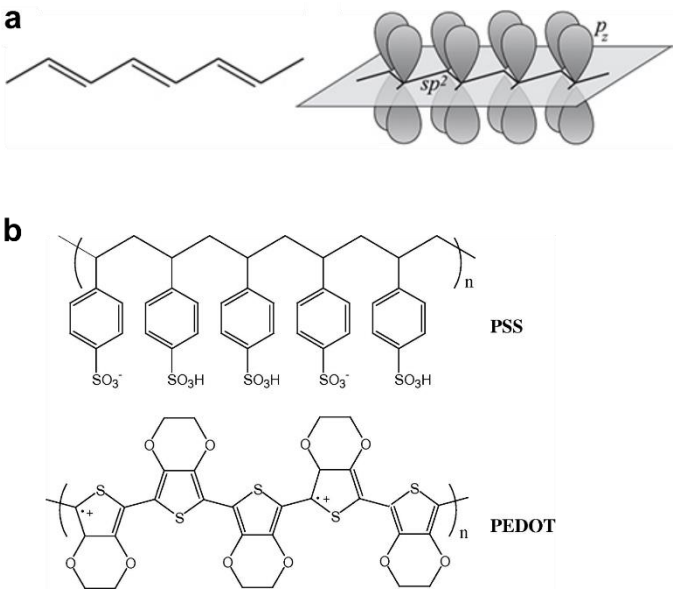


Figure 1.4: Structure of conducting polymers

(a) Example structure of a conductive polymer, polyacetylene (left), with schematic indicating unhybridized p-orbitals along the backbone (right). **(b)** Structure of PEDOT:PSS.

Chapter 2: Development of a single cell RNA-sequencing based pooled overexpression screening method to study reprogramming

2.1 Abstract

Understanding the complex effects of genetic perturbations on cellular state and fitness in has been challenging using traditional pooled screening techniques, which typically rely on the delivery of a single perturbation per cell and unidimensional phenotypic readouts. Here, we use barcoded open reading frame (ORF) overexpression libraries coupled with single-cell RNA sequencing (scRNA-seq) to assay cell state and fitness, a technique we call SEUSS (**S**calabl**E** f**U**nctional **S**creening by **S**equencing) (**Figure 2.1**). Using this system, we perturbed hPSCs with a library of developmentally critical transcription factors (TFs) and assayed the impact of TF overexpression on fitness and transcriptomic states across multiple conditions. We further leveraged the versatility of the ORF library approach to systematically assay mutant gene libraries and whole gene families. From the transcriptomic responses, we built genetic co-regulatory networks to identify key altered gene modules and we found that *KLF4* and *SNAI2* drive opposing effects along the epithelial-mesenchymal transition axis. From the fitness response, we identified *ETV2* as a driver of reprogramming towards an endothelial-like state, highlighting the power of our method to characterize the effects of genetic perturbations.

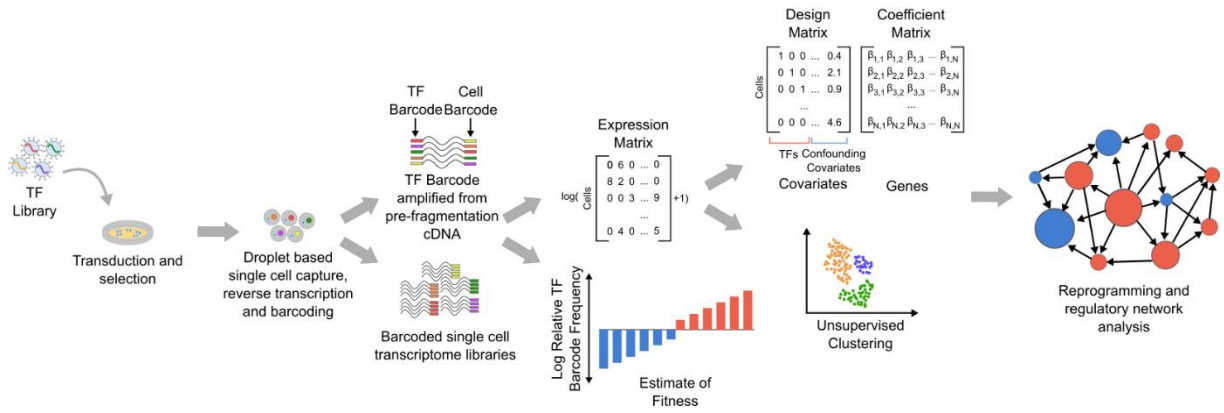


Figure 2.1: Schematic of SEUSS workflow

Experimental and analytical framework for evaluation of effects of transcription factor (TF) overexpression in hPSCs: Individual TFs are cloned into the barcoded ORF overexpression vector, pooled and packaged into lentiviral libraries for transduction of hPSCs. Transduced cells are harvested at a fixed time point to be assayed as single cells using droplet based scRNA-seq to evaluate transcriptomic changes. Cells are genotyped by amplifying the overexpression transcript from scRNA-seq cDNA prior to fragmentation and library construction and identifying the overexpressed TF barcode for each cell. The cell count for each genotype is used to estimate fitness. Gene expression matrices from scRNA-seq are used to obtain differential gene expression and clustering signatures, which in turn are used for evaluation of cell state reprogramming and gene regulatory network analysis.

2.2 Introduction

Cellular reprogramming via the overexpression of transcription factors (TF), has widely impacted biological research, from the direct conversion of adult somatic cells[41,42] and the induction of pluripotent stem cells[32,33,99–102], to the differentiation of hPSCs[43–47,103]. The discovery of TFs that drive reprogramming has previously involved both prior knowledge of their role in development and cellular transformation, and systematic trial-and-error. A scalable screening method to assess the effects of TF overexpression would advance fundamental understanding of reprogramming and enable the rapid discovery of novel reprogramming factors.

Recently, screens combining genetic perturbations with scRNA-seq[104] readouts have emerged as promising alternatives to traditional screens[105,106], enabling high-throughput, high-content screening by simultaneously profiling the transcriptomic response of tens of thousands of individual cells to genetic perturbations. These scRNA-seq screens are scalable and enable direct

readout of transcriptomic changes, providing a powerful tool in unraveling transcriptional networks and cascades. While other groups have demonstrated CRISPR-Cas9 based knock-out and knock-down scRNA-seq screens[107–111], to our knowledge, scRNA-seq based gene overexpression screens have yet to be demonstrated.

Here, we use barcoded open-reading frame (ORF) overexpression libraries with a coupled scRNA-seq and fitness screen, a technique we call SEUSS, to systematically overexpress a pooled library of TFs and assay both the transcriptomic and fitness effects on hPSCs. While CRISPRa offers some advantages, including easier scale-up, and the ability to mimic endogenous activation[112,113], we chose ORF constructs for several reasons. ORF overexpression yields strong, stable expression of the gene of interest, and enables the expression of specific isoforms as well as engineered or mutant forms of genes, aspects not accessible through endogenous activation.

We harnessed the SEUSS approach to assay the effects of TF overexpression on the pluripotent cell state and to find reprogramming factors.

2.3 Materials and methods

2.3.1 Cell Culture

H1 male hESC cell line was 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).

HEK 293T cells were maintained in high glucose DMEM supplemented with 10% fetal bovine serum (FBS).

HUVECs were maintained in endothelial growth medium (EGM-2, Lonza) and were not used beyond passage 10.

2.3.2 Library Preparation

The lentiviral backbone plasmid for the TF-Hygro vector format was constructed containing the EF1α promoter, mCherry transgene flanked by BamHI restriction sites, followed by a P2A peptide and hygromycin resistance enzyme gene immediately downstream. Each transcription factor in the library was individually inserted in place of the mCherry transgene. Since the ectopically expressed transcription factor would lack a poly-adenylation tail due to the presence of the 2A peptide immediately downstream of it, the transcript will not be captured during single-cell transcriptome sequencing which relies on binding the poly-adenylation tail of mRNA. Thus, a barcode sequence was introduced to allow for identification of the ectopically expressed transcription factor. The backbone was digested with HpaI, and a pool of 20 bp long barcodes with flanking sequences compatible with the HpaI site, was inserted immediately downstream of the hygromycin resistance gene by Gibson assembly. The vector was constructed such that the barcodes were located only 200 bp upstream of the 3'-LTR region. This design enabled the barcodes to be transcribed near the poly-adenylation tail of the transcripts and a high fraction of barcodes to be captured during sample processing for scRNA-seq.

To create the transcription factor library, individual transcription factors were PCR amplified out of a human cDNA pool (Promega Corporation) or obtained as synthesized double-stranded DNA fragments (gBlocks, IDT Inc) with flanking sequences compatible with the BamHI restriction sites. *MYC* mutants were obtained as gBlocks with a 6-amino acid GSGSGS linker substituted in place of deleted domains (**Table S1**). The lentiviral backbone was digested with BamHI HF (New England Biolabs) at 37 ºC for 3 hours in a reaction consisting of: lentiviral backbone, 4 µg, CutSmart buffer, 5 µl, BamHI, 0.625 µl, H_20 up to 50 µl. After digestion, the vector

was purified using a QIAquick PCR Purification Kit (Qiagen). Each transcription factor vector was then individually assembled via Gibson assembly. The Gibson assembly reactions were set up as follows: 100 ng digested lentiviral backbone, 3:10 molar ratio of transcription factor 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 and to extract barcode sequences. One overexpression vector was created for each TF, thus a single unique barcode was associated with each TF.

To assemble the library, individual transcription factor vectors were pooled together in an equal mass ratio along with a control vector containing the mCherry transgene which constituted 10% of the final pool.

To create the neural transcription factor library without a hygromycin resistance transgene, in the TF-NoHygro format, individual transcription factor coding sequences were PCR amplified from plasmids containing them, with flanking sequences compatible with the upstream BamHI restriction site and downstream HpaI restriction site. 20 bp barcode sequences were added after the transcription factor stop codon via the primers compatible with the HpaI restriction site. Swapping the locations of the ORF and selection marker was avoided so that residues from 2A peptide cleavage were not added to the N-terminal of the overexpressed TF. The lentiviral backbone was digested with BamHI HF and HpaI (New England Biolabs) at 37 ºC for 3 hours in a reaction consisting of: lentiviral backbone, 3 µg, CutSmart buffer, 5 µl, BamHI, 0.625 µl, HpaI, 2 µl, H_20 up to 50 µl. After digestion, the vector was purified using a QIAquick PCR Purification Kit (Qiagen). Each transcription factor vector was then individually assembled via Gibson

assembly. The Gibson assembly reactions were set up as follows: 100 ng digested lentiviral backbone, 3:10 molar ratio of transcription factor 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 and to extract barcode sequences. One overexpression vector was created for each TF, thus a single unique barcode was associated with each TF.

To assemble the library, individual transcription factor vectors were pooled together in an equal mass ratio along with a control vector containing the mCherry transgene which constituted 10% of the final pool.

2.3.3 Viral Production

HEK 293T cells were maintained in high glucose DMEM supplemented with 10% fetal bovine serum (FBS). In order to produce lentivirus particles, cells were seeded in a 15 cm dish 1 day prior to transfection, such that they were 60-70% confluent at the time of transfection. For each 15 cm dish 36 µl of Lipofectamine 2000 (Life Technologies) was added to 1.5 ml of Opti-MEM (Life Technologies). Separately 3 µg of pMD2.G (Addgene #12259), 12 µg of pCMV delta R8.2 (Addgene #12263) and 9 µg of an individual vector or pooled vector library was added to 1.5 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 15 cm dish was replaced with 25 ml of fresh, pre-warmed medium. After the incubation period, the mixture was added dropwise to each dish 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) to a final volume of 600-800 µl, divided into aliquots and frozen at -80 ºC.

For high MOI transduction of the neural transcription factor library, the lentivirus was further concentrated using Amicon Ultra-0.5 centrifugal ultrafilters with a 100,000 NMWL cutoff (Millipore) such that the final volume used for transduction was less than 20% of the total volume per well of a 6 well plate.

For analysis of barcode shuffling, lentivirus particles were produced by seeding HEK 293T cells in 6-well plates such that they were 60-70% confluent at the time of transfection. For each well, 7.5 µl of Lipofectamine 2000 was added to 125 ul of Opti-MEM. Separately, 625 ng of pMD2.G, 2.5 µg of pCMV delta R8.2 and 1.875 µg of an individual or pooled neural transcription factor library was added to 125 µl 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 well was replaced with 2 ml of fresh, pre-warmed medium. After the incubation period, the mixture was added dropwise to each well of HEK 293T cells. For the Arrayed Neural TF library, viral particles of each individual transcription factor were produced in separate wells of 6-well plates. For the Pooled Neural TF library, the plasmids were pooled along with a mCherry vector constituting 10% of the final pool and lentiviral particles produced in an equal number of wells as the Unpooled production. Supernatant containing the viral particles was harvested after 48 and 72 hours, pooled and filtered with 0.45 µm filters (Steriflip, Millipore), and further concentrated using Amicon Ultra-15 centrifugal ultrafilters with a 100,000 NMWL cutoff (Millipore) to a final volume of 600-800 µl, divided into aliquots and frozen at -80 ºC. The lentivirus was further concentrated using Amicon Ultra-0.5

centrifugal ultrafilters with a 100,000 NMWL cutoff (Millipore) such that the final volume used for transduction was less than 20% of the total volume per well of a 6 well plate.

2.3.4 Viral Transduction

For viral transduction, on day -1, H1 cells were dissociated to a single cell suspension using Accutase (Innovative Cell Technologies) and seeded into Matrigel-coated plates in mTeSR containing ROCK inhibitor, Y-27632 (10 μM, Sigma-Aldrich). For transduction with the TF library, cells were seeded into 10 cm dishes at a density of $6x10^6$ cells for screens conducted in mTeSR or $4.5x10^6$ cells for screens conducted in endothelial growth medium (EGM) or multilineage (ML) medium (DMEM + 20% FBS.) For transduction with the neural TF library, *KLF* gene family library and *c-MYC* mutants library, cells were seeded at a density of $1x10^6$ cells per well of a 6-well plate. For transduction with the neural TF library not containing a hygromycin resistance transgene, cells were seeded at a density of $0.5x10^6$ cells per well of a 6-well plate. For transduction with the neural TF library, with or without the hygromycin resistance transgene, multiple wells were transduced at varying titers, along with companion wells transduced with equal titer of control mCherry virus only, to control for effects of viral toxicity and high serum content. The highest transduction titer for which no toxicity or differentiation was seen in the control wells was used for scRNA-seq experiments. For transduction with individual transcription factors cells were seeded at a density of $4x10^5$ cells per well of a 12 well plate or $2x10^5$ cells per well of a 24 well plate for experiments conducted in mTeSR, while for experiments conducted in the alternate media cells were seeded at a density of $3x10^5$ cells per well of a 12 well plate or $1.5x10^5$ cells per well of a 24 well plate.

On day 0, medium was replaced with fresh mTeSR to allow cells to recover for 6-8 hours. Recovered cells were then transduced with lentivirus added to fresh mTeSR containing polybrene (5 μg/ml, Millipore). On day 1, medium was replaced with the appropriate fresh medium: mTeSR,

endothelial growth medium (EGM-2, Lonza) or high glucose DMEM + 20% FBS. For all vectors and libraries containing a hygromycin resistance transgene, hygromycin (Thermo Fisher Scientific) selection was started from day 2 onward at a selection dose of 50 µg/ml, medium containing hygromycin was replaced daily.

2.3.5 Single Cell Library Preparation

For screens conducted in mTeSR cells were harvested 5 days after transduction while for alternate media, EGM or ML, cells were harvested 6 days after transduction with the TF library. Cells were dissociated to single cell suspensions using Accutase (Innovative Cell Technologies). For samples sorted with magnetically assisted cell sorting (MACS), cells were labelled with anti-TRA-1-60 antibodies or with dead cell removal microbeads and sorted as per manufacturer's instructions (Miltenyi Biotec). Samples were then resuspended in 1XPBS with 0.04% BSA at a concentration between 600-2000 per µl. Samples were loaded on the 10X Chromium system and processed as per manufacturer's instructions (10X Genomics). Unused cells were centrifuged at 300 rcf for 5 minutes and stored as pellets at -80 ºC until extraction of genomic DNA.

Single cell libraries were prepared as per the manufacturer's instructions using the Single Cell 3' Reagent Kit v2 (10X Genomics). Prior to fragmentation, a fraction of the sample post-cDNA amplification was used to amplify the transcripts containing both the TF barcode and cell barcode. Single cell RNA-seq libraries and barcode amplicons were sequenced on an Illumina HiSeq platform.

2.3.6 Quantification of Barcode Shuffling

To quantify the extent of barcode shuffling, lentivirus particles of the Neural TF library were produced in an arrayed or pooled manner in 6-well plates as described previously. Multiple wells were transduced at varying titers, the highest transduction titer for which no toxicity or differentiation was seen in the control wells was used for downstream processing. Hygromycin

(Thermo Fisher Scientific) selection was started from day 2 onward at a selection dose of 50 µg/ml, medium containing hygromycin was replaced daily. Cells were harvested 5 days after transduction, spun at 300 rcf for 5 min to obtain cell pellets and genomic DNA extracted.

2.3.7 Barcode Amplification

Barcodes were amplified from cDNA generated by the single cell system as well as from genomic DNA from cells not used for single cell sequencing. Barcodes were amplified from both types of samples and prepared for deep sequencing through a two-step PCR process.

For amplification of barcodes from cDNA, the first step was performed as three separate 50 µl reactions for each sample. 2 µl of the cDNA was input per reaction with Kapa Hifi Hotstart ReadyMix (Kapa Biosystems). The PCR primers used were, Nexterai7_TF_Barcode_F: GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGAGAACTATTTCCTGGCTGTTACGCG and NEBNext Universal PCR Primer for Illumina (New England Biolabs). The thermocycling parameters were 95 °C for 3 min; 24-26 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 (~500 bp) of 3 reactions for each sample were pooled, size-selected and purified with Agencourt AMPure XP beads 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. Nextera XT 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 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 three separate 50 µl reactions for each sample. 2 µg of genomic DNA was input per reaction with Kapa Hifi Hotstart ReadyMix. The PCR primers used were, NGS_TF-Barcode_F: ACACTCTTTCCCTACACGACGCTCTTCCGATCTAGAACTATTTCCTGGCTGTTACGCG and NGS_TF-Barcode_R: GACTGGAGTTCAGACGTGTGCTCTTCCGATCTTGTCTTCGTTGGGAGTGAATTAGC. The thermocycling parameters were: 95 °C for 3 min; 26-28 cycles of (98 °C for 20 s; 55 °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 (200 bp) of 3 reactions for each sample were pooled, size-selected with Agencourt AMPure XP beads (Beckman Coulter, Inc.) at a ratio of 0.8, and the supernatant from this was further size-selected and purified at a ratio of 1.6. The second step of PCR was performed as two separate 50 µl reactions with 50 ng of first step purified PCR product per reaction. Next Multiplex Oligos for Illumina (New England Biolabs) 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 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 0.8, and the supernatant from this was further size-selected and purified 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 MiSeq platform.

For amplification of barcodes from genomic DNA for barcode shuffling analysis, genomic DNA was extracted from stored cell pellets with a DNeasy Blood and Tissue Kit (Qiagen). The barcode for the mCherry control only was amplified for next generation sequencing in three steps. The first step PCR was performed as three separate 50 µl reactions for each sample. 2 µg of

genomic DNA was input per reaction with Kapa Hifi Hotstart ReadyMix. The PCR primers used were, mCh_BC_Shuffling_F: CACCATCGTGGAACAGTACGAAC and TF_BC_Shuffling_R: GACTGGAGTTCAGACGTGTGCTCTTCCGATCTCACTGTTTAACAAGCCCGTCAGTAG. The thermocycling parameters were: 95 °C for 3 min; 24-26 cycles of (98 °C for 20 s; 65 °C for 15 s; and 72 °C for 90 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 of 3 reactions for each sample were pooled, and dimers removed by size-selecting with Agencourt AMPure XP beads (Beckman Coulter, Inc.) at a ratio of 1.6. The second step of PCR was performed as two separate 50 µl reactions with 50 ng of first step purified PCR product per reaction with Kapa Hifi Hotstart ReadyMix. The PCR primers used were, TF_BC_Shuffling_Step2_F: ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGGTTGACGGCAATTTCGATG and TF_BC_Shuffling_R: GACTGGAGTTCAGACGTGTGCTCTTCCGATCTCACTGTTTAACAAGCCCGTCAGTAG. The thermocycling parameters were: 95 °C for 3 min; 6-8 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. The amplicons from these two reactions for each sample were pooled, size-selected with Agencourt AMPure XP beads at a ratio of 0.8, and the supernatant from this was further size-selected and purified at a ratio of 1.6. The third step of PCR was performed as two separate 50 µl reactions with 50 ng of second step purified PCR product per reaction with Kapa Hifi Hotstart ReadyMix. Next Multiplex Oligos for Illumina (New England Biolabs) 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 0.8, and the supernatant from this was further size-selected and purified 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 HiSeq platform.

For amplification of barcodes from genomic DNA from cells transduced with the neural TF library in the TF-NoHygro format, genomic DNA was extracted from stored cell pellets with a DNeasy Blood and Tissue Kit (Qiagen). The barcode for the mCherry control only was amplified for next generation sequencing in two steps. The first step PCR was performed as three separate 50 µl reactions for each sample. 2 µg of genomic DNA was input per reaction with Kapa Hifi Hotstart ReadyMix. The PCR primers used were, No-Hygro_gDNA_mCh_Barcode_F: ACACTCTTTCCCTACACGACGCTCTTCCGATCCACCATCGTGGAACAGTACGAAC and No-Hygro_gDNA_Barcode_R: GACTGGAGTTCAGACGTGTGCTCTTCCGATCTTTCGATGCATGGGGTCGTGC. The thermocycling parameters were: 95 °C for 3 min; 28-30 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 of 3 reactions for each sample were pooled, size-selected with Agencourt AMPure XP beads at a ratio of 0.8, and the supernatant from this was further size-selected and purified at a ratio of 1.6. The second step of PCR was performed as two separate 50 µl reactions with 50 ng of first step purified PCR product per reaction. Next Multiplex Oligos for Illumina (New England Biolabs) 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 0.8, and the supernatant from this was further size-selected and purified 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 MiSeq platform.

2.3.8 Single cell RNA-seq Processing and Genotype Deconvolution

Using the 10X genomics CellRanger pipeline, we aligned Fastq files to hg38, counted UMIs to generate counts matrices, and aggregated samples across 10X runs with cellranger aggr. All cellranger commands were run using default settings.

To assign one or more transcription factor genotypes to each cell, we aligned the plasmid barcode reads to hg38 using BWA, and then labeled each read with its corresponding cell and UMI tags. To remove potential chimeric reads, we used a two-step filtering process. First, we only kept UMIs that made up at least 0.5% of the total amount of reads for each cell. We then counted the number of UMIs and reads for each plasmid barcode within each cell, and only assigned that cell any barcode that contained at least 10% of the cell's read and UMI counts. Barcodes were mapped to transcription factors within one edit distance of the expected barcode.

To quantify barcode shuffling, we simply extracted the plasmid barcode from each read and counted the number of reads corresponding to each genotype.

2.3.9 Clustering and Cluster Enrichment

Clustering was performed on the aggregated counts matrices using the Seurat pipeline[114]. We first filtered the counts matrix for genes that are expressed in at least 1% of cells, and cells that express at least 200 genes. We then normalized the counts matrix, found overdispersed genes, and used a negative binomial linear model to regress away library depth, batch effects, and mitochondrial gene fraction. We performed PCA on the overdispersed genes, keeping the first 20 principal components. We then used the PCs to generate a K Nearest Neighbors graph, with K = 30, used the KNN graph to calculate a shared nearest neighbors graph, and used a modularity optimization algorithm on the SNN graph to find clusters. Clusters were recursively merged until all clusters could be distinguished from every other cluster with an out of the box

error (oobe) of less than 5% using a random forest classifier trained on the top 15 genes by loading magnitude for the first 20 PCs. We used tSNE on the first 30 PCs to visualize the results.

Cluster enrichment was performed using Fisher's exact test, testing each genotype for both over-enrichment and under-enrichment in each cluster.

2.3.10 Differential Expression and Identification of Significant Genotypes

We used a modified version of the MIMOSCA linear model[108] to analyze the differentially expressed genes for each genotype. In our model, we used the R glmnet package with the multigaussian family, with alpha (the lasso vs ridge parameter) set to 0.5. Lambda (the coefficient magnitude regularization parameter) was set using 5-fold cross validation. We also used mCherry as a control genotype, computing gene expression changes for each genotype against the mCherry control. Our method outputs a genes by genotypes matrix of regression coefficients, where each coefficient corresponds to the effect of each genotype on each gene relative to the mCherry control. P-values were calculated empirically by randomly permuting the genotype assignments, and then false discovery rates were calculated using the Benjamani-Hochberg procedure.

TFs were chosen as significant for downstream analysis if they were enriched for at least one cluster with a p-value of less than 10^{-12}, or if the TF drove statistically significant differential expression of greater than 50 genes. Our threshold for calling a differentially expressed gene is that the false discovery rate was less than 0.05, and the absolute coefficient magnitude was greater than 0.025.

2.3.11 Gene Co-perturbation Network and Module Detection

We took the genes by genotypes coefficients matrix from the regression analysis with trimmed genotypes and used it to calculate the Euclidean distance between genes, using the significant genotypes as features. We then built a k-nearest neighbors graph from the Euclidean

distances between genes, with k = 30. From this kNN graph, we calculated the fraction of shared nearest neighbors (SNN) for each pair of genes to build and SNN graph. For example, if two genes share 23/30 neighbors, we create an edge between them in the SNN graph with a weight of 23/30 = 0.767.

To identify gene modules, we used the Louvain modularity optimization algorithm[115]. For each gene module, we identified enriched Gene Ontology terms using Fisher's exact test (**Table S4**). We also ranked genes in each gene module by the number of enriched Gene Ontology terms the gene is part of, to identify the most biologically significant genes in each module (**Table S4**). Gene module identities were assigned based on manual inspection of enriched GO terms and the genes within each module. The effect of each genotype on a gene module was calculated by taking the average of the regression coefficients for the genotype and the genes within the module. Gene modules where no genotype had an average absolute coefficient of at least 0.05 were dropped from further analysis in order to exclude gene modules that did not show variation across our dataset.

2.3.12 Replicate Correlation

For each of the medium conditions, we had two replicate screens. To establish the reproducibility of our screens, we correlated the regression coefficients of the replicates, where each coefficient represents the effect of a single TF on a single gene using a Pearson correlation. Because the vast majority of coefficients were either zero or very close to zero, we only correlated coefficients that were nonzero with an false discovery rate of less than 0.5 in *at least one* replicate (not both). This essentially filters out the coefficients that are zero or close to zero in both replicates.

To compare the results of our screen vs the bulk microarray overexpression screen, we used GSEA to assess the enrichment in the TF-gene effects (in the form of regression

coefficients) with the downstream targets for that same TF as determined by the bulk microarray screen.

2.3.13 Fitness Effect Analysis

To calculate fitness effects from genomic DNA reads, we first used MagECK[116] to align reads to genotype barcodes and count the number of reads for each genotype in each sample, resulting in a genotypes by samples read counts matrix. We normalized the read counts matrix by dividing each column by the sum of that column, and then calculated log fold-change by dividing each sample by the normalized plasmid library counts, and then taking a log2 transform. For the stem cell media, we averaged the log fold change across the non MACS sorted samples.

To calculate fitness effects from genotype counts identified from single cell RNA-seq, we used a cell counts matrix instead of a read counts matrix, and repeated the above protocol. To correlate the fitness replicates, used a Pearson correlation of the log fold-changes.

2.3.14 Epithelial mesenchymal transition analysis

We took genes from the Hallmark Epithelial Mesenchymal Transition geneset from MSigDB[117] and ran PCA on those genes with the stem cell medium dataset, visualizing the first two principal components. The two principal components resulted in an EMT-like signature, and we used the gene loadings from those principal components, along with literature research to identify a relevant panel of EMT related genes to display.

2.3.15 RNA Extraction, qRT-PCR and bulk RNA-seq library preparation

RNA was extracted from cells using the 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 1:5 with nuclease-free water.

qRT-PCR reactions were setup as: 2 µl cDNA, 400 nM of each primer, 2X Kapa SYBR Fast Master Mix (Kapa Biosystems), H_2O 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. Primers used for qRT-PCR are listed in **Table S5**. For confirmation of overexpression by qRT-PCR, primers were chosen such that they amplified a portion of the transcript in the hygromycin resistance region. This was done to avoid amplification of any endogenous transcripts, and since the overexpression is driven by a single promoter the TF, P2A peptide and the hygromycin resistance are on a single transcript.

Bulk RNA-seq libraries were prepared from 150 ng of RNA using the NEBNext Ultra RNA Library Prep kit for Illumina (New England Bioloabs) as per the manufacturer's instructions. Libraries were sequenced on an Illumina HiSeq platform.

2.3.16 Bulk RNA-Seq Analysis and Correlation

We mapped the bulk RNA-Seq fastq files to GRCh38 and quantified read counts mapping to each gene's exon using Gencode v28 and STAR aligner[118]. We used total counts normalization to adjust for library size effects, and then took a log-transform to adjust for heteroscedasticity. To quantify the effect of each TF versus mCherry, we took the log fold-change (logFC) of each TF's normalized expression versus the mCherry normalized expression. We compared this bulk logFC to the single cell RNA-seq regression coefficients using Pearson correlation.

2.3.17 Immunofluorescence

Cells were fixed with 4% (wt/vol) paraformaldehyde in PBS at room temperature for 30 minutes. Cells were then incubated with a blocking buffer: 5% donkey serum, 0.2% Triton X-100 in PBS for 1 hour at room temperature followed by incubation with primary antibodies diluted in

the blocking buffer at 4 °C overnight. Primary antibodies used were: VE-Cadherin (D87F2, RRID: AB_2077969, Cell Signaling Technology; 1:400), EPCAM (VU1D9, RRID: AB_558797, Thermo Fisher, 1:200), Vimentin (AF2105, RRID: AB_355153, R&D Systms, 1:50). Secondary antibodies used were: DyLight 488 labelled donkey anti-rabbit IgG (ab96891, Abcam; 1:250), DyLight 488 labelled donkey anti-goat IgG (ab96931, Abcam, 1:250), AlexaFluor 488 labelled goat anti-mouse IgG (A-11001, Thermo Fisher, 1:500).

After overnight incubation with primary antibodies, cells were labelled with secondary antibodies diluted in 1% BSA in PBS for 1 hour at 37 °C. Nuclear staining was done by incubating cells with DAPI for 5 minutes at room temperature. All imaging was conducted on a Leica DMi8 inverted microscope equipped with an Andor Zyla sCMOS camera and a Lumencor Spectra X multi-wavelength fluorescence light source.

2.3.18 Endothelial Tube Formation Assay

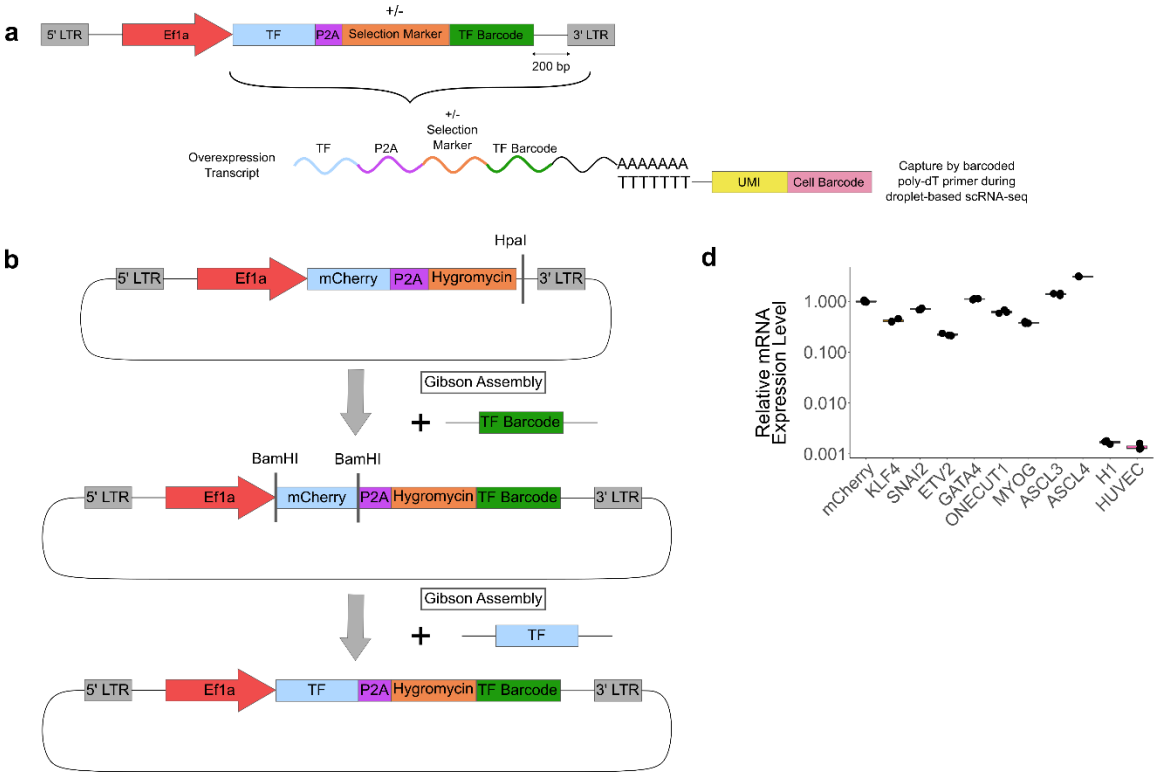
A mCherry expressing H1 cell line was created by transducing H1 cells with a lentivirus containing the EF1α promoter driving expression of the mCherry transgene, internal ribosome entry site (IRES) and a puromycin resistance gene. Cells were then maintained under constant puromycin selection at a dose of 0.75 µg/ml. mCherry labelled H1 cells were transduced with either *ETV2* lentivirus or control mCherry lentivirus, hygromycin selection was started on day 2 and cells were used for tube formation assay on day 6.

Growth-factor reduced Matrigel (Corning) was thawed on ice and 250 µl was deposited cold per well of a 24-well plate. The deposited Matrigel was incubated for 60 minutes at 37 °C, 5% CO_2, to allow for complete gelation and the *ETV2*-transduced or control cells were then seeded on it at a density of $3.2x10^5$ cells per well in a volume of 500 µl EGM. Imaging was conducted 24 hours after deposition of the cells.

2.4 Results

2.4.1 Engineering a single cell RNA-seq based overexpression screening system

We designed an ORF overexpression vector (TF-Hygro, **Figure 2.2.a**) such that each TF was paired with a unique 20 bp barcode sequence located 200 bp upstream of the lentiviral 3'-long terminal repeat (LTR) region. This yields a polyadenylated transcript bearing the barcode proximal to the 3' end, facilitating efficient capture and detection in scRNA-seq (**Figure 2.1**). To construct the ORF library, transcription factors were amplified out of a multi-tissue human cDNA pool or directly synthesized as double-stranded DNA fragments, and individually cloned into a backbone vector (**Figure 2.2.b**). The final library consisted of 61 developmentally critical or pioneer TFs (**Figure 2.2.c**), and overexpression was confirmed for select TFs by qRT-PCR (**Figure 2.2.d**).



c Library Composition

ASCL1	FLI1	HAND2	LMX1A	NEUROG1	SIX2	TFAP2C
ASCL3	FOXA1	HNF1A	MEF2C	NEUROG3	SNAI2	mCherry
ASCL4	FOXA2	HNF1B	MESP1	NRL	SOX10	
ASCL5	FOXA3	HNF4A	MITF	ONECUT1	SOX2	
ATF7	FOXP1	HOXA1	MYC	OTX2	SOX3	
CDX2	GATA1	HOXA10	MYCL	PAX7	SPI1	
CRX	GATA2	HOXA11	MYCN	POU1F1	SPIB	
ERG	GATA4	HOXB6	MYOD1	POU5F1	SPIC	
ESRRG	GATA6	KLF4	MYOG	RUNX1	SRY	
ETV2	GLI1	LHX3	NEUROD1	SIX1	TBX5	

Figure 2.2: Overexpression vector design and expression validation

(a) Schematic of lentiviral overexpression vector and capture of overexpression transcript during scRNA-seq. While the vector used in the screens contained a hygromycin resistance selection marker, it may also be designed without a selection marker. **(b)** The construction of TF-Hygro involved two steps: (i) insertion of a pool of barcodes into the backbone after digestion with HpaI, (ii) individually substituting mCherry with TFs after digestion with BamHI. **(c)** Confirmation of exogenous expression of select overexpressed TFs by qRT-PCR analysis. Data for all assays were normalized to *GAPDH* and expressed relative to control mCherry-transduced cells. Untransduced H1 hESCs and HUVECs were used as negative controls. Primers were chosen such that they amplified a portion of the transcript in the hygromycin resistance region. This was done to avoid amplification of any endogenous transcripts, and since the overexpression is driven by a single promoter the TF, P2A peptide and the hygromycin resistance are on a single transcript. **(d)** Composition of TF library.

We conducted the overexpression screens by transducing lentiviral ORF libraries into human embryonic stem cells (hESCs), maintaining them under antibiotic selection after transduction for 5 days for screens in hPSC medium, and 6 days for screens in unilineage and multilineage medium. We then performed scRNA-seq on the transduced and selected cells. TF barcodes were recovered and associated with scRNA-seq cell barcodes by targeted PCR amplification from the unfragmented cDNA, allowing genotyping of each cell for downstream analysis (**Figure 2.1**). Genotyped cell counts, while limited in sample size, also allowed us to estimate fitness (**Figure 2.3.a**).

We then used the Seurat computational pipeline to cluster cells based on their gene expression profiles, and identified over-enrichment of TFs in specific clusters using Fisher's exact test (**Figure 2.3.b**)[114]. We used a linear model to identify genes whose expression levels are appreciably changed by the perturbation[108]. For downstream analysis, we focused on TFs that were either significantly enriched for at least one cluster (p-value < 10^{-12}) or had at least 50 significant differentially expressed genes (**Figure 2.3.b**).

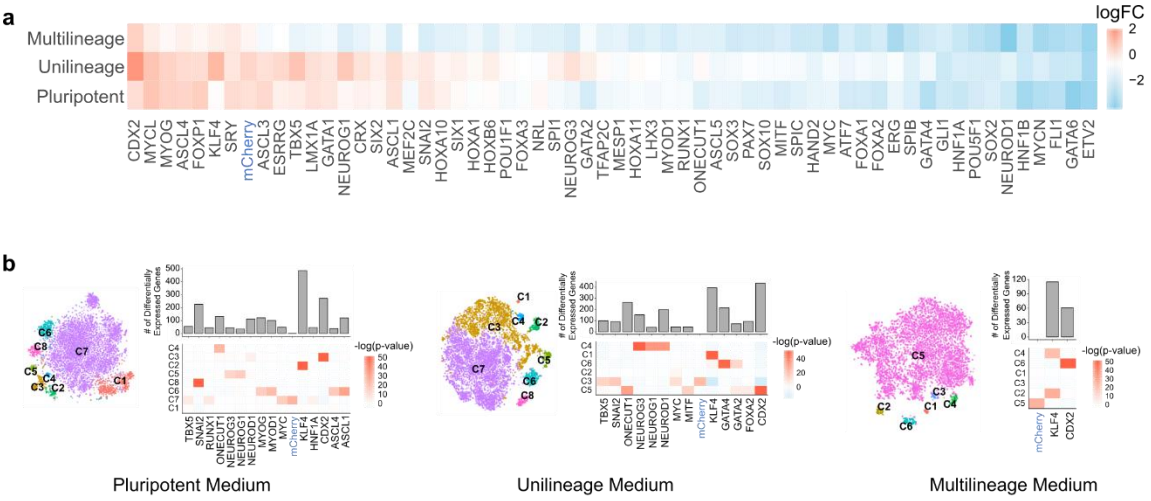


Figure 2.3: Fitness and transcriptomic effects of TF overexpression

(a) Fitness effect of TFs: log fold change of individual TFs, calculated as cell counts normalized against plasmid library read counts. **(b)** t-SNE projection (left panels), and differential gene expression and cluster enrichment of significant TFs (right panels) from screens in different growth medium conditions: pluripotent stem cell medium, unilineage medium and multilineage medium. The TFs were chosen as significant with the following criteria: cluster enrichment with a p-value of less than 10^{-12}, or if the TF drove differential expression of more than 50 genes.

2.4.2 Studying growth condition-specific effects of TF overexpression

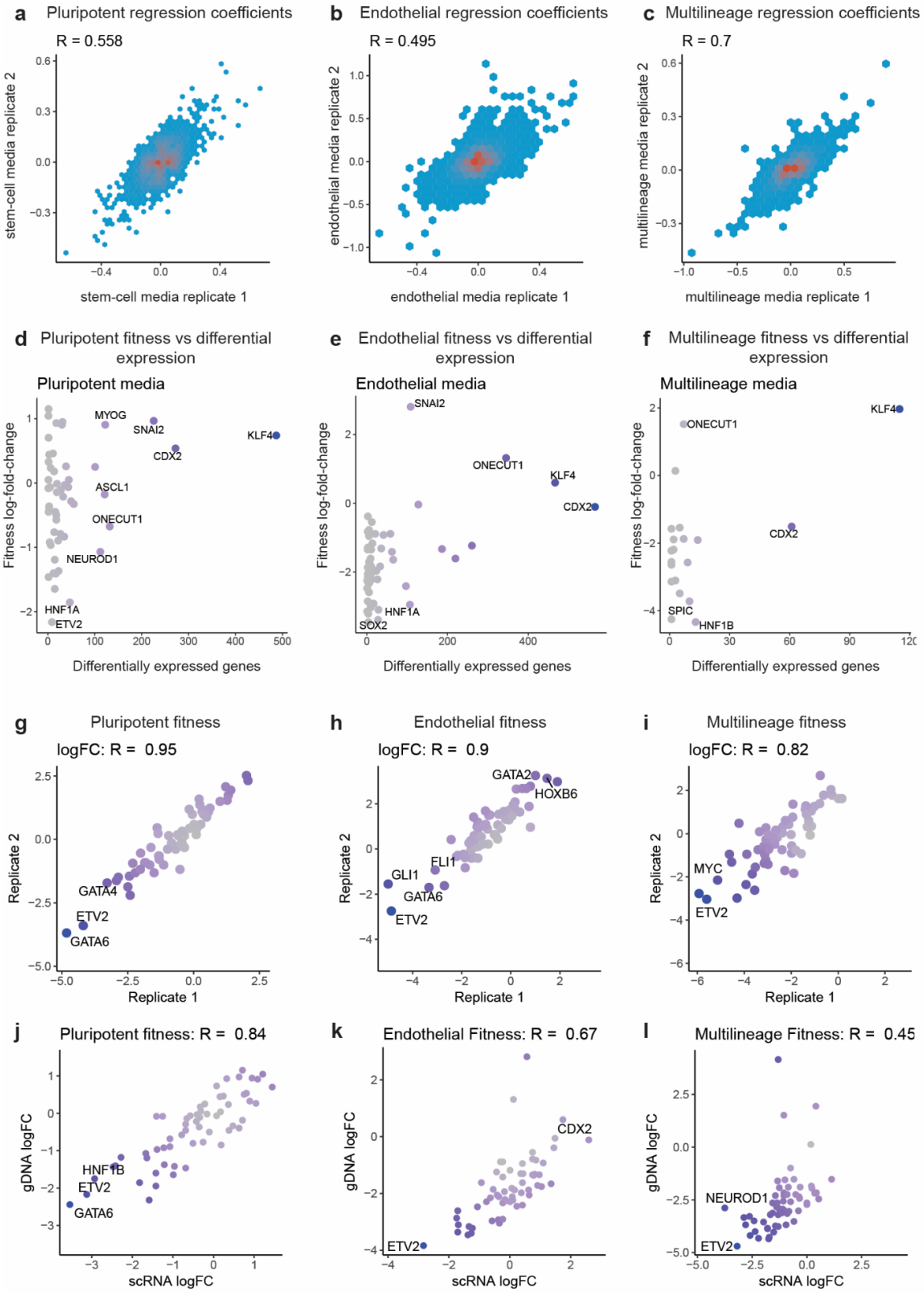
We used the SEUSS framework to assess the pluripotent cell state response to TF overexpression, and to study the interplay of TF overexpression and growth conditions. We conducted two replicate screens with the 61 element TF-Hygro library in each of three different medium conditions: hPSC medium, a unilineage medium, specifically endothelial growth medium (EGM), and a multilineage (ML) differentiation medium, specifically a high serum growth medium. We aggregated 7,728 cells across the hPSC medium screens, 10,137 cells across the unilineage medium screens, and 6,807 cells across the ML medium screens. The experimental replicates for each medium condition were well correlated (**Figure 2.4.a-c**), implying overall reproducibility.

We found that transcriptomic changes do not necessarily correlate with changes in fitness (**Figure 2.4.d-f**), demonstrating that profiling both fitness and the transcriptome is necessary for

understanding cell state changes. We also found that fitness estimates from cell counts were well correlated between replicates (**Figure 2.4.g-i**) and correlated with bulk fitness from genomic DNA despite the limited sampling of cell counts via scRNA-seq (**Figure 2.4.j-l**). Among the most significantly depleted TFs across medium conditions, was the haemato-endothelial master regulator *ETV2*, (**Figure 2.3.a**).

Figure 2.4: Correlation between replicates, between fitness and transcriptomic effects, and between bulk and scRNA-seq derived fitness

(a) Correlation between coefficients in the pluripotent medium screens. **(b)** Correlation between coefficients in the unilineage medium screens. **(c)** Correlation between coefficients in the multilineage medium screens. For **(a)-(c)**, correlation was between regression coefficients, with each coefficient representing the effect of a TF on an individual gene. We subset to coefficients that are nonzero with an adjusted p-value (FDR) of less than 0.5 in *either* replicate to filter out coefficients that are zero in both replicates. **(d)-(f)** Correlation of the number of differentially expressed genes for each TF vs the fitness effect (log-FC) for: **(d)** Pluripotent medium. **(e)** Unilineage medium (EGM). **(f)** Multilineage medium. **(g)-(i)** Correlation between replicates of fitness estimates from scRNA-seq genotyped cell counts for: **(g)** Pluripotent medium. **(h)** Unilineage medium (EGM). **(i)** Multilineage medium. **(j)-(l)** Correlation between log fold change of TF counts vs plasmid library control for genomic DNA reads vs cell counts fitness for: **(j)** Pluripotent medium **(k)** Unilineage medium (EGM) **(l)** Multilineage medium


a
Pluripotent regression coefficients
R = 0.558
stem-cell media replicate 2
stem-cell media replicate 1
b
Endothelial regression coefficients
R = 0.495
endothelial media replicate 2
endothelial media replicate 1
c
Multilineage regression coefficients
R = 0.7
multilineage media replicate 2
multilineage media replicate 1
d
Pluripotent fitness vs differential expression
Pluripotent media
MYOG
SNAI2
KLF4
CDX2
ASCL1
ONECUT1
NEUROD1
HNF1A
ETV2
Fitness log-fold-change
Differentially expressed genes
e
Endothelial fitness vs differential expression
Endothelial media
SNAI2
ONECUT1
KLF4
CDX2
HNF1A
SOX2
Fitness log-fold-change
Differentially expressed genes
f
Multilineage fitness vs differential expression
Multilineage media
KLF4
ONECUT1
CDX2
SPIC
HNF1B
Fitness log-fold-change
Differentially expressed genes
g
Pluripotent fitness
logFC: R = 0.95
GATA4
ETV2
GATA6
Replicate 2
Replicate 1
h
Endothelial fitness
logFC: R = 0.9
GATA2
HOXB6
FLI1
GLI1
GATA6
ETV2
Replicate 2
Replicate 1
i
Multilineage fitness
logFC: R = 0.82
MYC
ETV2
Replicate 2
Replicate 1
j
Pluripotent fitness: R = 0.84
HNF1B
ETV2
GATA6
gDNA logFC
scRNA logFC
k
Endothelial Fitness: R = 0.67
CDX2
ETV2
gDNA logFC
scRNA logFC
l
Multilineage Fitness: R = 0.45
NEUROD1
ETV2
gDNA logFC
scRNA logFC

To interpret the effects of significant TFs, we used the regression coefficients of the linear model to build a weighted gene-to-gene co-regulatory network, where genes with a highly weighted edge between them respond to TF overexpression in a similar manner (**Figure 2.5.a**). Applying this approach to the screens, we identified 9 altered gene modules via a graph clustering algorithm[115]. Many of these gene modules showed a strong enrichment for Gene Ontology (GO) terms, and gene module identity was assigned using GO enrichment paired with manual inspection of genes in each module (**Figure 2.5.b**).

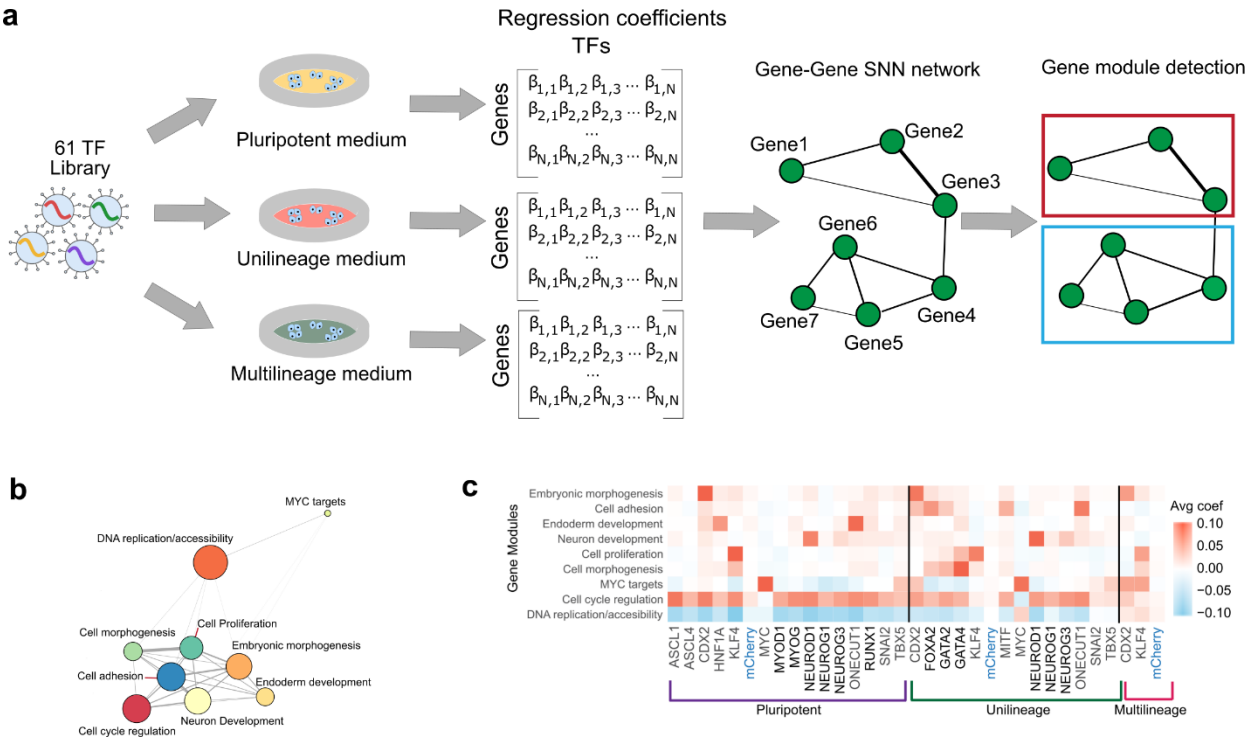


Figure 2.5: Gene co-perturbation network and gene module effects of TF overexpression

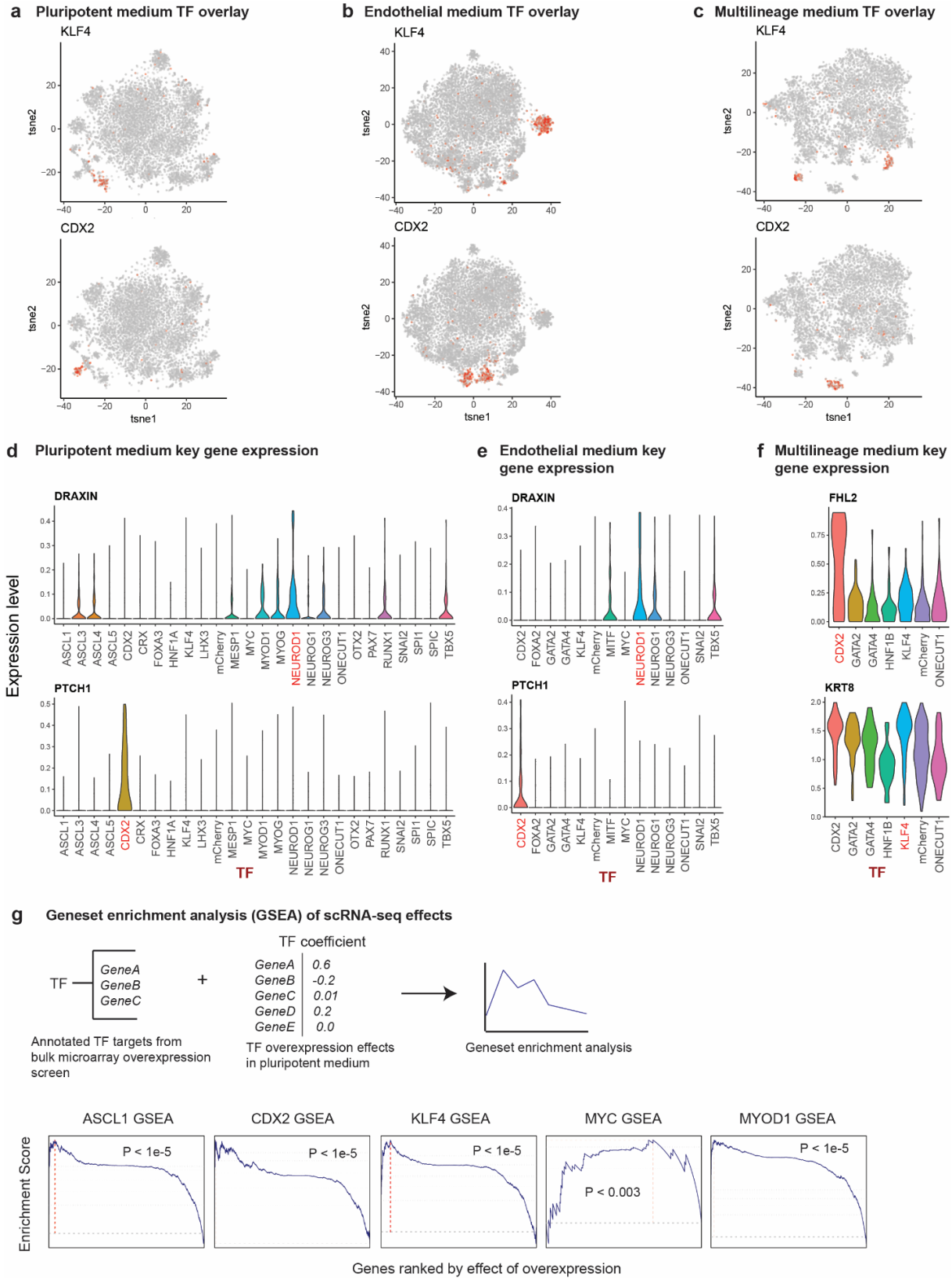
(a) A SNN network is built from the linear model coefficients and the network is then segmented into gene modules. Genes have a highly weighted edge between them if they respond similarly to TF overexpression. **(b)** Gene module network: Node size indicates the number of genes in the module; Edge size indicates distance between modules. (**c**) Effect of TF overexpression on gene modules in different medium conditions, effect size was calculated as the average of the linear model coefficients for a given TF perturbation across all genes within a module.

We next calculated the effect of each significant TF on the gene modules (**Figure 2.5.c**). While certain TFs (*CDX2*, *KLF4*) show strong cluster enrichment (**Figure 2.3.b, 2.6.a-c**) and consistent gene module effects across media conditions, some TFs have growth medium-specific

effects. For instance, some *HNF1A* effects were specific to hPSC medium and some *GATA4* effects were specific to EGM (**Figure 2.5.c**). We found that the annotated neural specifier *NEUROD1*[119–122] shows strong effects on the neuron differentiation module and upregulates genes known to play a role in neuronal development (**Figure 2.6.d-e**). In the pluripotent medium condition, we also find that *HNF1A* and *ONECUT1*, strongly upregulate the endoderm development module, reflecting their known roles in endoderm development[123–130] (**Figure 2.5.c**). Across medium conditions, *CDX2* upregulates the embryonic morphogenesis module and genes known to play a role in extra-embryonic patterning (**Figure 2.6.d-f**), potentially reflecting its role in trophectoderm development[131,132]. To benchmark our results, we compared the effects for significant TFs in hPSC medium with a previously reported bulk microarray screen of TF overexpression in mESCs[133]. We used GSEA to assess the overlap between the overexpression effects in our screens and the annotated TF targets in the bulk microarray screen. We found significant enrichment between the TFs present in both screens (**Figure 2.6.g**).

Figure 2.6: Analysis and validation of significant TFs

(a) tSNE plot of screens in pluripotent medium, color coded by select TFs (*KLF4*, *CDX2*) **(b**) tSNE plot of screens in unilineage medium, color coded by select TFs (*KLF4*, *CDX2*) **(c**) tSNE plot of screens in multilineage medium, color coded by select TFs (*KLF4*, *CDX2*) **(d**) Expression level of *DRAXIN* and *PTCH1* in screens in pluripotent medium with the TFs expected to upregulate expression of each highlighted in red **(e**) Expression level of *DRAXIN* and *PTCH1* in screens in unilineage medium with the TFs expected to upregulate expression of each highlighted in red **(f**) Expression level of *FHL2* and *KRT8* in screens in unilineage medium with the TFs expected to upregulate expression of each highlighted in red. **(g**) Geneset enrichment analysis for homologous genes in mESCs upon overexpression of TFs[133]. TFs present in both datasets – *ASCL1*, *CDX2*, *MYC, KLF4* and *MYOD1* – display a highly significant degree of overlap in their effects.


a Pluripotent medium TF overlay
KLF4
CDX2
tsne1
tsne2
b Endothelial medium TF overlay
c Multilineage medium TF overlay
d Pluripotent medium key gene expression
DRAXIN
PTCH1
Expression level
TF
e Endothelial medium key gene expression
f Multilineage medium key gene expression
FHL2
KRT8
g Geneset enrichment analysis (GSEA) of scRNA-seq effects
TF
GeneA
GeneB
GeneC
Annotated TF targets from bulk microarray overexpression screen
TF coefficient
GeneA 0.6
GeneB -0.2
GeneC 0.01
GeneD 0.2
GeneE 0.0
TF overexpression effects in pluripotent medium
Geneset enrichment analysis
ASCL1 GSEA
CDX2 GSEA
KLF4 GSEA
MYC GSEA
MYOD1 GSEA
P < 1e-5
P < 0.003
Enrichment Score
Genes ranked by effect of overexpression

2.4.3 Biological Effects of Significant TFs

We then sought to investigate the effects of two TFs, *SNAI2* and *KLF4*, which significantly perturb the hPSC transcriptome (**Figure 2.3.b, Figure 2.5.c**). Since *KLF4* and *SNAI2* are known to play critical and opposing roles in EMT[134–136] we assessed whether they cause changes along an EMT-like axis in hPSCs as well. A PCA analysis using 200 genes from a consensus EMT geneset from MSigDB[117] demonstrated a stratification of *KLF4*-transduced cells towards an epithelial-like state and *SNAI2*-transduced cells towards a mesenchymal-like state (**Figure 2.7.a**). The scRNA-seq data also demonstrated expression level changes in signature genes consistent with EMT (**Figure 2.7.b**), which we confirmed with qRT-PCR (**Figure 2.7.c-e**). We also confirmed protein expression changes with immunofluorescence staining of *EPCAM* and *VIM* (**Figure 2.7.f**).

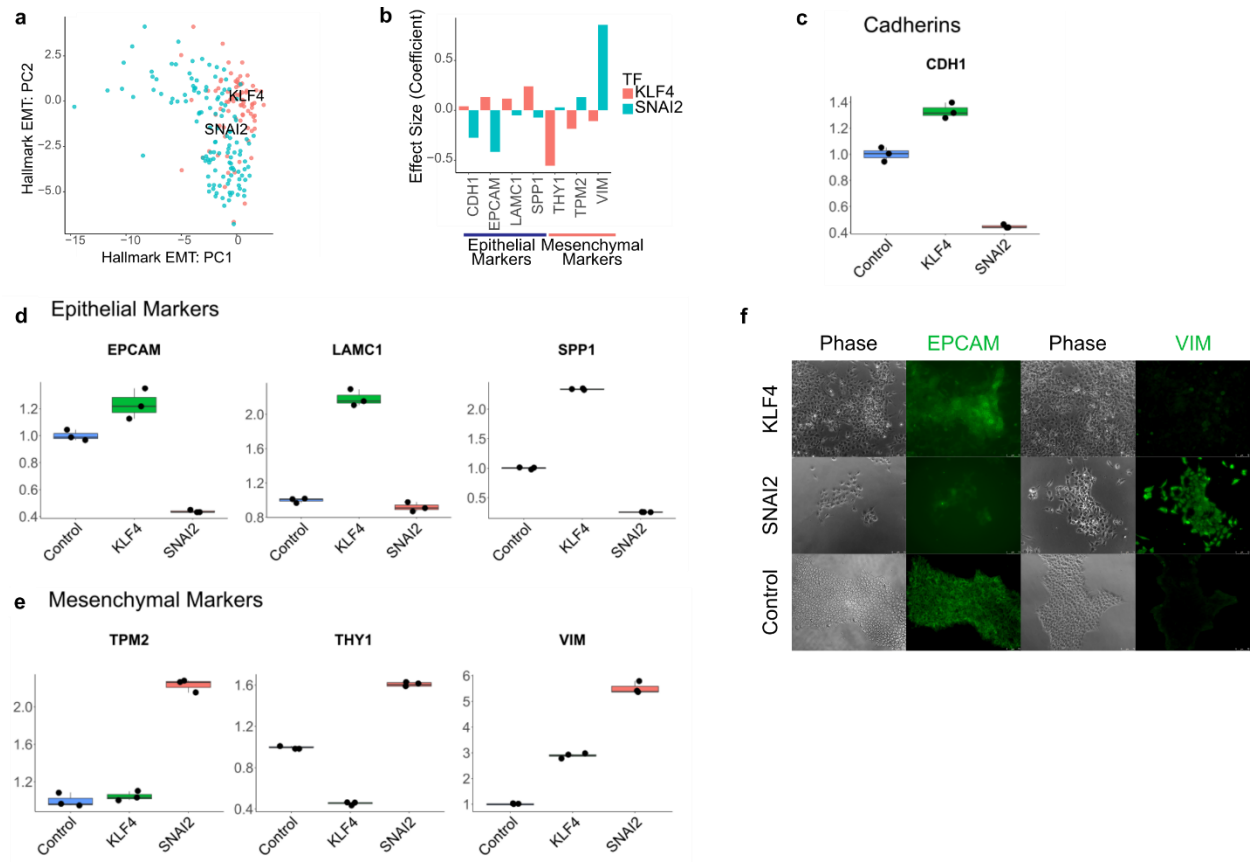


Figure 2.7: Biological effects of TF overexpression: *KLF4* and *SNAI2* as opposing drivers in EMT

(a) PC plot of performing PCA on 200 genes from the Hallmark Epithelial Mesenchymal Transition geneset from MSigDB[117]. **(b**) Effect of *KLF4* and *SNAI2* on selected epithelial and mesenchymal markers. **(c**) qRT-PCR analysis of signature cadherin during EMT: *CDH1* at day 5 post-transduction in pluripotent stem cell medium. **(d**) qRT-PCR analysis of signature epithelial marker genes during EMT: *EPCAM*, *LAMC1* and *SPP1* at day 5 post-transduction in pluripotent stem cell medium. **(e**) qRT-PCR analysis of signature mesenchymal marker genes during EMT: *TPM2*, *THY1* and *VIM* at day 5 post-transduction in pluripotent stem cell medium. Data for all assays were normalized to *GAPDH* and expressed relative to control cells. **(f**) Transmission and immunofluorescence micrographs of *EPCAM*- and *VIM*-labelled day 5 *KLF4*-, *SNAI2*- or mCherry-transduced cells.

To demonstrate the power of fitness effects in discovering TFs with significant impact on reprogramming, we focused on *ETV2*, which has the greatest average fitness loss across medium conditions (**Figure 2.3.a**). Since *ETV2* is known to drive reprogramming from fibroblasts[137] and promote differentiation of endothelial cells from hPSCs[138,139], we hypothesized that the reduced fitness could be due to a proliferation disadvantage if *ETV2*-transduced cells are undergoing reprogramming without division. Focused experiments revealed that while *ETV2*-transduced cells undergo extensive cell death in pluripotent medium, there is a morphology change indicative of

an endothelial phenotype in EGM (**Figure 2.8.a**). Immunofluorescence revealed a distinct distribution of *CDH5*, with localization at cell-cell junctions similar to human umbilical vein endothelial cells (HUVECs) (**Figure 2.8.b**). Functional testing confirmed tube formation (**Figure 2.8.c**), while qRT-PCR assays demonstrated strong upregulation of the key endothelial markers *CDH5*, *PECAM1* and *VWF* (**Figure 2.8.d**), thus suggesting that a single TF, *ETV2*, may be able to drive differentiation from a pluripotent to an endothelial-like state.

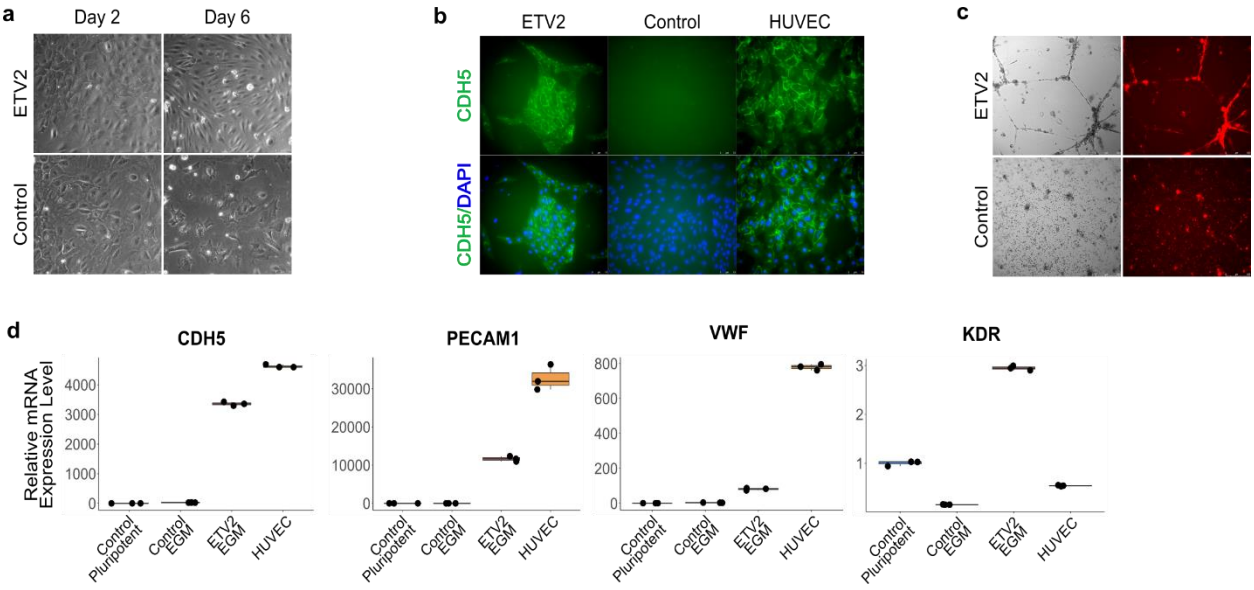


Figure 2.8: Biological effects of TF overexpression: *ETV2* as a driver of reprogramming to an endothelial-like state

(a) Morphology change for cells transduced with either *ETV2* or mCherry in EGM. **(b**) Immunofluorescence micrograph of *CDH5* labelled day 6 *ETV2*- or mCherry-transduced cells and HUVECs. **(c**) Tube formation assay for day 6 *ETV2*- or mCherry-transduced cells. **(d**) qRT-PCR analysis of signature endothelial genes *CDH5*, *PECAM1*, *VWF* and *KDR*, at day 6 post-transduction. Data were normalized to *GAPDH* and expressed relative to control cells in pluripotent stem cell medium.

2.4.4 Screening Mutant Gene Libraries and Gene Families

Since *MYC* was found to drive significant transcriptomic changes in hPSC medium (**Figure 2.3.b**), we chose to assay *MYC* mutants to demonstrate the ability of SEUSS to systematically screen mutant forms of proteins. We constructed a library of mutant *MYC* proteins, with both functional domain deletions and (**Figure 2.9.a**), and hotspot mutations[140]. Screening

this Myc-Hygro library in hPSC medium, we found that hotspot mutations, and deletion of the nuclear localization signal (NLS) sequence maintain an effect similar to the wild type *MYC*, suggesting that these mutations have minimal effect. A majority of the domain deletions show a marked reduction in both the number of differentially expressed genes, and the activation of MYC target module genes, suggesting that deleting entire functional domains is deleterious for function of *MYC* in hPSCs, as one might expect (**Figure 2.9.b**).

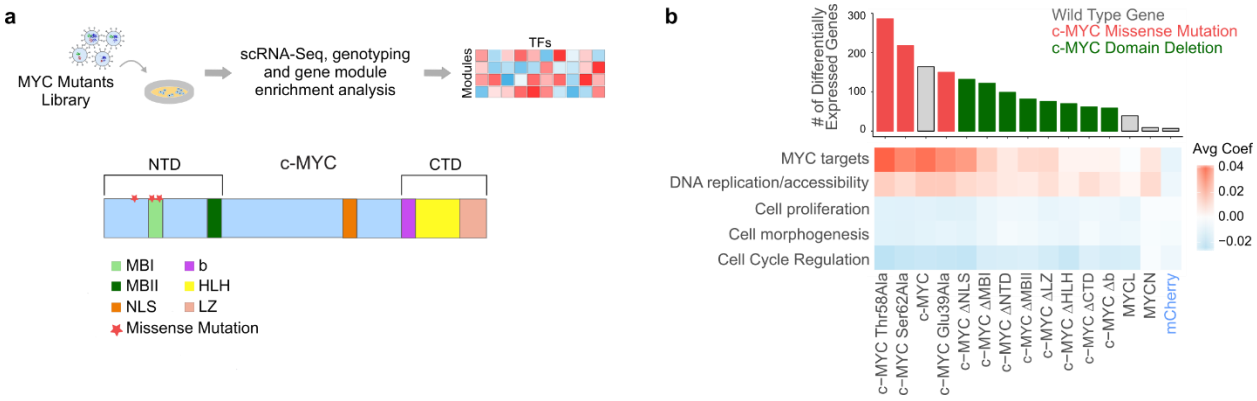


Figure 2.9: Screening mutant genes

(a) Schematic of workflow for *c-MYC* mutant library screen and schematic of functional domains of *c-MYC*: MYC Box I (MBI) and MYC Box II (II) which are essential for transactivation of target genes are housed in the amino-terminal domain (NTD); the basic (b) helix-loop-helix (HLH) leucine zipper (LZ) motif, which is required for heterodimerization with the MAX protein is housed in the carboxy-terminal domain (CTD); the nuclear localization signal domain (NLS) is located in the central region of the protein. **(b)** Effect of *MYC* mutant overexpression on number of differentially expressed genes and on gene modules. For heatmap, effect size was calculated as the average of the linear model coefficients for a given TF perturbation across all genes within a module.

The consistent and strong effects of *KLF4* overexpression motivated the investigation of the *KLF* zinc finger transcription factor family[141] (**Figure 2.10.a**) as a demonstration of SEUSS' utility in studying perturbation patterns across gene families. A KLF-Hygro screen including all 17 members of this protein family was conducted in hPSC medium. Gene module analysis showed that Group I *KLF* family members, including *KLF4* and *KLF5*, have similar effects, as does *KLF17* (**Figure 2.10.b**), which may reflect their similar role in promoting epithelial cell states[142–144]. We benchmarked the results obtained in the *KLF* gene family screen via SEUSS, against bulk RNA-

seq for *KLF4*, *KLF5* and *KLF17*, the 3 *KLF* family members with the highest number of differentially expressed genes and found the results to be correlated (**Figure 2.10.c**).

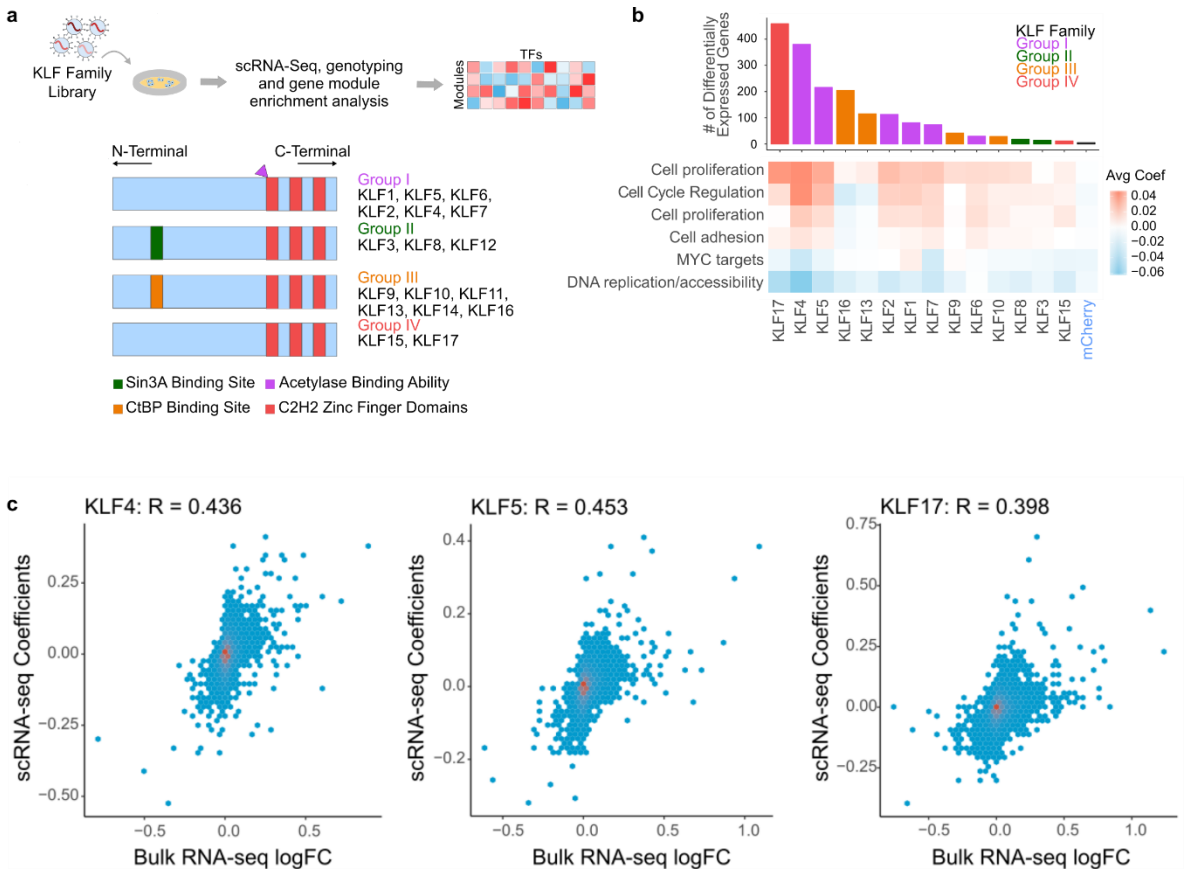


Figure 2.10: Screening gene families

(a) Schematic of workflow for *KLF* gene family screen and schematic of *KLF* gene family protein structure grouped by common structural and functional features **(b**) Effect of *KLF* family overexpression on number of differentially expressed genes and on gene modules. For heatmap, effect size was calculated as the average of the linear model coefficients for a given TF perturbation across all genes within a module. **(c)** Correlation of scRNA-seq regression coefficients versus log fold change in bulk RNA-seq for *KLF4*, *KLF5* and *KLF17*. Log fold change was calculated from bulk RNA-seq data for cells transduced with a TF versus cells transduced with control mCherry virus.

2.4.5 Drawbacks – Barcode Shuffling

Since barcode shuffling has been identified as a key factor limiting the power and sensitivity of single-cell based screening[107,145–148], we also explored an alternate version of the overexpression vector (TF-NoHygro, **Figure 2.11.a**) to minimize template-switching events during

lentiviral packaging. The rate at which the association between the ORF and barcode is lost due to template-switching is proportional to the length of the constant region between the ORF and barcode, which in this case is the selection marker. In the TF-NoHygro vector, the selection marker was excised and the ORF was placed immediately adjacent to the barcode sequence, with only a 25 bp priming sequence between them (**Figure 2.11.a**). The absence of a selection marker does not impact analysis of results from the screens since we could exactly determine which perturbation was delivered to each cell.

We then assessed barcode shuffling rates for pooled virus production for the TF-Hygro and TF-NoHygro designs, compared to a control where each library element in the TF-Hygro format was packaged individually and then pooled (**Figure 2.11.a**). We used a 14-element subset of the full library (**Figure 2.11.b**) to quantify these shuffling rates. While the TF-Hygro design yielded a ~36% barcode shuffling rate, the TF-NoHygro design had a negligible shuffling rate (**Figure 2.11.c**).

Given the degree of barcode shuffling in the TF-Hygro format, we also assessed the reproducibility of the assay results obtained using the TF-Hygro format versus the TF-NoHygro format for the 14 TF sub-library (**Figure 2.11.a**). Virus for both vector formats was produced in a pooled manner, so that the TF-Hygro format would suffer from barcode shuffling, while the TF-NoHygro format would not. We conducted scRNA-seq screens in both vector formats and found that regression coefficients for cells overexpressing a single TF were well correlated between the vector formats (**Figure 2.11.d**). Hence, results obtained from pooled screens with TF-Hygro are still valid, albeit with a reduction in power.

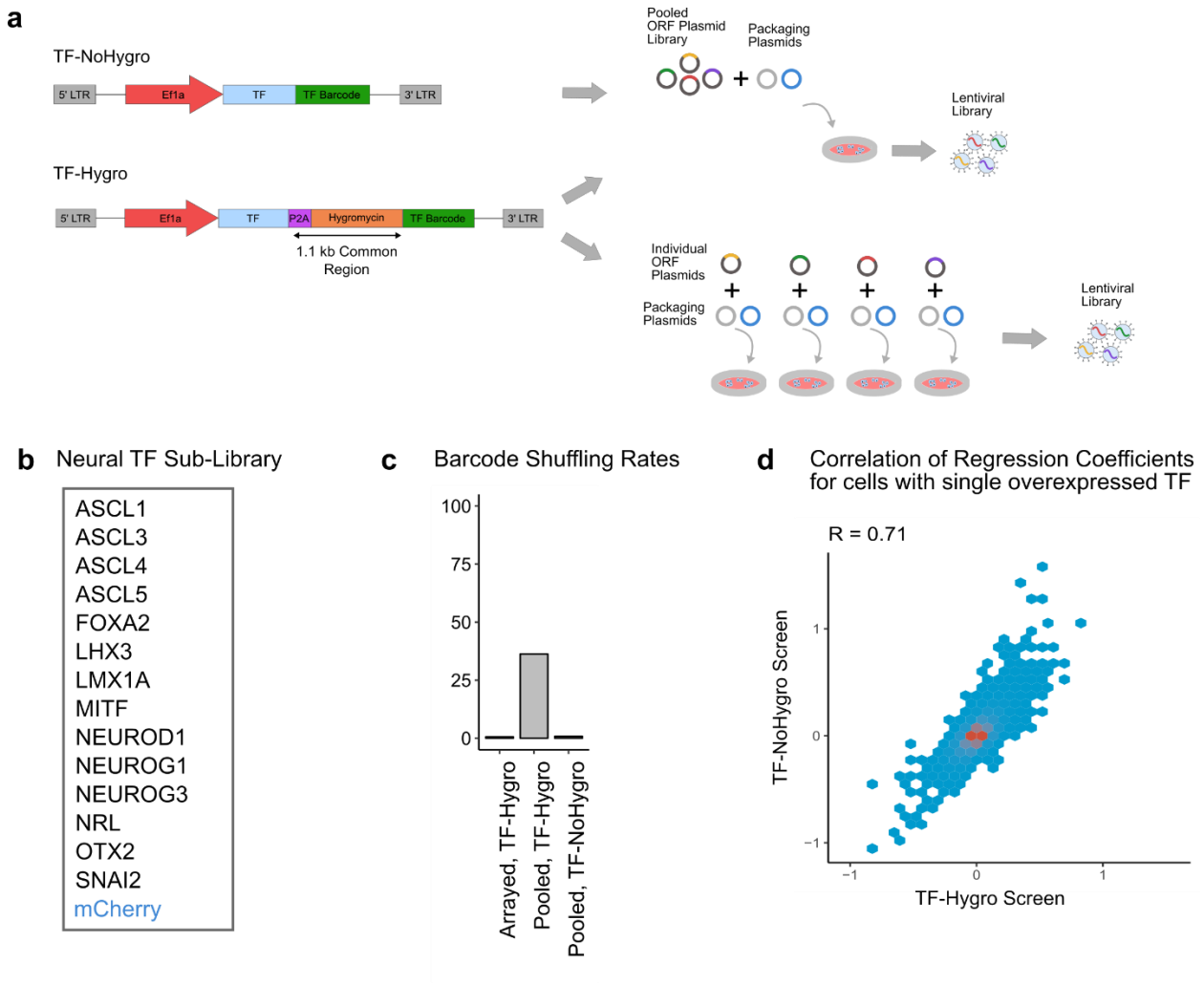


Figure 2.11: Barcode shuffling

(a) Schematic of different lentiviral ORF overexpression vector designs and pooled vs arrayed packaging methods. **(b)** 14-element Neural TF sub-library. **(c)** Quantification of shuffling rates for individually packaged virus using the TF-Hygro design, pooled virus packaging using the TF-Hygro design, and pooled virus packaging using the TF-NoHygro design. **(d)** Correlation between single TF regression coefficients for scRNA-seq screens on the Neural TF library using the original and updated vectors.

2.5 Discussion

To our knowledge, this is the first demonstration of a high-throughput gene over-expression screening approach that simultaneously assays both fitness and transcriptomic effects. Our use of ORF overexpression drove strong phenotypic effects, allowing us to capture

subtle transcriptomic signals with fewer cells per perturbation than some of the CRISPR-based screens, while the versatility of SEUSS was demonstrated by mapping context dependent effects of TFs, assaying mutant forms of a TF, as well as all of the the TFs in a gene family.

Our studies also revealed important considerations in the execution and interpretation of such screens. Consistent with recent studies[145,146], we observed shuffling of distally located barcodes due to recombination during pooled lentiviral packaging. While viral particles may be produced without risk of recombination by individually packaging each vector, this hinders screen scalability. We also noted that there are limitations on the maximum infectivity for stem cells with lentiviral vectors[149,150], reducing the probability of observing combinatorial overexpression effects. Further engineering of the overexpression vector is necessary to enable large-scale and combinatorial overexpression screens. We also observed that overexpression levels will vary with the gene being expressed, which could affect screens sensitive to such variations. Further, in our assays, since hPSCs were transduced with pooled libraries, transcriptomic changes driven by cell-cell interactions could increase variability. We also aggregated replicates to increase our power after demonstrating that the replicate experiments showed similar effects. This may not always be possible, as screens in specific cell types or culture conditions may result in more variability amongst replicates. Furthermore, in these experiments, we chose a compact library size to ensure that within a single scRNA-seq run of up to 10,000 cells, each perturbation was represented by a statistically significant number of cells. However, given the development of combinatorial indexing methods[151,152] that can profile hundreds of thousands of cells simultaneously, we anticipate SEUSS to be scalable to all known TFs.

Taken together, SEUSS has broad applicability to study the effects of overexpression in diverse cell types and contexts; it may be extended to novel applications such as screening of protein mutagenesis or screening the effects of synthetic proteins. In combination with other

methods of genetic and epigenetic perturbation it may allow us to generate a comprehensive understanding of the pluripotent and differentiation landscape.

2.6 Acknowledgements

We acknowledge Dr Tse Nga Ng for her support of U.P., members of the Mali and Zhang lab for helpful discussions, and the UCSD Institute for Genomic Medicine sequencing core for their support on scRNA-seq sample preparation and sequencing. This work was generously supported by the following sources: UCSD Institutional Funds, the Burroughs Wellcome Fund (1013926), the March of Dimes Foundation (5-FY15-450), the Kimmel Foundation (SKF-16-150), and NIH grants (R01HG009285, RO1CA222826, RO1GM123313).

Chapter 2 in part is a reprint of the material Parekh, U.*, Wu, Y.*, Zhao, D., Worlikar, A., Shah, N., Zhang, K. & Mali, P. Mapping Cellular Reprogramming via Pooled Overexpression Screens with Paired Fitness and Single Cell RNA-Sequencing Readout, *Cell Systems*, 7 (5), 548-555.e8 (2018) (* equal contribution). The dissertation author was one of the two primary authors of the study.

Chapter 3: Development of a multiplexed screening method to study oncogenicity of genes and variants across lineages

3.1 Abstract

Deconstructing tissue-specific effects of genes and variants on proliferative advantage is critical to understanding cellular transformation and to systematic selection of cancer therapeutics. Dissecting these specificities at scale requires integrated methods for multiplexed genetic screens tracking fitness across time, across human cell types, and in a suitable cellular niche since functional differences also depend on physiological cues. Towards this, we present a novel approach, harnessing single-cell cancer driver screens in teratomas coupled with hit enrichment by serial teratoma reinjection, to simultaneously screen drivers across multiple lineages *in vivo*. Using this system, we analyzed population shifts and lineage-specific enrichment for 51 cancer associated genes and gene variants, profiling over 100,000 cells spanning over 20 lineages, across two rounds of serially injected teratomas. We confirmed that *c-MYC* alone or combined with myristoylated *AKT1* potently drives proliferation in progenitor neural lineages, demonstrating signatures of malignancy. These drivers directed teratoma development to lineages representative of pediatric tumors such as medulloblastoma and rhabdomyosarcoma. Additionally, mutant $MEK1^{S218D/S222D}$ provides a proliferative advantage in mesenchymal lineages like fibroblasts. Our method provides a powerful new platform for multi-lineage longitudinal study of oncogenesis.

3.2 Introduction

The onset of cancer is often posited to be an evolutionary process enabled by the acquisition of somatic mutations and genetic lesions over time[153,154]. Yet, these mutations lead to survival advantages and cancer only if accumulated in the relevant types of cells at the appropriate stage of differentiation, at the appropriate time and in the appropriate cell state[31,155,156]. Understanding this process of tumorigenesis and dissecting the tissue-specific molecular mechanisms which govern neoplastic transformation is a longstanding goal in cancer biology. In particular, with the explosion in cancer genome sequencing data[157–162], understanding the tissue-specific oncogenicity of the growing list of genetic variants of unknown significance may lead to significant advances in building early detection systems, which would improve patient outcomes[163], as well as inform the development and application of therapeutic strategies[164].

In this regard, xenograft and genetically engineered animal models[1,2,165] have been especially useful to recapitulate the process by which healthy cells undergo transformation, but such models often do not completely capture human biology, transformation or tumorigenesis[1,2,5,6,165]. On the other hand, immortalized cell lines and primary cells have been important workhorses of cancer research[7,8], aiding in mechanistic studies, the uncovering of therapeutic vulnerabilities and understanding resistance mechanisms[166]. Yet, elucidating the wide spectrum of tissue-specific programs governing transformation[167–174] remains a challenge[175], especially as access to and culture of diverse types of primary cells is challenging. Such *in vitro* systems also exclude the environmental and physiological context which are key modulators of driver-specificity, and often lack the ability to perturb cells along their differentiation trajectories or in distinct states, an important factor in driver-specific transformation[176].

Systems which allow us to ethically gain access to developing human tissue models, recapitulating the architecture and signaling programs of native tissue, in a suitable physiological

niche could be an invaluable tool to deconstruct tissue-specific drivers of neoplastic transformation. In this regard, there has been significant work toward the use of directed differentiation of stem cells[58] in 2-dimensional monolayer culture as well as into organoids[54,55,59,177–181] to model cancer. These models capture various cell states along the developmental trajectory, while organoids also capture the diversity of cells present in specific tissue, organized in native-like architecture. However, these systems are cultured in specialized conditions, which may not represent *in vivo* settings, and lack vasculature which is a central characteristic of cancers[182]. *In vivo* engraftment of suitably differentiated human stem cells provides an avenue to study the human-specific dynamics of oncogenesis and cancer evolution in an appropriate milieu[183–186]. But, each of these systems provides access to a single or a few lineages at a time, thus, to test drivers in a panoply of lineages to assess their oncogenic potential in each is a laborious and slow process.

Recently, our group demonstrated that human pluripotent stem cell (hPSC) derived teratomas, which are typically benign tumors with differentiated cells from all three germ layers and regions of organized tissue-like architecture[187], can enable high throughput genetic screens simultaneously across cell types of all germ layers[74]. Leveraging this, we propose here a method that uses the teratoma model, in combination with single cell RNA-sequencing (scRNA-seq) based open reading frame (ORF) overexpression screens[188] and serial tumor propagation, to massively parallely assess the tissue-specific oncogenicity of genes and gene variants across a diverse set of lineages.

3.3 Materials and methods

3.3.1 Cell Culture

H1 male hESC cell line was maintained under feeder-free conditions in mTeSR1 medium (Stem Cell Technologies) supplemented with 1% antibiotic-antimycotic (Thermo Fisher Scientific).

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). Prior to lentiviral transduction, hESCs were passaged using Accutase (Innovative Cell Technologies) and plated as dissociated cells to achieve higher transduction efficiency. When passaged with Accutase, cells were plated in mTeSR1 containing Y27632 (10 μM, Tocris). H1 hESCs used started at P30 and were passaged a maximum of 4 passages before injection.

HEK 293T cells were maintained in high glucose DMEM supplemented with 10% fetal bovine serum (FBS) and 1 % antibiotic-antimycotic. HEK 293T cells were passaged using 0.05% Trypsin (Thermo Fisher Scientific).

3.3.2 Library Preparation

The lentiviral backbone plasmid, compatible with detection in scRNA-seq was constructed as previously described[1] containing the EF1α promoter, mCherry transgene flanked by BamHI restriction sites, followed by a P2A peptide and hygromycin resistance enzyme gene immediately downstream. Each driver ORF in the library was individually inserted in place of the mCherry transgene. A barcode sequence was introduced to allow for identification of the ectopically expressed transcription factor. The backbone plasmid was digested with HpaI, and a pool of 20 bp long barcodes with flanking sequences compatible with the HpaI site, was inserted immediately downstream of the hygromycin resistance gene by Gibson assembly. The vector was constructed such that the barcodes were located only 200 bp upstream of the 3'-LTR region. This design enabled the barcodes to be transcribed near the poly-adenylation tail of the transcripts and a high fraction of barcodes to be captured during sample processing for scRNA-seq.

To create the driver ORF library, individual drivers were PCR amplified out of the Cancer Pathways kit (Addgene #1000000072)[166], individual plasmids (Addgene #9053, #85140, #82262, #82297, #82175, #61852, #39872, #23776, #23688, #10745, #23231)[189–194], a human cDNA pool (Promega Corporation), or obtained as synthesized double-stranded DNA fragments (gBlocks, IDT Inc) with flanking sequences compatible with the BamHI restriction sites. The barcoded lentiviral backbone was digested with BamHI HF (New England Biolabs) at 37 ºC for 3 hours in a reaction consisting of: lentiviral backbone, 4 µg, CutSmart buffer, 5 µl, BamHI, 0.625 µl, H_20 up to 50 µl. After digestion, the vector was purified using a QIAquick PCR Purification Kit (Qiagen). Each transcription factor vector was then individually assembled via Gibson assembly. The Gibson assembly reactions were set up as follows: 100 ng digested lentiviral backbone, 3:10 molar ratio of transcription factor 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 and to extract barcode sequences. One overexpression vector was created for each ORF, thus a single unique barcode was associated with each ORF.

3.3.3 Viral Production

HEK 293T cells were maintained in high glucose DMEM supplemented with 10% fetal bovine serum (FBS).

To prepare lentivirus for the full library, to avoid barcode shuffling[145–147,188] lentivirus for each ORF was packaged independently. Cells were seeded in a 6-well plate 1 day prior to

transfection, such that they were 60-70% confluent at the time of transfection. For each well of a 6-well plate, 2.25 µl of Lipofectamine 2000 (Life Technologies) was added to 125 µl of Opti-MEM (Life Technologies). Separately 187.5 ng of pMD2.G (Addgene #12259), 750 ng of pCMV delta R8.2 (Addgene #12263) and 562.5 ng of an individual vector was added to 125 µl 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 2 ml of fresh, pre-warmed medium per well. After the incubation period, the mixture was added dropwise to each well 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). For library lentiviral production, supernatant of all ORF wells was mixed together and concentrated to a final volume of 600-800 µl, divided into aliquots and frozen at -80 ºC. To prepare lentivirus for individual vectors, the production protocol as described above was scaled up to a 15 cm dish.

3.3.4 Viral Transduction

For viral transduction, on day -1, H1 cells were dissociated to a single cell suspension using Accutase and seeded into Matrigel-coated 6-well plates at a density of $3x10^5$ cells per well in mTeSR containing ROCK inhibitor, Y27632 (10 µM, Tocris). 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). On day 1, medium was replaced with fresh mTeSR1. Hygromycin (Thermo Fisher Scientific) selection was started from day 2 onward at a selection dose of 50 µg/ml, medium containing hygromycin was replaced daily.

3.3.5 Animals

Housing, husbandry and all procedures involving animals used in this study were performed in compliance with protocols 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.

3.3.6 Teratoma Formation

A subcutaneous injection of 6-8 million PSCs in a slurry of growth factor reduced Matrigel (Corning) and mTeSR medium (1:1) was made in the right flank of anesthetized $Rag2^{-/-};\gamma c^{-/-}$ immunodeficient mice. Weekly monitoring of teratoma growth was made by quantifying approximate elliptical area (mm^2) with the use of calipers measuring outward width and height.

3.3.7 Teratoma Processing

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 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 1-2mm in diameter and subjected to standard GentleMACS (Miltenyi) protocols: Human Tumor Dissociation Kit (medium tumor settings) and Red Blood Cell Lysis Kit. From round 1 teratomas, 6-8 million dissociated live cells were resuspended in growth factor reduced Matrigel (Corning) and subcutaneously injected in the right flank of anesthetized $Rag2^{-/-};\gamma c^{-/-}$ immunodeficient mice to form round 2 tumors. For single cell RNA-seq, samples were also

processed with Dead Cell Removal Kit (Miltenyi). Single cells were then resuspended in PBS + 0.04% BSA for processing on the 10X Genomics Chromium[195] 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 ºC or resuspended in RNALater (Thermo Fisher Scientific), incubated overnight at 4 ºC, RNALater removed and then frozen at -80 ºC.

3.3.8 Barcode Amplification

Barcodes were amplified from cDNA generated by the single cell system, and gDNA from validation tumors and 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). The PCR primers used were, NEB_EC2H_Barcode_F: GACTGGAGTTCAGACGTGTGCTCTTCCGATCTAGAACTATTTCCTGGCTGTTACGCG 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 (~500 bp) of 4 reactions for each sample were pooled, size-selected and purified with Agencourt AMPure XP beads 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. 2 µg of genomic DNA was input per reaction with Kapa Hifi Hotstart ReadyMix. The PCR primers used were, EC2H_gDNA_Barcode_F: ACACTCTTTCCCTACACGACGCTCTTCCGATCTACTGTCGGGCGTACACAAATC and EC2H_gDNA_Barcode_R: GACTGGAGTTCAGACGTGTGCTCTTCCGATCTCACTGTTTAACAAGCCCGTCAGTAG. 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 (200 bp) of the two reactions for each sample were pooled, size-selected with Agencourt AMPure XP beads (Beckman Coulter, Inc.) 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.

3.3.9 Single cell RNA-seq Processing

Fastq files were aligned to a combined hg19 and mm10 reference and expression matrices generated using the count command in cellranger v3.0.1 (10X Genomics). cellranger commands were run using default settings.

3.3.10 Data Integration and Clustering

Data integration was performed on counts matrices from the following samples: 4 Round 1 teratomas perturbed with the full library, 2 Round 2 teratomas perturbed with the full library, 3 Round 1 teratomas perturbed with a library without *c-MYC* or *myr-AKT1*, and 2 Round 2 teratomas perturbed with a library without *c-MYC* or *myr-AKT1*. Integration was done using the Seurat v3 pipeline[196]. Expression matrices were filtered to remove any cells expressing less than 200 genes or expressing greater than 20% mitochondrial genes, as well as to remove any genes that are expressed in less than 0.1% of cells. DoubletFinder[197] 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 then scaled, and nUMI as well as mitochondrial gene fraction was regressed out. Principal component analysis (PCA) was performed on this matrix and 22-30 PCs were identified as significant using an elbow plot. The 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 calculate change in the cell type abundance, the number of cells of each type was summarized for each wild type and each driver library sample. This table was input to edgeR[198,199], which was then used to determine $\log_2$ fold change, p-value and false discovery rate (FDR).

3.3.11 Barcode Assignment

To assign one or more barcodes to each cell, we used a previously described method[188]. Briefly, we extracted each barcode by identifying its flanking sequences, resulting in reads that contain cell, UMI, and barcode tags. To remove potential chimeric reads, we used a two-step filtering process. First, we only kept barcodes that made up at least 0.5% of the total amount of reads for each cell. We then counted the number of UMIs and reads for each plasmid barcode within each cell, and only assigned that cell any barcode that contained at least 10% of the cell's read and UMI counts.

3.3.12 Bulk RNA Extraction and RNA-seq library preparation

RNA was extracted from cells using the 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).

Bulk RNA-seq libraries were prepared from 1000 ng of RNA using the NEBNext Ultra RNA Library Prep kit for Illumina or the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (New England Biolabs) as per the manufacturer's instructions. Libraries were sequenced on an Illumina NovaSeq platform.

3.3.13 Bulk RNA-Seq Analysis

We mapped the bulk RNA-Seq fastq files to GRCh38 and quantified read counts mapping to each gene's exon using Ensembl v99 and STAR aligner[118]. The genes by counts matrix was then processed via DESeq2[200] to normalize counts and estimate differential expression. Log fold change in gene expression of *c-MYC* and *c-MYC + myr-AKT1* driven tumors compared to wild type teratomas was used as input for geneset enrichment analysis.

Data from the TARGET initiative was obtained from the UCSC Xena browser. TARGET data as well as experimental tumor samples were processed using edgeR[198] and limma[201] to filter and normalize the data and to remove heteroscedascity. The 2000 most highly variable genes across the TARGET tumors were then selected for performing principal component analysis.

3.3.14 Barcode Counting

Barcodes amplified and sequenced from genomic DNA were aligned using Bowtie2[202,203], and then counted using MAGeCK[116].

3.3.15 Gas chromatograph-Mass spectrometry (GC-MS) sample preparation and analysis

Metabolites were extracted, analyzed, and quantified, as previously described in detail[31]. Frozen tissue was pulverized using a cellcrusher and ten to fifteen mg of frozen tissue were then homogenized with a ball mill (Retsch Mixer Mill MM 400) at 30 Hz for 5 minutes in 500 µL -20°C methanol and 200 µL of ice-cold MiliQ water. The mixture was then transferred into a 2 mL Eppendorf tube containing in 500 µL of chloroform, vortexed for 5 min followed by centrifugation with 16 000 x *g* for 5 min at 4 °C. To determine the abundances of nucleobases the interface was centrifuged with 1 mL −20 °C methanol at 16,000×*g* for 5 min at 4 °C, the pellet was hydrolyzed with 1 mL 6N HCL for 2h at 80°C, centrifuged at 16,000×*g* for 5 min and the supernatant was dried at 60 °C under airflow.

Metabolite derivatization was performed using a Gerstel MPS. Dried metabolites were dissolved in 15 µl of 2% (w/v) methoxyamine hydrochloride (Thermo Scientific) in pyridine and incubated for 60 min at 45°C. An equal volume of 2,2,2-trifluoro-N-methyl-N-trimethylsilyl-acetamide (MSTFA) (nucleobases) was added and incubated further for 30 min at 45 °C. Derivatized samples were analyzed by GC-MS using a DB-35MS column (30 m x 0.25 mm i.d. x 0.25 µm, Agilent J&W Scientific) installed in an Agilent 7890A gas chromatograph (GC) interfaced

with an Agilent 5975C mass spectrometer (MS) operating under electron impact ionization at 70 eV. The MS source was held at 230 °C and the quadrupole at 150 °C and helium was used as a carrier gas at a flow rate of 1 mL/min. The GC oven temperature was held at 80 °C for 6 min, increased to 300 °C at 6 °C/min and after 10 min increased to 325 °C at 10 °C/min for 4 min.

3.3.16 Immunofluorescence

Sections were fixed with 4% (wt/vol) paraformaldehyde in PBS at room temperature for 10 minutes. Sections were then incubated with a blocking buffer: 5% donkey serum, 0.2% Triton X-100 in PBS for 1 hour at room temperature followed by incubation with primary antibodies diluted in the blocking buffer at 4 °C overnight. Primary antibodies used were: Ki-67 (D3B5, Cell Signaling Technology; 1:100).

After overnight incubation with primary antibodies, cells were labelled with secondary antibodies diluted in 5% donkey serum in PBS for 1 hour at 37 °C. Nuclear staining was done by incubating cells with DAPI for 5 minutes at room temperature. All imaging was conducted on a Leica DMi8 inverted microscope equipped with an Andor Zyla sCMOS camera and a Lumencor Spectra X multi-wavelength fluorescence light source.

3.4 Results

3.4.1 Design of a multiplex *in vivo* screening platform

To enable a screening platform which would simultaneously allow the determination of lineages along with detection of perturbations, we implemented pooled overexpression screens in teratomas with scRNA-seq readout, using a lentiviral overexpression vector we previously developed for compatibility with scRNA-seq based pooled genetic screens[188] (**Figure 3.1.a**). The vector contains a unique 20 bp barcode for each library element, located 200 bp upstream of the lentiviral 3' long terminal repeat (LTR). This yields a polyadenylated transcript with the barcode proximal to the 3' end, thus allowing detection in droplet based scRNA-seq systems which rely on

poly-A capture. Each cancer driver was cloned into the lentiviral backbone vector, packaged individually into lentivirus particles and this individually packaged lentivirus then combined, to avoid barcode shuffling due to the recombination driven template switching inherent in pooled lentiviral packaging[145–147,188]. To allow for combinatorial driver transduction, these driver libraries were then transduced into H1 human embryonic stem cells (hESCs), which we had previously characterized to be of normal karyotype[74]. The hESCs were transduced at the highest titer where we did not see visible morphological changes, and maintained under antibiotic selection for 4 days after transduction (**Figure 3.1.b**). In addition, for cells profiled prior to injection by scRNA-seq, for the majority of drivers we observed no significant changes per cluster in the proportion of cells expressing drivers as compared to those expressing the internal negative control (**Figure 3.2.a-c**), confirming that the driver library was not significantly perturbing the hESCs from their basal state. We also confirmed that all library elements were detected in the transduced cells prior to injection (**Figure 3.2.d**).

To form teratomas with these library-transduced hESCs, we subcutaneously injected 6-8 million of these cells suspended in a 1:1 mixture of Matrigel and the pluripotent stem cell medium mTeSR1 in the right flank of anesthetized $Rag2^{-/-};\gamma c^{-/-}$ immunodeficient mice (**Figure 3.1.b**). As growth controls, wild type, unmodified H1 hESCs were similarly injected in a separate group of $Rag2^{-/-};\gamma c^{-/-}$ immunodeficient mice. The growth of all teratomas was monitored via weekly caliper-based measurements of elliptical area (**Figure 3.1.b**). Once the teratomas reached a threshold size for excision, they were extracted for downstream processing. After extraction, tumors were weighed and measured, and representative pieces were frozen for cryosectioning and histological analysis. The remaining pieces were dissociated into single cell suspensions. A part of this suspension was processed for single cell RNA-sequencing (scRNA-seq) using the droplet based 10X Genomics Chromium system, while a second part consisting of 6-8 million cells was reinjected into $Rag2^{-/-};\gamma c^{-/-}$ immunodeficient mice for serial proliferation to further enrich

transformed lineages and drivers of proliferative advantage. These reinjected round 2 tumors were monitored and processed similarly to the initially formed round 1 teratomas (**Figure 3.1.b**), to obtain histology information as well as transcriptomic profiles via scRNA-seq.

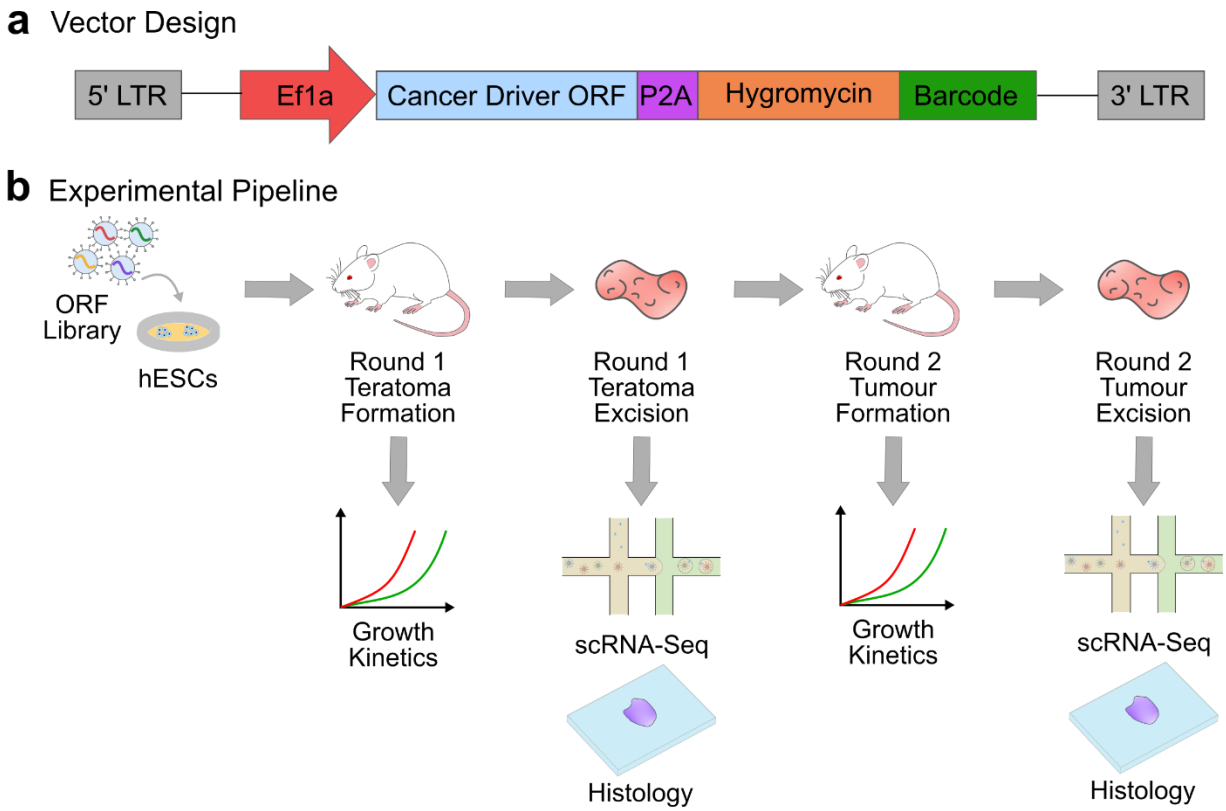
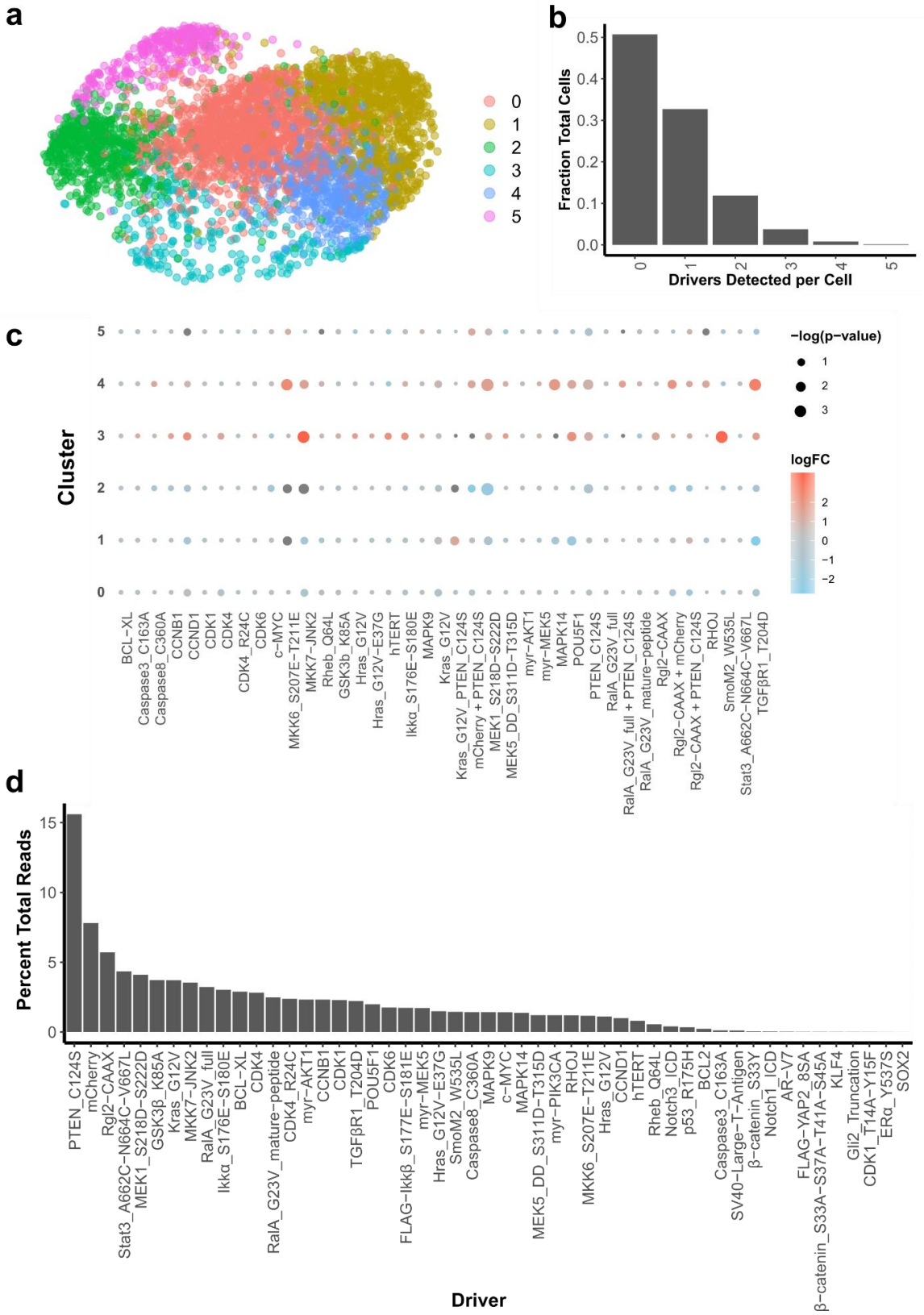


Figure 3.1: Overview of experimental design and library construction

(a) Schematic of lentiviral overexpression vector. **(b)** Schematic of experimental framework for evaluation of effects of cancer driver overexpression in developing teratomas and reinjected tumors: Individual cancer driver ORFs are cloned into the barcoded ORF overexpression vector, packaged into lentivirus then pooled for transduction of hESCs. Transduced cells are then injected subcutaneously into $Rag2^{-/-};\gamma c^{-/-}$ immunodeficient mice to form the round 1 teratomas. Teratoma growth is assayed by caliper-based measurement of approximate elliptical area of the tumor. Once the teratoma size reaches a threshold value, teratomas are excised and assayed by single-cell RNA-seq and histology, and a fraction of cells are reinjected in $Rag2^{-/-};\gamma c^{-/-}$ immunodeficient mice to form round 2 tumors for further enrichment of drivers. Round 2 tumor growth is also monitored via caliper-based measurements and at the end point tumors are assayed via scRNA-seq and histology.

Figure 3.2: Characterization of driver library transduced cells prior to injection

(**a**) UMAP visualization of cells labelled by cluster identity. (**b**) Fraction of cells for each number of drivers detected in scRNA-seq for driver library transduced H1 hESCs. (**c**) Log2 fold change and significance of proportion of cells in a cluster expressing a driver versus proportion expressing the internal control (mCherry), calculated by permutation testing. Only drivers expressed in 20 or more cells are included. (**d**) Percentage of reads detecting each library element for barcodes amplified from genomic DNA extracted from driver-library transduced H1 hESCs prior to injection.


a
0
1
2
3
4
5
b
Fraction Total Cells
0.0
0.1
0.2
0.3
0.4
0.5
0
1
2
3
4
5
Drivers Detected per Cell
c
Cluster
5
4
3
2
1
0
−log(p−value)
1
2
3
logFC
2
1
0
−1
−2
BCL−XL
Caspase3_C163A
Caspase8_C360A
CCNB1
CCND1
CDK1
CDK4
CDK4_R24C
CDK6
c−MYC
MKK6_S207E−T211E
MKK7−JNK2
Rheb_Q64L
GSK3b_K85A
Hras_G12V
Hras_G12V−E37G
hTERT
Ikkα_S176E−S180E
MAPK9
Kras_G12V
Kras_G12V_PTEN_C124S
mCherry + PTEN_C124S
MEK1_S218D−S222D
MEK5_DD_S311D−T315D
myr−AKT1
myr−MEK5
MAPK14
POU5F1
PTEN_C124S
RalA_G23V_full
RalA_G23V_full + PTEN_C124S
RalA_G23V_mature−peptide
Rgl2−CAAX
Rgl2−CAAX + mCherry
Rgl2−CAAX + PTEN_C124S
RHOJ
SmoM2_W535L
Stat3_A662C−N664C−V667L
TGFβR1_T204D
d
Percent Total Reads
0
5
10
15
PTEN_C124S
mCherry
Rgl2−CAAX
Stat3_A662C−N664C−V667L
MEK1_S218D−S222D
GSK3β_K85A
Kras_G12V
MKK7−JNK2
RalA_G23V_full
Ikkα_S176E−S180E
BCL−XL
CDK4
RalA_G23V_mature−peptide
CDK4_R24C
myr−AKT1
CCNB1
CDK1
TGFβR1_T204D
POU5F1
CDK6
FLAG−Ikkβ_S177E−S181E
myr−MEK5
Hras_G12V−E37G
SmoM2_W535L
Caspase8_C360A
MAPK9
c−MYC
MAPK14
MEK5_DD_S311D−T315D
myr−PIK3CA
RHOJ
MKK6_S207E−T211E
Hras_G12V
CCND1
hTERT
Rheb_Q64L
Notch3_ICD
p53_R175H
BCL2
Caspase3_C163A
SV40−Large−T−Antigen
β−catenin_S33Y
Notch1_ICD
AR−V7
FLAG−YAP2_8SA
β−catenin_S33A−S37A−T41A−S45A
KLF4
Gli2_Truncation
CDK1_T14A−Y15F
ERα_Y537S
SOX2
Driver

3.4.2 Deconstructing fitness effects across lineages

Using this method, we screened 51 genes and gene variants representing major signaling pathways and oncogenes[166] involved in oncogenesis, tumor proliferation, survival, cell cycle regulation and stemness regulation (**Figure 3.3**).

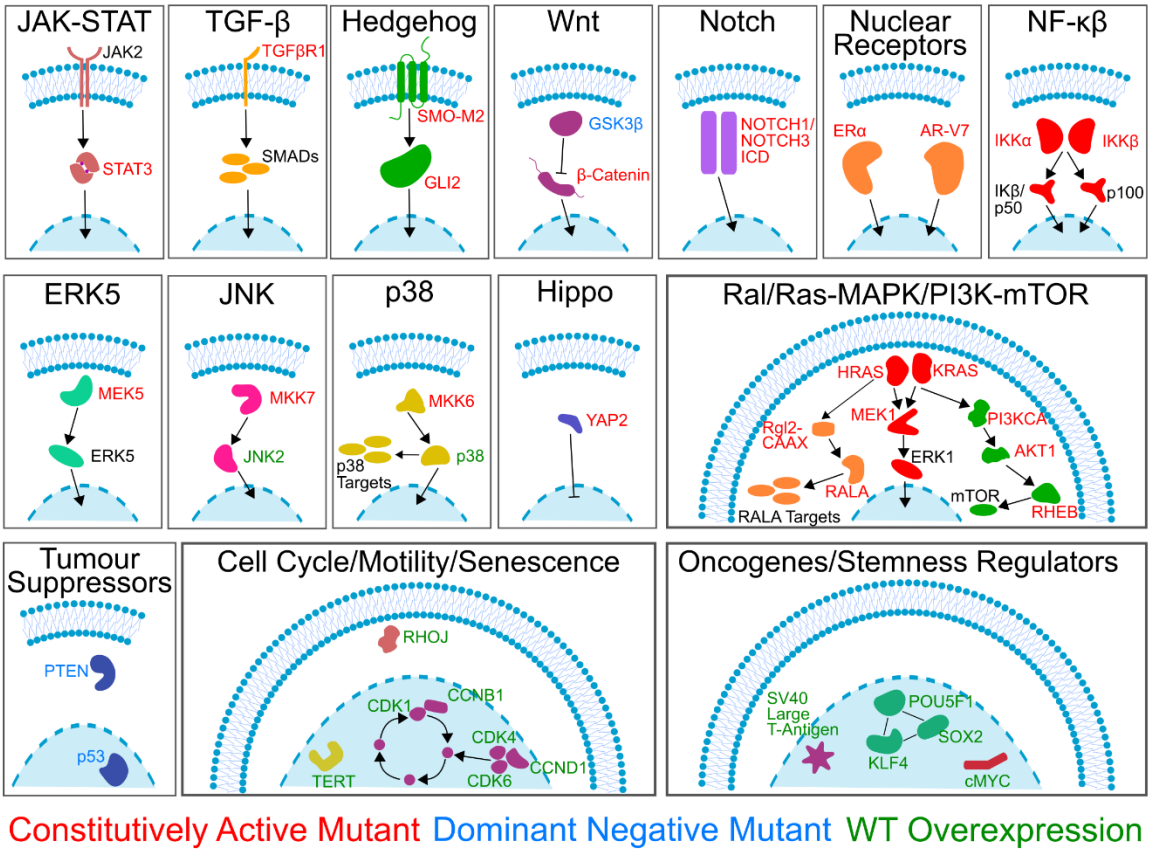


Figure 3.3: Library composition

Composition of the cancer driver ORF library, encompassing major signaling pathways, oncogenes and stemness regulators involved in oncogenesis and cancer progression.

To screen these drivers, we generated four teratomas from the perturbed hPSCs. In these first-round driver library teratomas, we observed palpable and measurable tumors between 20-27 days after injection, which grew to a size sufficient for extraction between 41-60 days after injection. In comparison, out of eight teratomas formed from unperturbed wild type cells, one did not form a tumor in 90 days of monitoring, while the seven remaining teratomas were palpable

and measurable between 27-39 days after injection and six of them grew to a size sufficient for extraction between 55-78 days after injection (**Figure 3.4.a**).

Once these round 1 driver library teratomas were excised and processed via scRNA-seq, gene expression matrices were generated from the resultant data using 10X Genomics cellranger. We then used these expression matrices with the Seurat pipeline[196] to integrate data from all four driver library teratomas and compensate for batch effects. Using this integrated data matrix, cells were then clustered using a shared nearest neighbor algorithm. Cell types were classified by the Seurat label transfer process using a previously classified set of teratomas generated from wild type H1 hESCs[74] as a reference. Clusters were projected as a uniform manifold approximation and projection (UMAP) scatterplot (**Figure 3.4.b**), with cell types well distributed across all the teratomas (**Figure 3.4.c**). 17 out of 23 cell types detected in wild type teratomas were detected in the driver library teratomas.

Drivers were assigned to individual cells by amplifying paired scRNA-seq cell barcodes and library barcodes from the unfragmented scRNA-seq cDNA, enabling genotyping of each cell. Barcodes were detectably expressed in nearly 45% of cells in this round, and we hypothesize that stochastic silencing of the lentiviral cassette and potentially sparse capture during scRNA-seq may be leading to barcode association for a fraction of the captured cells. 54% of these genotyped cells expressed a single detectable barcode while the remaining genotyped cells expressed two or more barcodes (**Figure 3.4.d-f**).

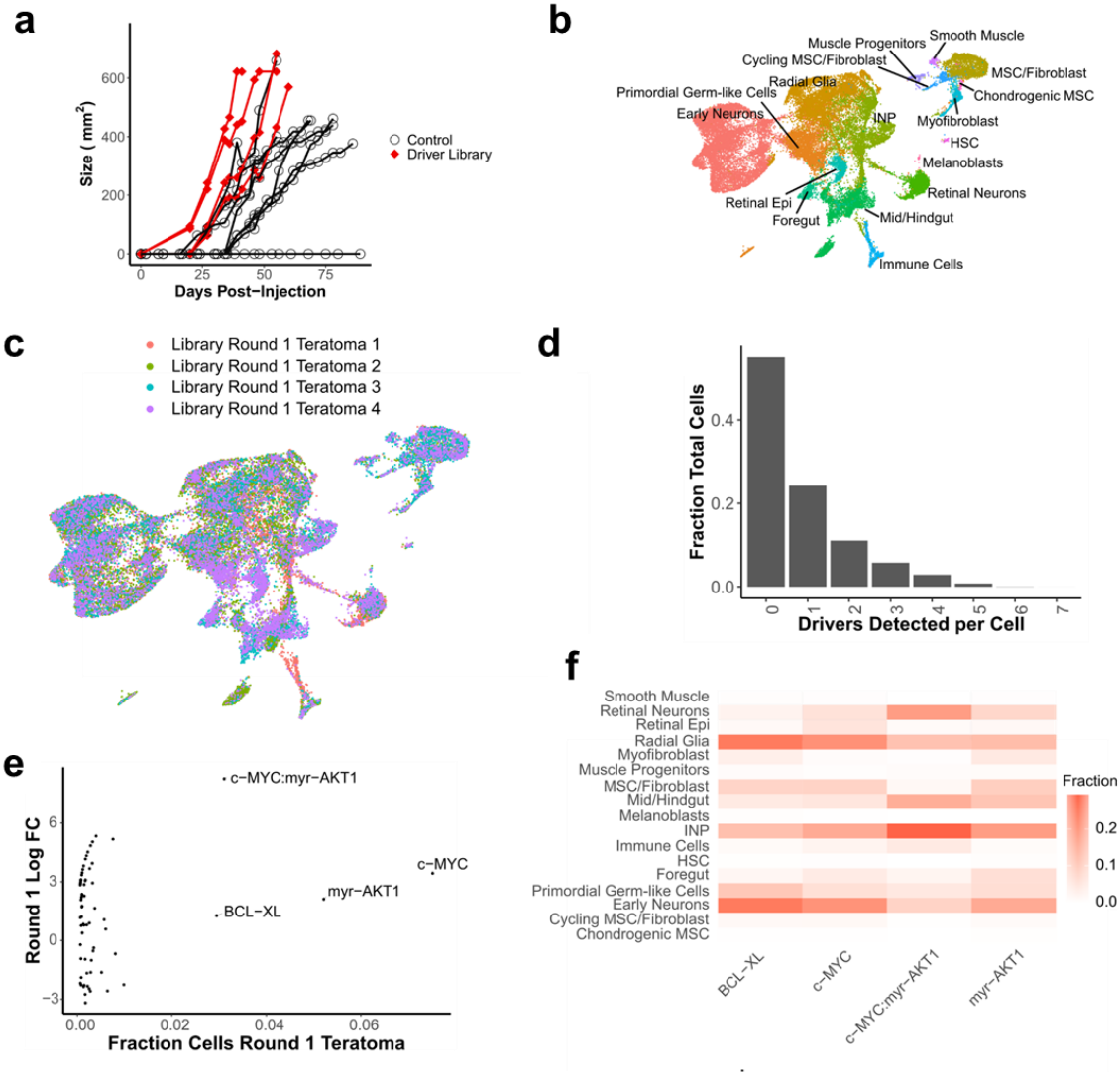


Figure 3.4: Characterization of round 1 teratomas

(**a**) Growth kinetics of round 1 teratoma formation for injections with driver library transduced hESCs vs control WT hESCs (**b**) UMAP visualization of cell types from round 1 teratomas formed by driver library transduced hESCs. (**c**) UMAP visualization of cells by sample from round 1 teratomas formed by driver library transduced hESCs. (**d**) Fraction of cells for each number of drivers detected in driver library round 1 teratomas. (**e**) Scatterplot of log fold change of drivers in round 1 teratomas over pre-injection cells, versus fraction of cells in round 1 teratomas for each driver. Only those drivers detected in at least 25 cells across all 4 round 1 teratomas are plotted. Those drivers detected in at least 1% of round 1 teratoma cells are annotated. (**f**) Heatmap of fraction of cells of each type detected for top detected drivers in round 1 teratomas.

A subset of the dissociated cells from each of the round 1 teratomas were reinjected to grow round 2 tumors, to further enrich dominant drivers and lineages. From these four reinjected tumors, two tumors were processed via scRNAseq and histology. Three of the fastest growing

wild type teratomas were also dissociated into single cell suspensions and reinjected to form round 2 control tumors.

For the two driver library tumors which were processed, tumors were palpable and measurable by 21 days after injection and reached the size threshold for excision by 35-39 days after injection. This was a markedly faster growth rate than the control round 2 tumors, which in contrast reached a detectable and measurable size at least 42 days after injection and did not reach the size threshold for excision in up to 120 days of monitoring (**Figure 3.5.a**). Due to the hardness and density of the tissue, none of the control round 2 tumors could be dissociated into single cell suspensions using the standard dissociation protocols.

Post-excision, the round 2 driver library tumors were processed in a manner similar to the round 1 teratomas. Mapping clusters to cell types in wild type teratomas revealed only 7 out of the 23 cell types detected in the wild type teratomas, a sign of potential fitness advantage in these lineages leading to them dominating the tumor composition (**Figure 3.5.b**). A significant difference was observed between the two tumor samples even after integrating the data with batch correction via the Seurat integration pipeline (**Figure 3.5.c**). This challenge in integrating scRNA-seq datasets is also observed in clinical tumor samples, which display patient- or tumor-specific batch effects[204–206] due to inter- and intra-tumor heterogeneity. Barcodes were detectably expressed in 95% of cells in this round (**Figure 3.5.d**), suggesting that cells with a survival and proliferation advantage were those where the lentiviral cassette was not silenced and was robustly expressed.

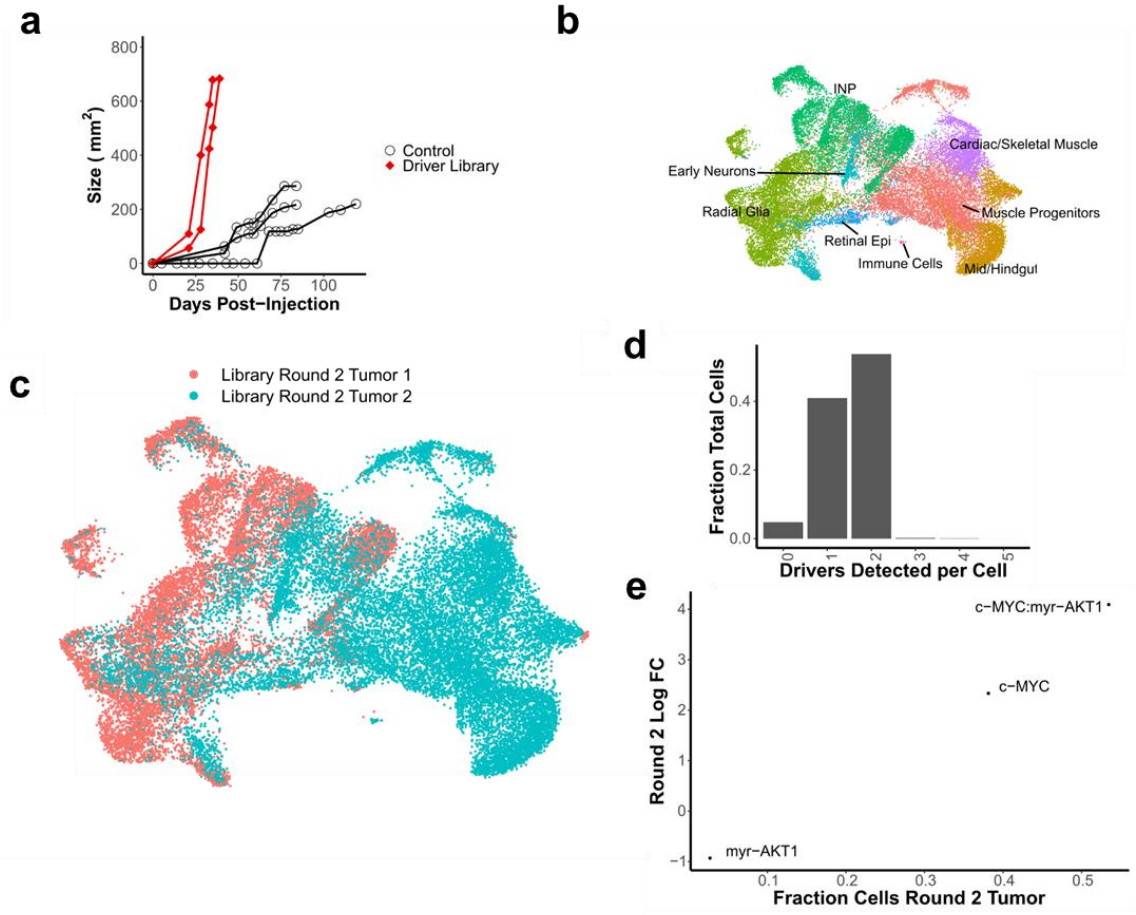


Figure 3.5: Characterization of round 2 teratomas

(**a**) Growth kinetics of round 2 tumors formed from re-injected cells from round 1 teratomas formed by driver library transduced hESCs vs WT hESCs. Control measurements are from a common set of tumors grown from the parent WT hESC line, which were used as growth controls for all experiments in this study. (**b**) UMAP visualization of cell types from round 2 tumors formed by re-injected cells from round 1 teratomas of driver library transduced hESCs. (**c**) UMAP visualization of cells by sample from round 2 tumors formed by driver library transduced hESCs. (**d**) Fraction of cells for each number of drivers detected in driver library round 2 tumors. (**e**) Scatterplot of log fold change of drivers in round 2 tumors over round 1 teratomas, versus fraction of cells in round 2 tumors for each driver. Only those drivers detected in at least 25 cells across both round 2 tumors are plotted. Those drivers detected in at least 1% of round 2 tumor cells are annotated.

We then examined the cell type populations which were present in these driver library tumors, compared to teratomas derived from wild type hESCs. In the round 1 teratomas, immature neural lineages and neural progenitors dominated the composition of the tumor with early neurons, radial glia and intermediate neural progenitors (INPs) making up 60% of the teratoma, and primordial germ-like cells another 10% (**Figure 3.6.a**). Compared to the control teratomas

this represented a 4-6 fold increase in proportion of the neural cell types, and a nearly 100 fold increase in proportion of the primordial germ-like cells. In contrast, the mesenchymal lineages were significantly reduced as a proportion of cells in the tumors (**Figure 3.6.a**). In the round 2 tumors, cell type populations were further shifted. The majority of these round 2 tumors consisted of neural progenitor-like cells, primarily radial glia and INPs, which made up 45% of these tumors, and muscle progenitor-like cells, which made up 25% of the tumors (**Figure 3.6.b**). Here we observed significant sample-specific effects with tumor 1 composed of a majority of neural-like cells, while tumor 2 was composed primarily of muscle, muscle progenitor-like and gut-like lineages (**Figure 3.5.c**).

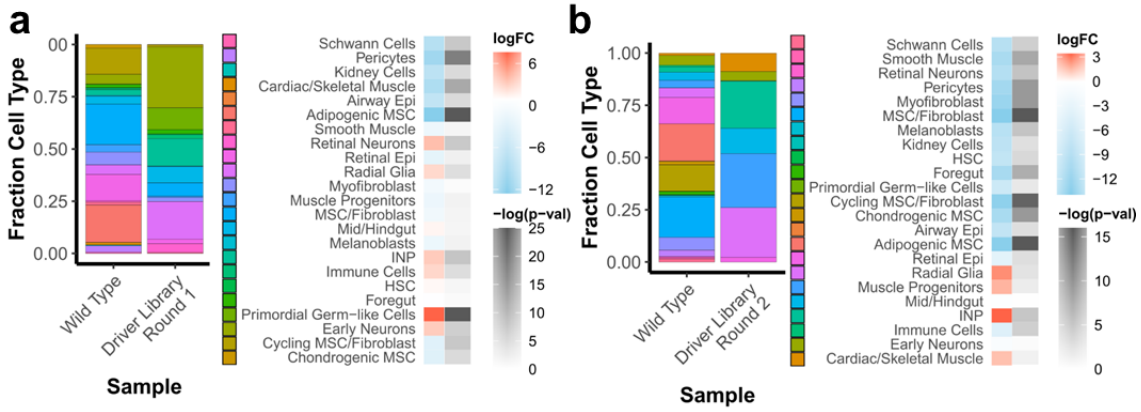


Figure 3.6: Cell type composition of driver library tumors

(**a**) Relative fraction of each cell type in round 1 teratomas formed from library transduced hESCs and WT hESCs, and log fold change of each cell type for driver library teratomas vs WT teratomas. (**b**) Relative fraction of each cell type in round 2 tumors formed from library transduced hESCs and WT hESCs, and log fold change of each cell type for driver library tumors vs WT tumors.

There was a dramatic redistribution of driver populations detected over the course of these serial injections. Prior to injection, *c-MYC* and myristoylated *AKT1* (*myr-AKT1*) constituted 1.3% and 2.4% respectively of total cells while a combinatorial transduction of the two, *c-MYC* + *myr-AKT1*, was undetectable in 5949 genotyped cells. In the first round of teratomas *c-MYC*, *myr-AKT1* and *c-MYC* + *myr-AKT1* constituted 16.7%, 11.6% and 6.9% of cells respectively, while strikingly in the second round of tumors *c-MYC* + *myr-AKT1* and *c-MYC* made up 56% and 40%

of cells respectively (**Figure 3.7.a**). These observations are consistent with the observation of elevated expression or amplification of *c-MYC*, or *c-MYC* dependent survival and proliferation, in subsets of embryonal tumors such as medulloblastoma[185,186,207,208] and atypical rhabdoid/teratoid tumors[209,210], and pediatric soft tissue tumors like rhabdomyosarcoma[211–213]. The observed enhanced combinatorial effect of the drivers is also in line with the cooperative action of *c-MYC* and *AKT1*, where the action of *c-MYC* is negatively regulated by the *FOXO* group of transcription factors, which in turn are regulated by the *PI3K/AKT* pathway[214]. A constitutively active form of *AKT1*, such as *myr-AKT1*, phosphorylates *FOXO* transcription factors and abrogates their function, thus allowing for uninhibited *c-MYC* activity[186,214]. In our observations with the teratoma based system, *c-MYC* seemed to also drive a muscle-like phenotype which was reduced as a fraction of the population when *myr-AKT1* was expressed in combination. This may also explain the sample-specific clustering observed between the two tumors which had different drivers as the top enriched hits leading to fitness advantages (**Figure 3.7.b-c**).

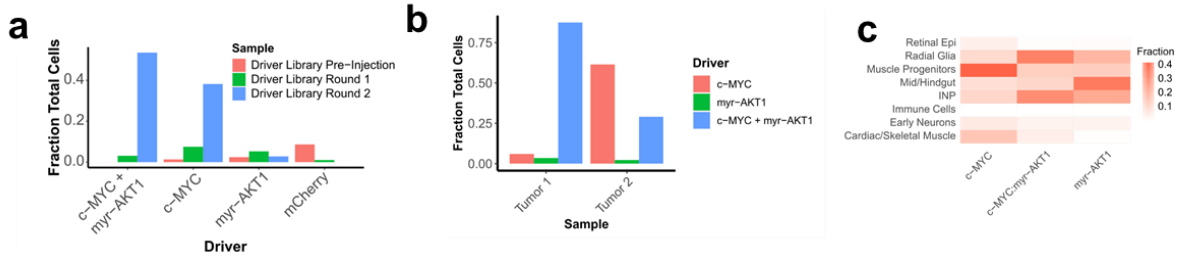


Figure 3.7: Driver enrichment in driver library tumors

(**a**) Relative fraction of top enriched drivers prior to injection and in each round of tumor formation. (**b**) Top enriched hits for each round 2 tumor sample. (**c**) Heatmap of fraction of cells of each type detected for top detected drivers in round 2 tumors.

Histology displayed clear differences between the control and driver library tumors. While control tumors had a diversity of cell types from all three germ layers in well-organized architectures (**Figure 3.8**), the driver library tumors had far more homogenous composition and displayed distinct signs of malignancy such as nuclear pleiomorphism, areas of high mitotic rate and areas of necrosis (**Figure 3.8**), suggesting the de novo creation of a transformed phenotype

driven by *c-MYC* overexpression with or without the accompanying dysregulation of *AKT1* signaling.

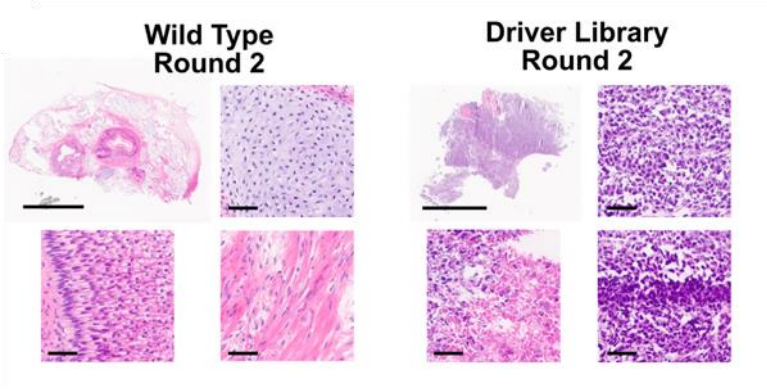


Figure 3.8: Histology of driver library tumors

H&E stained sections of round 2 tumors formed from WT and driver library transduced hESCs. WT tumors display mature cell types from all three germ layers, such as cartilage (top right), muscle (bottom right) and dermis-like epithelium (bottom left). Driver library tumors display disorganized and more homogenous composition along with markers of transformation such as nuclear pleiomorphism (top right), areas of high mitotic rates (bottom right) and areas of necrosis (bottom left). Scale bars for full section are 5 mm, scale bars for magnified images are 50 μm.

3.4.3 Screening less-dominant drivers by removing dominant hits

In the previously described screens, the proliferative advantage conferred by the top hits caused cells expressing those drivers to overwhelm all other cells, such that other drivers were not significantly detected in the second round of tumors formed by serial reinjection. To assay the fitness effects of other drivers, we hypothesized that removing the top hits from the library would allow the detection of enrichment of other drivers.

Towards this we repackaged a lentiviral driver sub-library, removing *c-MYC* and *myr-AKT1* from the pool. Using this sub-library we conducted the screening process similarly to that for the full driver library. We injected and monitored six round 1 teratomas, out of which we characterized the three fastest growing ones via scRNA-seq. Similar to before, round 1 teratomas were

reinjected for driver enrichment, and tumors were monitored for 75 days, with the two fastest growing tumors assayed via scRNA-seq.

The round 1 driver sub-library teratomas were measurable between 18-25 days after injection and reached a size sufficient for extraction between 40-60 days after injection (**Figure 3.9.a**). We again assessed the cell type populations in these sub-library screens by mapping the cells to the wild type teratomas as a reference. Using this, we determined that the round 1 teratomas contained 13 major cell types, a majority of which were neural cell types, in particular early neurons, radial glia and INPs, distributed across all three samples (**Figure 3.9.b-c**). In these samples, barcodes were detectable in 27% of cells (**Figure 3.9.d**).

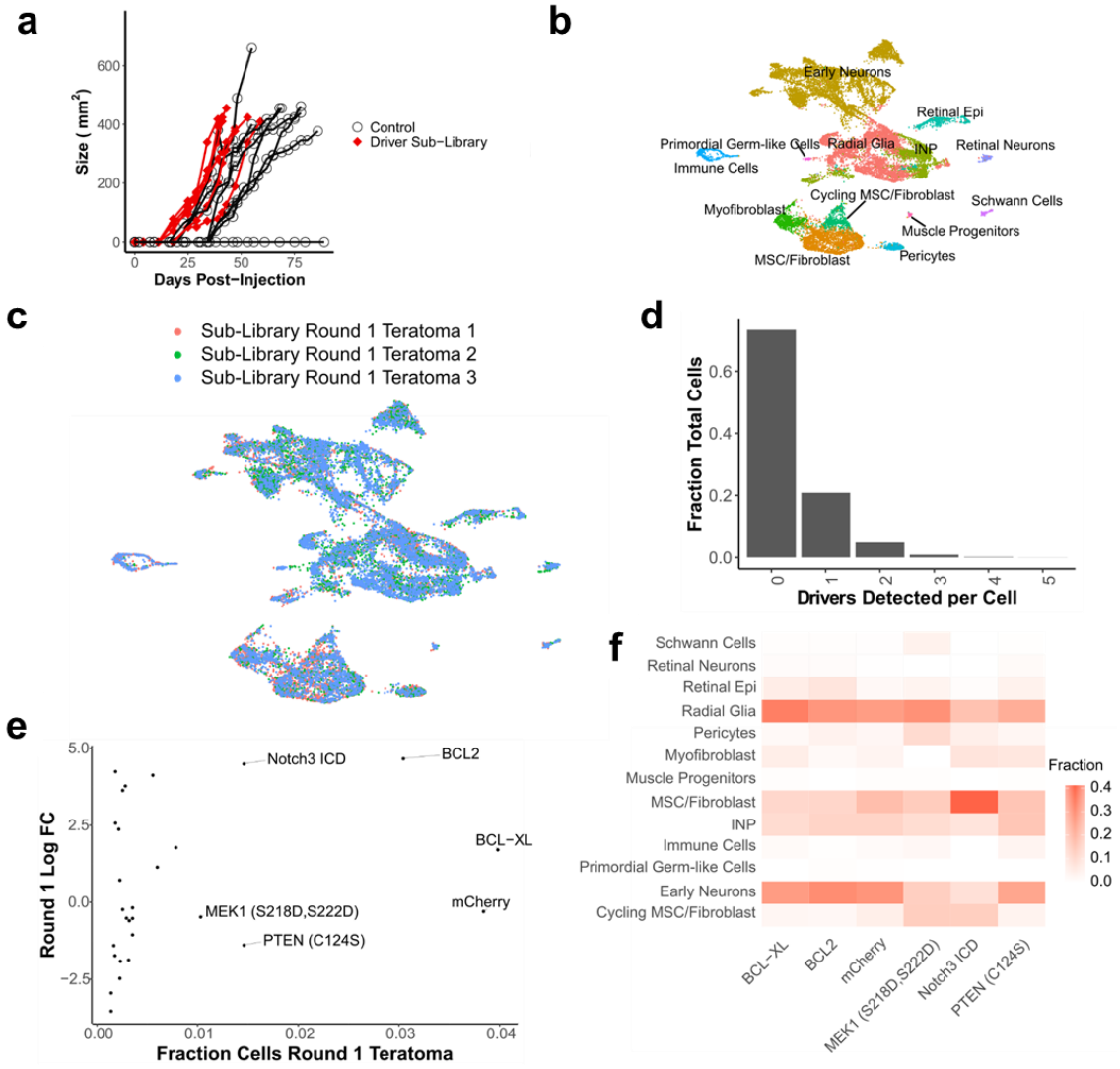


Figure 3.9: Characterization of driver sub-library round 1 teratomas

(**a**) Growth kinetics of round 1 teratoma formation for injections with driver sub-library transduced hESCs vs WT hESCs (**b**) UMAP visualization of cell types from round 1 teratomas formed by driver sub-library transduced hESCs. (**c**) UMAP visualization of cells by sample from round 1 teratomas formed by driver sub-library transduced hESCs. (**d**) Fraction of cells for each number of drivers detected in driver sub-library round 1 teratomas. (**e**) Scatterplot of log fold change of drivers in round 1 sub-library teratomas over pre-injection cells, versus fraction of cells in round 1 teratomas for each driver. Only those drivers detected in at least 25 cells across all 3 round 1 sub-library teratomas are plotted. Those drivers detected in at least 1% of round 1 teratoma cells are annotated. (**f**) Heatmap of fraction of cells of each type detected for top detected drivers in round 1 teratomas.

In contrast to the full driver library, the round 2 driver sub-library teratomas showed a slower growth rate, reaching a measurable size between 27-38 days after injection. While one of the round 2 tumors reached a sufficient size for extraction at 75 days after injection (**Figure**

3.10.a), with the fastest growing tumors extracted and assayed via scRNA-seq. In these round 2 sub-library tumors, we again detected 13 major cell types, but as opposed to neural or muscle progenitor lineages, the majority were mesenchymal cell types with adipogenic mesenchymal stromal cells (MSCs), MSC/fibroblasts and chondrogenic MSCs (**Figure 3.10.b-c**). Barcodes were detectable in 46% of cells in this round of tumors (**Figure 3.10.d**), with the increase in the fraction of genotyped cells suggesting that surviving cells which express barcodes may have a fitness or survival advantage conferred by the expressed driver. We again subsetted the drivers to visualize only those detected in 25 cells or more (**Figure 3.10.e-f**), the majority of which were driven wholly or in part by *MEK1*$^{S218D/S222D}$, while the remaining were driven by *RHOJ*, a small GTP-binding protein known to play a role in cell migration, which has recently been found to confer proliferative advantage in multiple lineages[175].

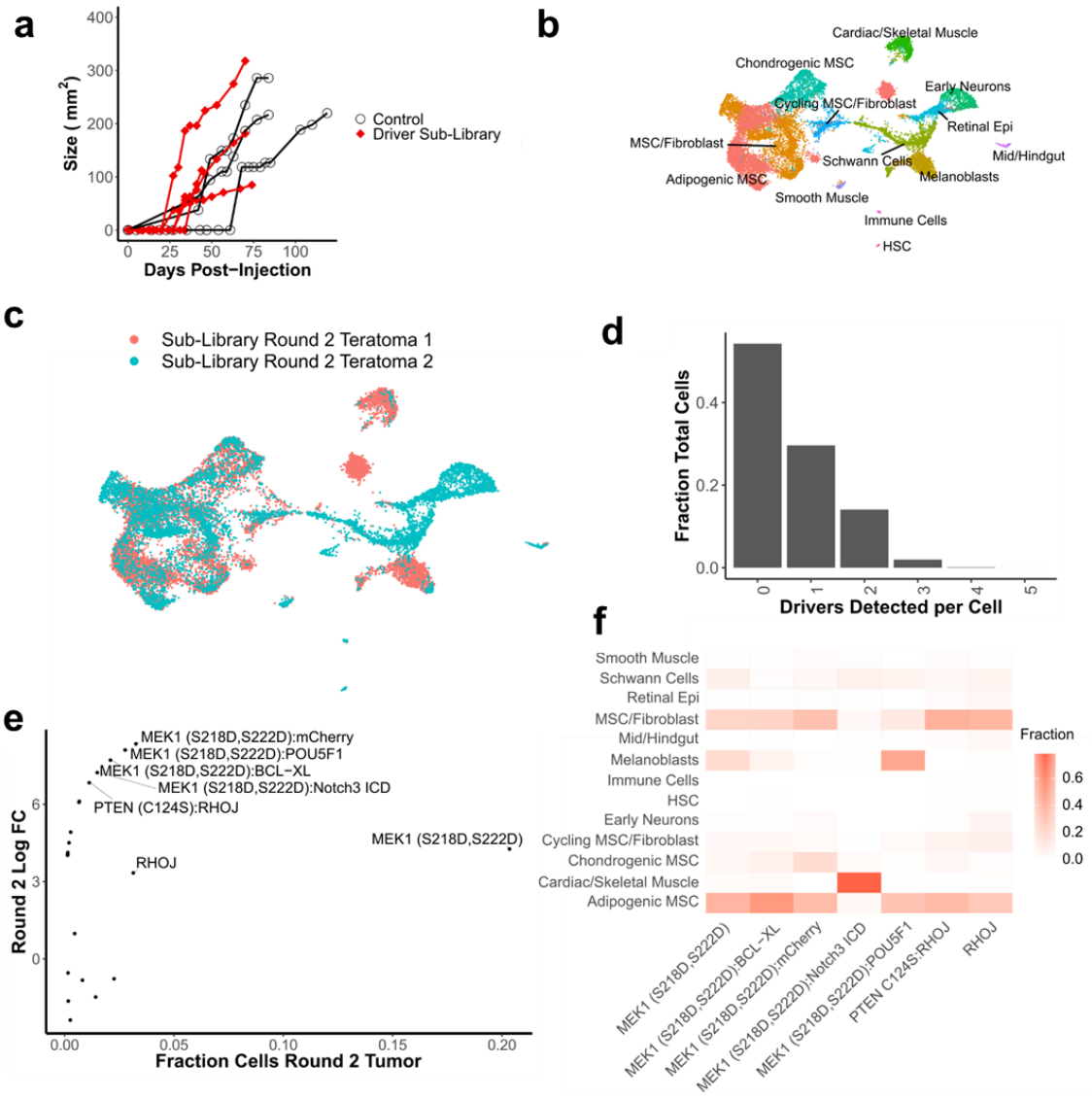


Figure 3.10: Growth kinetics of round 2 sub-library tumors

(**a**) Growth kinetics of round 2 tumors formed from re-injected cells from round 1 teratomas formed by driver library transduced hESCs vs WT hESCs. Control measurements are from the common set of tumors grown from the parent WT hESC line, which were used as growth controls for all experiments in this study. (**b**) UMAP visualization of cell types from round 2 tumors formed by re-injected cells from round 1 teratomas of driver library transduced hESCs. (**c**) UMAP visualization of cells by sample from round 2 teratomas formed by driver sub-library transduced hESCs. (**d**) Fraction of cells for each number of drivers detected in driver sub-library round 2 tumors. (**e**) Scatterplot of log fold change of drivers in round 2 tumors over round 1 teratomas, versus fraction of cells in round 2 tumors for each driver. Only those drivers detected in at least 25 cells across both round 2 tumors are plotted. Those drivers detected in at least 1% of round 2 tumor cells are annotated. (**f**) Heatmap of fraction of cells of each type detected for top detected drivers in round 2 tumors.

We then evaluated the cell type populations in comparison to teratomas formed from wild type H1 hESCs. In the round 1 sub-library teratomas, immature and progenitor neural cell types

constituted 70% of the total population of the teratoma, and fibroblast-like mesenchymal lineages making up a further 20% of the teratoma (**Figure 3.11.a**). However, in the round 2 tumors, in a striking difference compared to the full library screens, the tumors were composed primarily of fibroblast and MSC phenotypes, which constituted 65% of the tumors (**Figure 3.11.b**). Along with the fibroblast lineages which constituted a majority of these tumors, a small neural component persisted via an expansion in Schwann cells and melanoblasts, which may be derived from Schwann cell precursors[215]. The neural progenitor-like lineages present in the *c-MYC* and *myr-AKT1* driven tumors were not present in these tumors. This change in composition of the tumors was accompanied by an enrichment of the constitutively active mutant $MEK1^{S218D/S222D}$ which was present in 3.3% of cells prior to injection but constituted 20% of all cells in the round 2 tumors. In comparison, the fraction of cells expressing the internal negative control, mCherry, fell from 16% of cells prior to injection to 2% of cells in the round 2 tumors (**Figure 3.11.c**). This role of $MEK1^{S218D/S222D}$ in supporting proliferation and survival of fibroblasts is consistent with previously reported results in literature where expression of constitutively active versions of *MEK1* were sufficient to trigger proliferative states and even transformation in fibroblasts *in vitro*[216–218].

Histology images from the two rounds of tumors further confirmed the cellular composition. In round 1 teratomas we observed a majority of neuroectoderm-like cell types, while in round 2 tumors the majority of cells were mesenchymal in nature (**Figure 3.11.d**). In these sub-library screens, we did not observe clear indications of malignancy as we observed in the *c-MYC* and *myr-AKT1* driven full library screens, suggesting that while $MEK1^{S218D/S222D}$ may drive proliferative advantage, it was a less potent driver of cellular transformation on its own.

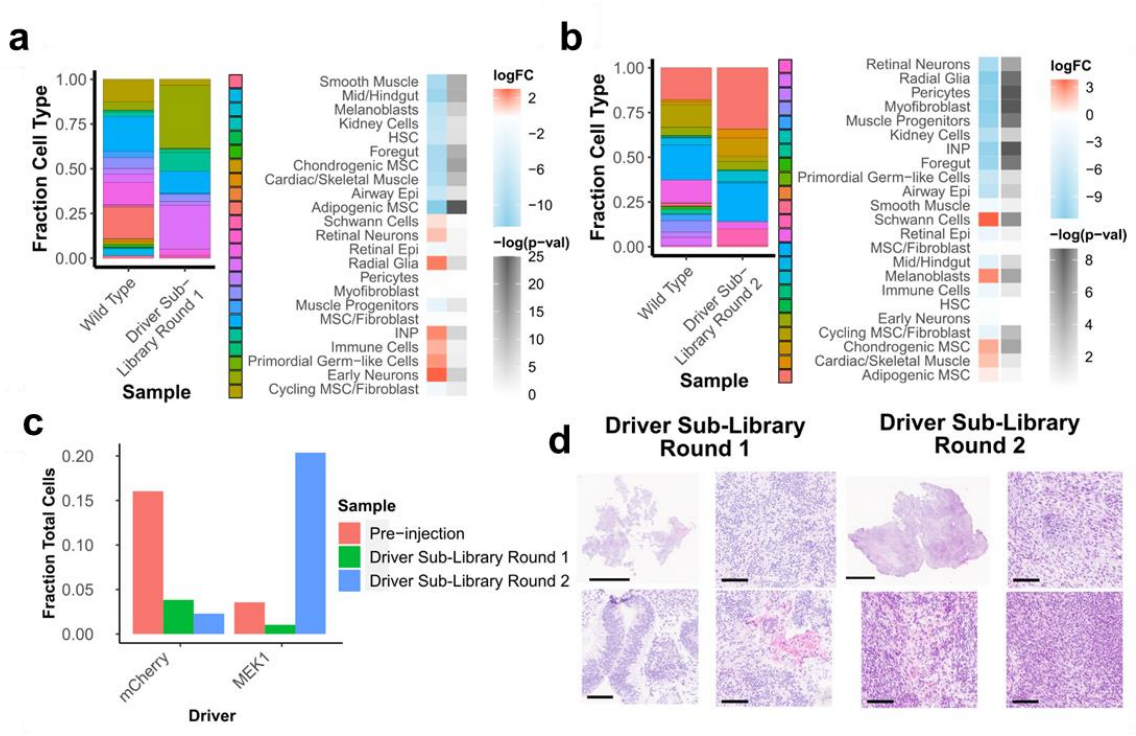


Figure 3.11: Cell type composition and driver enrichment in sub-library screen tumors

(a) Relative fraction of each cell type in round 1 teratomas formed from library transduced hESCs and WT hESCs, and log fold change of each cell type for driver library teratomas vs WT teratomas. **(b**) Relative fraction of each cell type in round 2 tumors formed from library transduced hESCs and WT hESCs, and log fold change of each cell type for driver library tumors vs WT tumors. **(c**) Relative fraction of top enriched drivers prior to injection and in each round of tumor formation. **(d**) H&E stained sections of round 1 and round 2 tumors formed from hESCs transduced with driver libraries without *c-MYC* or *myr-AKT1*. Round 1 tumors contain neuroectodermal and epithelial cell types as the majority of cells, while round 2 tumors contain mesenchymal cell types as the majority of cells. Scale bars for full section are 5 mm, scale bars for magnified images are 100 μm.

3.4.4 Validating enriched driver hits

To validate the hits obtained from the two sets of screens we individually tested the effects of *c-MYC*, *myr-AKT1*, *c-MYC* + *myr-AKT1* and *MEK1*$^{S218D/S222D}$. hPSCs were transduced with either one of these driver vectors or a negative control vector expressing mCherry. Prior to injection the driver and control transduced cells were mixed in a 1:1 ratio, with a portion of these mixed cells pelleted and genomic DNA extracted to assess barcode distribution, and the remaining cells injected for teratoma formation in three Rag2$^{-/-}$;γc$^{-/-}$ immunodeficient mice for each driver to be validated. Teratomas were formed from these injected cells, excised when they

reached sufficient size, representative pieces of the excised tumors were immediately flash frozen in liquid nitrogen and some representative pieces preserved for cryosectioning, while the remaining tissue was dissociated. Dissociated cells were divided to be stored for genomic DNA extraction, RNA extraction and for serial reinjection to form round 2 tumors. Round 2 tumors were also then allowed to grow to a size sufficient for extraction and then excised and dissociated as for round 1. Dissociated cells were stored for genomic DNA extraction and RNA extraction.

Consistent with the observations during the screens, teratomas formed by hESCs transduced with *c-MYC* and *c-MYC* + *myr-AKT1* grew at the highest rate compared to all other tumors. In the round 1 tumors, the *c-MYC*, *myr-AKT1* and *c-MYC* + *myr-AKT1* were measurable between 27-32 days after injection and were at an extractable size between 42-48 days after injection. While for *MEK1*$^{S218D/S222D}$ round 1 teratomas, tumors were measurable 33-40 days after injection and at an extractable size 59-78 days after injection (**Figure 3.12.a**). In the round 2 tumors, those driven by *c-MYC* + *myr-AKT1* grew at the fastest rate, followed by *c-MYC*, both of which showed a significantly enhanced growth rate compared to the control round 2 tumors. Round 2 tumors driven by *myr-AKT1* alone or those driven by *MEK1*$^{S218D/S222D}$ did not show a significantly enhanced growth rate compared to the control tumors (**Figure 3.12.a**), although tumors driven by *MEK1*$^{S218D/S222D}$ clustered toward the higher end of the control tumor growth rate.

To assess the barcode distribution from cells prior to injection through both rounds of tumor formation, we PCR amplified barcodes from genomic DNA isolated from stored cell pellets and quantified them via deep sequencing. We find that negative control barcodes decrease from 21-35% of total mapped reads prior to injection, to less than 2% of reads for *MEK1*$^{S218D/S222D}$, and less than 0.1% of reads for all other drivers by the second round of tumors (**Figure 3.12.b**).

We then performed bulk RNA-sequencing on these validation tumors to assess their composition. In agreement with the results of the screens, we found that the *c-MYC* driven tumors

were composed of a mix of neural and muscle lineages displaying elevated expression of neural-related genes *TUBB3*, *NEUROD1*, *SYP*, *CAMKV* and *NEFH* (**Figure 3.12.c**). In comparison, tumors driven by a combination of *c-MYC* and *myr-AKT1* were primarily composed of a neural progenitor like phenotype, expressing *NEFH* and *CAMKV* and elevated levels of *DANCR*, a non-coding RNA which acts to suppress differentiation and is present in many cancers[219,220] (**Figure 3.12.c**). On the other hand, tumors driven by *MEK1*$^{S218D/S222D}$ displayed elevated expression of mesenchymal markers *VIM*, *ITM2A* and *CD44* (**Figure 3.12.c**), which is again consistent with the results from the sub-library screens.

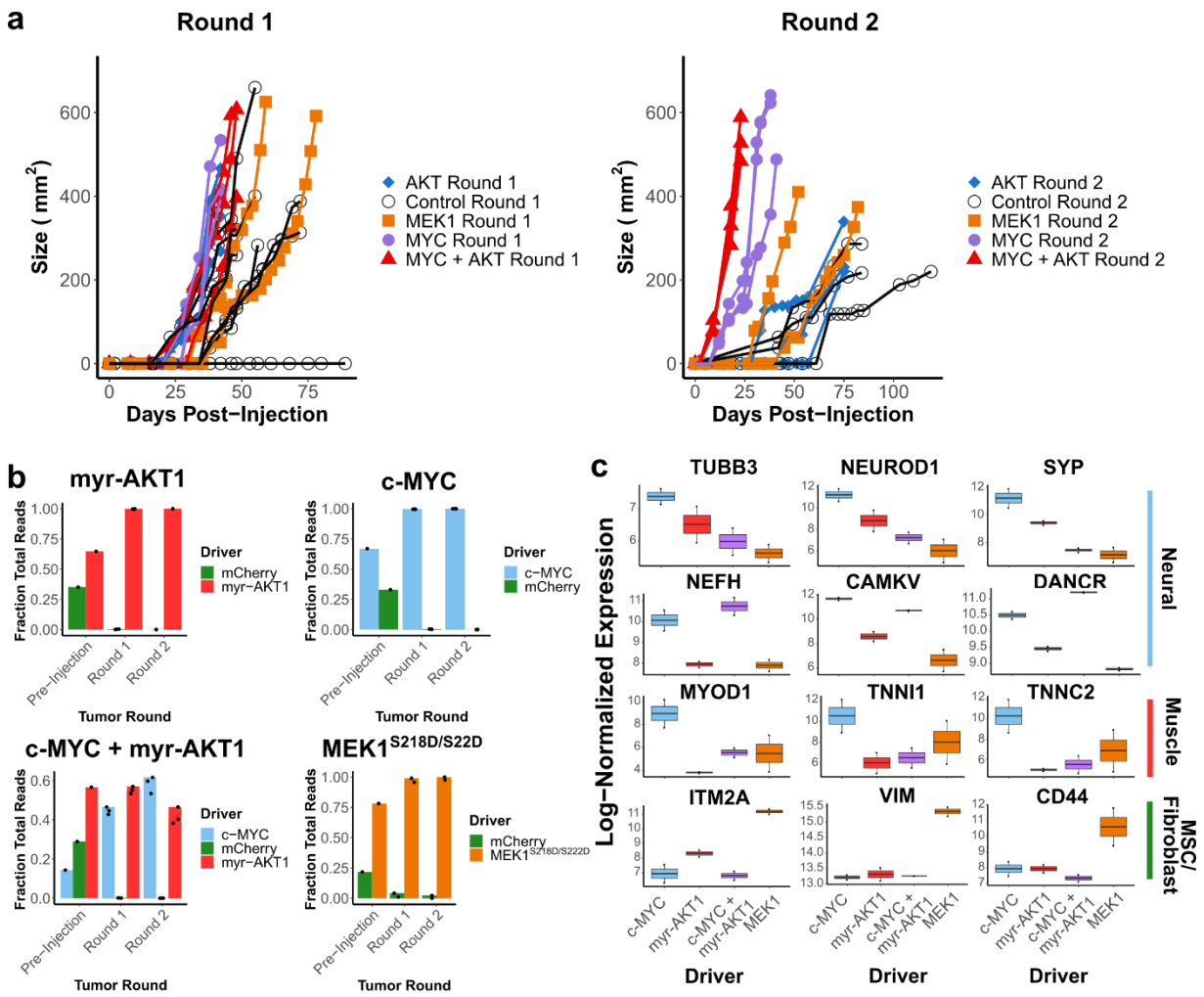


Figure 3.12: Characterization of validation tumors

(a) Growth kinetics of round 1 and round 2 tumors formed from a mixture of hESCs transduced with driver hits (c-MYC, myr-AKT1, c-MYC + myr-AKT1 or MEK1 (S218D, S222D) and hESCs transduced with a negative control (mCherry). Control measurements are from the common set of tumors grown from the parent WT hESC line, which were used as growth controls for all experiments in this study. **(b)** Fraction of reads detecting either the driver or negative control barcodes at each stage: pre-injection, round 1 tumor formation and round 2 tumor formation. Barcodes were amplified from genomic DNA. **(c)** Gene expression of lineage-specific markers for round 2 tumors driven by individual hits. Expression values are normalized and log transformed.

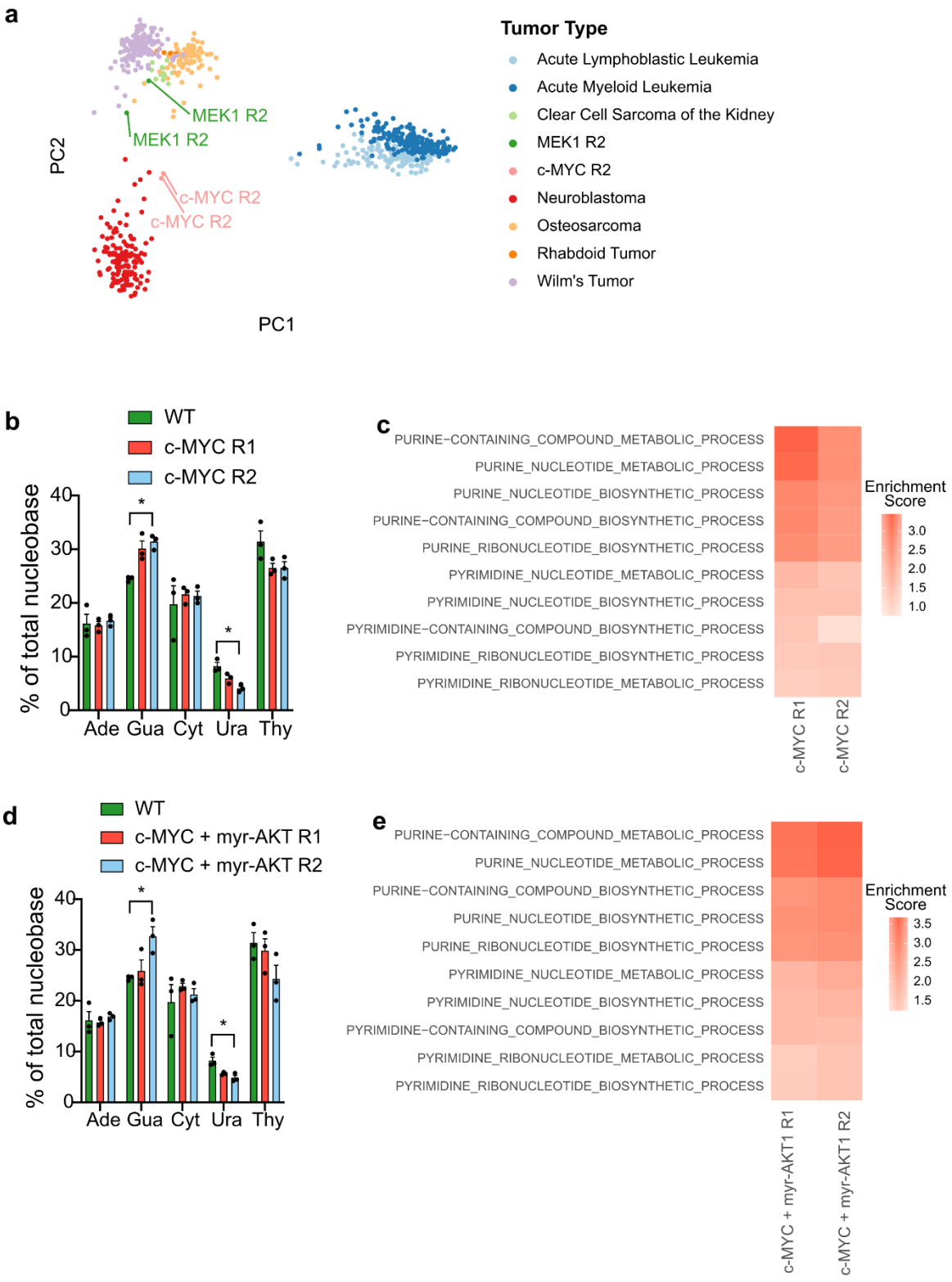
We also compared the *c-MYC* and *MEK1*$^{S218D/S222D}$ driven tumors to pediatric cancers from the TARGET initiative. Using the 2000 most variable genes across the TARGET tumors we performed a principal component analysis (PCA) , (**Figure 3.13.a**). Plotting the two PCs capturing the majority of variation, we observed that *c-MYC* driven tumors clustered toward the TARGET

neuroblastoma tumors, while those driven by *MEK1*$^{S218D/S222D}$ clustered toward the kidney tumors (Wilm's tumor and clear cell sarcoma of the kidney) and osteosarcoma.

c-MYC and *myr-AKT1* also contribute to metabolic reprogramming in tumors by promoting nucleotide biosynthesis[221,222]. To further characterize *c-MYC* and *c-MYC* + *myr-AKT1* driven tumors, we quantified nucleobase levels via mass spectrometry. Compared to WT teratomas, *c-MYC* and *c-MYC* + *myr-AKT1* driven tumors increased the abundance of purine nucleobases, especially guanine (**Figure 3.13.b-c**). Further, this was supported by broad enrichment of nucleobase synthesis terms, and particularly purine synthesis related Gene Ontology terms in genes which were upregulated compared to WT teratomas (**Figure 3.13.d-e**). This is consistent with previous studies which have demonstrated the regulation of nucleotide metabolism by *c-MYC*[223,224], with de novo purine synthesis especially implicated in tumor maintenance[225] and response to therapy[226].

Figure 3.13: Correlation to clinical tumors and metabolic effects of *c-MYC* and *c-MYC + myr-AKT1*

(a) PC1 vs PC2 plot of *c-MYC* and *MEK1*$^{S218/S222D}$ driven tumors with pediatric tumors from the TARGET initiative. (**b**) Nucleobase abundance in Control (Wild Type) teratomas, *c-MYC* driven Round 1 tumors and *c-MYC* driven Round 2 tumors. (**c**) Heatmap of enrichment scores for Gene Ontology genesets related to purine and pyrimidine nucleotide metabolism and synthesis for *c-MYC* driven Round 1 and Round 2 tumors compared to wild type teratomas. (**d**) Nucleobase abundance in Control (Wild Type) teratomas, *c-MYC* + *myr-AKT1* driven Round 1 tumors and *c-MYC* + *myr-AKT1* driven Round 2 tumors. (**e**) Heatmap of enrichment scores for Gene Ontology genesets related to purine and pyrimidine nucleotide metabolism and synthesis for *c-MYC* + *myr-AKT1* driven Round 1 and Round 2 tumors compared to wild type teratomas.


a
Tumor Type
Acute Lymphoblastic Leukemia
Acute Myeloid Leukemia
Clear Cell Sarcoma of the Kidney
MEK1 R2
c-MYC R2
Neuroblastoma
Osteosarcoma
Rhabdoid Tumor
Wilm's Tumor
MEK1 R2
MEK1 R2
c-MYC R2
c-MYC R2
PC2
PC1
b
WT
c-MYC R1
c-MYC R2
% of total nucleobase
Ade Gua Cyt Ura Thy
c
PURINE−CONTAINING_COMPOUND_METABOLIC_PROCESS
PURINE_NUCLEOTIDE_METABOLIC_PROCESS
PURINE_NUCLEOTIDE_BIOSYNTHETIC_PROCESS
PURINE−CONTAINING_COMPOUND_BIOSYNTHETIC_PROCESS
PURINE_RIBONUCLEOTIDE_BIOSYNTHETIC_PROCESS
PYRIMIDINE_NUCLEOTIDE_METABOLIC_PROCESS
PYRIMIDINE_NUCLEOTIDE_BIOSYNTHETIC_PROCESS
PYRIMIDINE−CONTAINING_COMPOUND_BIOSYNTHETIC_PROCESS
PYRIMIDINE_RIBONUCLEOTIDE_BIOSYNTHETIC_PROCESS
PYRIMIDINE_RIBONUCLEOTIDE_METABOLIC_PROCESS
Enrichment Score
c-MYC R1
c-MYC R2
d
WT
c-MYC + myr-AKT R1
c-MYC + myr-AKT R2
% of total nucleobase
Ade Gua Cyt Ura Thy
e
PURINE−CONTAINING_COMPOUND_METABOLIC_PROCESS
PURINE_NUCLEOTIDE_METABOLIC_PROCESS
PURINE−CONTAINING_COMPOUND_BIOSYNTHETIC_PROCESS
PURINE_NUCLEOTIDE_BIOSYNTHETIC_PROCESS
PURINE_RIBONUCLEOTIDE_BIOSYNTHETIC_PROCESS
PYRIMIDINE_NUCLEOTIDE_METABOLIC_PROCESS
PYRIMIDINE_NUCLEOTIDE_BIOSYNTHETIC_PROCESS
PYRIMIDINE−CONTAINING_COMPOUND_BIOSYNTHETIC_PROCESS
PYRIMIDINE_RIBONUCLEOTIDE_METABOLIC_PROCESS
PYRIMIDINE_RIBONUCLEOTIDE_BIOSYNTHETIC_PROCESS
Enrichment Score
c-MYC + myr-AKT1 R1
c-MYC + myr-AKT1 R2

Histology showed that round 2 tumors driven by *myr-AKT1* alone showed poorly differentiated neural lineage cells interspersed with mesenchyme (**Figure 3.14.a**). Round 2 tumors driven by *c-MYC* (**Figure 3.14.b**) and those driven by *c-MYC* + *myr-AKT1* (**Figure 3.14.c**) were composed of poorly differentiated cells with signs of malignancy, including necrosis. Finally, as observed in screen results, round 2 tumors driven by $MEK1^{S218D/S222D}$ were composed primarily of mesenchymal, fibroblast-like cells interspersed with cartilage (**Figure 3.14.d**). These validation results strongly support the observations from the multiplexed screens and confirm the significant fitness advantage conferred by *c-MYC* and *c-MYC* + *myr-AKT1* on neural lineages, and of $MEK1^{S218D/S222D}$ on fibroblasts.

In addition, staining with the proliferation marker Ki-67 demonstrated a distinctly larger population of proliferating cells in the round 2 *c-MYC* and *c-MYC* + *myr-AKT1* driven tumors followed by round 2 $MEK1^{S218D/S222D}$ driven tumors, as compared to the WT teratomas (**Figure 3.14.e**). While round 2 tumors driven by *myr-AKT1* alone showed no higher proliferative capacity than WT tumors (**Figure 3.14.e**), confirming that *myr-AKT1* conferred transforming potential only in concert with *c-MYC*. These validation results strongly support the observations from the multiplexed screens and confirm the significant fitness advantage conferred by *c-MYC* and *c-MYC* + *myr-AKT1* on neural lineages, and of $MEK1^{S218D/S222D}$ on fibroblasts.

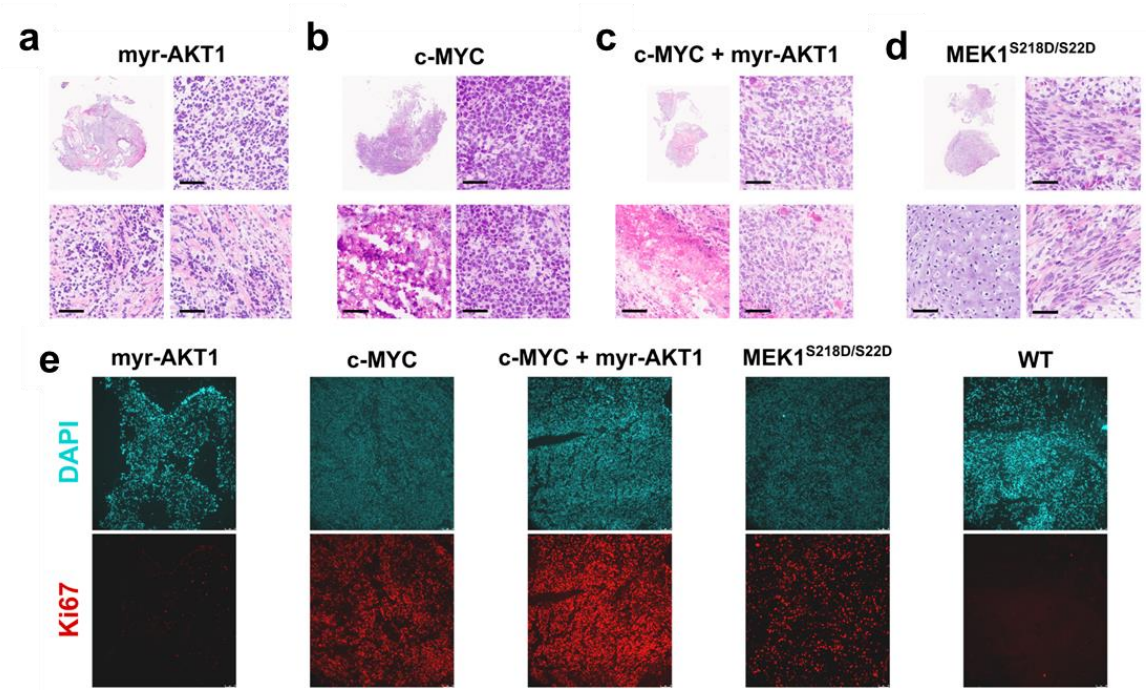


Figure 3.14: Histology of validation tumors

(a) H&E stain of round 2 tumor driven by myr-AKT1 showing regions of poor differentiation (top right) and necrosis (bottom right) interspersed with regions of organized tissue (bottom left). (**b**) H&E stain of round 2 tumor driven by c-MYC showing regions of poor differentiation (top and bottom right) and regions of necrosis (bottom left). (**c**) H&E stain of round 2 tumor driven by c-MYC + myr-AKT1 showing regions of poor differentiation (top and bottom right) and regions of necrosis (bottom left). (**d**) H&E stain of round 2 tumor driven by MEK1 (S218D, S222D) showing primarily regions of mature mesenchymal fibroblast-like tissue (top and bottom right) and cartilage-like (bottom left). Scale bars for magnified images are 50 μm. (**e**) Immunofluorescence micrographs of DAPI and Ki-67 stained sections. Scale bars are 75 μm.

3.5 Discussion

To study the important problem of what drives oncogenic transformation in human tissue, investigators have relied on *in vitro* and animal model systems. While immense progress has been made using these, limitations still remain. Animal models retain significant differences with human biology, while *in vitro* systems lack vasculature and the physiological cues present in the *in vivo* niche which are involved in regulating lineage-specific transformation. Additionally, currently prevalent models often necessitate the investigation of a single or a few lineages at a time, raising an impediment in studying tissue-specificity across multiple cell types.

In this study, we have developed a novel platform to study the effects of cancer drivers across lineages which harnesses hPSC-derived teratomas to access a diverse set of lineages, ORF overexpression libraries to express cancer drivers, scRNA-seq to read out transcriptomic profiles which determine the cell type and detect the perturbing driver, and serial proliferation of the tumors to enrich drivers and lineages enjoying fitness advantages. Using this platform, we initially screened key cancer drivers across more than 20 cell types and de novo generated a transformed phenotype in a neural lineage via the overexpression of *c-MYC* and *myr-AKT1*, a constitutively active form of *AKT1*, which dominated the serially reinjected tumors and displayed the hallmarks of malignancy. To screen less-dominant drivers, we repeated the screens with *c-MYC* and *myr-AKT1* removed from the library, and captured the proliferative advantage of other drivers, such as the one conferred on mesenchymal lineages like fibroblasts via the overexpression of constitutively active $MEK1^{S218D/S222D}$. These results were then validated by individually overexpressing these top hits during teratoma formation and serial reinjection, to confirm the enrichment of lineages detected in the screens.

While offering a powerful new approach in studying transformation, certain limitations and challenges merit consideration both in terms of interpretation of the resulting data, as well as in design of future implementations of the same. Firstly, in its current form, the overexpression vector is designed to be constitutively on. This may confound results since some drivers and lineages could be enriched due to their biased differentiation or inability to escape a state being governed by driver expression itself. In future versions of this platform, that effect may be mitigated by using an inducible expression system or one with recombinase-based control, both of which can be controllably turned on at specific time points, to ensure differentiation is not affected by driver expression. Additionally, background mutations and genetic alterations already present in hPSC lines may bias transformation phenotypes. Secondly, the predominantly embryonic versus adult state of the cell types might bias their innate transformation potential. In particular, we anticipate,

due to the embryonic origin of the starting cells, this system might be especially applicable to modeling of pediatric tumors. Thirdly, the differentiation of teratomas is a partially random stochastic process, which leads to an inherent heterogeneity in the cell types available in each sample. To overcome this, cell fate engineering methods may be paired with the screening platform to repeatably and predictably derive specific lineages of interest. Fourthly, while largely less discussed, dissociation and sample processing methods have a large impact on scRNA-seq results[227–229]. Improved protocols may allow for more fine-grained dissection of transcriptomes to assess cell state shifts within lineages. Fifth, the xenografted mouse models used for teratoma formation may be improved upon in two ways. In this study, we used a sub-cutaneous site of injection for teratoma formation, but the site of injection impacts teratoma differentiation[75] as well as the lineage-specificity of cancer drivers. Orthotopic injections for specific tissues may provide a more appropriate niche. In addition, the mice used in this study are immune deficient, thus precluding the study of any effects of immune system interaction. A possible route to addressing this may be through the use of humanized mouse models[230]. Sixth, while in this demonstration we use scRNA-seq to determine cell type and state, epigenetic factors such as chromatin state are critical to tumorigenesis, cancer evolution and progression. Combining techniques to map epigenetic characteristics, such as ATAC-seq, may lead to a more detailed understanding of the determinants of tumor formation. Seventh, our pooled screening method may lead to transcriptomic and cell state shifts due to cell-cell communication and paracrine signaling. These challenges may potentially be addressed by utilizing newly developed spatial transcriptomics[231–233] methods to tease apart cell endogenous versus exogenous effects. And finally, we have focused here primarily on deciphering the role of individual cancer drivers on oncogenic potential across lineages. Exploring combinatorial perturbations, especially in the background of tumor suppressor mutations will be crucial to dissecting both tumorigenicity as observed in native tumors, and also systematically studying variants-of-unknown significance, many of which individually may have only subtle phenotypes.

Taken together, we have demonstrated a proof of concept for a scalable, versatile platform which can screen multiple lineages and drivers in a single experiment, with a rich transcriptomic readout, thus providing a systematic path to studying the determinants and tissue-specificity of neoplastic transformation in human cells. We envision that refinements to this platform, coupled with the expanding array of available omics technologies will enable the comprehensive characterization of the trajectory of cells from normal to malignant states.

3.6 Acknowledgements

This work was generously supported by UCSD Institutional Funds and NIH grants (R01HG009285, RO1CA222826, RO1GM123313). This publication includes data generated at the UC San Diego IGM Genomics Center utilizing an Illumina NovaSeq 6000 that was purchased with funding from a National Institutes of Health SIG grant (#S10 OD026929).

Chapter 3 in part is a reprint of the material Parekh, U., McDonald, D., Dailamy, A., Wu, Y., Cordes, T., Zhang, K., Tipps, A., Metallo, C., Mali, P. Charting oncogenicity of genes and variants across lineages via multiplexed screens in teratomas, currently under peer review. The dissertation author is the primary author of the study.

Chapter 4: Programmatic Introduction of Parenchymal Cell Types in Blood Vessel Organoids

4.1 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 (**Figure 4.1**). 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.

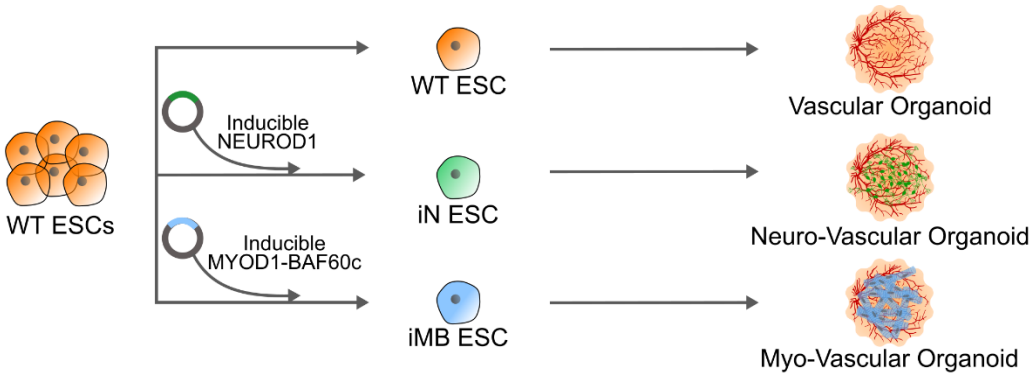


Figure 4.1: Schematic of parenchymal cell introduction in vascular organoids (VOs)

General strategy for the generation of vascularized organ tissues via introduction of parenchymal cell types in VOs.

4.2 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[48,53]. These advances have also enabled substantial progress toward building fully PSC-derived, functional human organs *in vitro*[53,234]. However, the absence of vasculature in most organoid derivation techniques limits their utility, principally in two ways. Firstly, it is widely accepted that vasculature plays a crucial role in development and disease[60,61]. Thus, avascular organoids are insufficient mimics to recapitulate organogenesis and disease. Secondly, vasculature is necessary to prevent necrosis in tissues that grow beyond 1 mm in size[62,63], which deems vasculature critical for building large-scale organoids and tissue models.

To address this, several groups have demonstrated progress on developing methods for vascularizing organoids[66,67,70,235–238]. Some groups have succeeded in vascularizing organoids

after transplanting them *in vivo*[66,67], but requiring an *in vivo* host limits experimental control, increases cost, and diminishes its potential to be used in clinical applications due to the chimeric nature of the tissue. Other groups have reported enhanced vascularization of kidney organoids[236–238], but these methods are not globally applicable to all organoid platforms, since they only enhance the vasculature that spontaneously emerges in regularly derived kidney organoids[239,240], rather than engineer a vasculature network *de novo*. Finally, another approach has demonstrated that transient *GATA6* overexpression can introduce a nascent vascular network in liver-bud organoids[235], however there is no current evidence that this system can be translated to other organoid systems.

Organoid platforms leverage knowledge of embryonic and organ development to provide temporally appropriate chemical cues, like growth factors or small molecule inhibitors, 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 spatial architectures. Thus, cells which do not belong to those specific organ compartments or arise from different germ layers are absent from the final organoid, e. g. vascular cell types which arise from the mesoderm are absent from cerebral organoids whose cell types arise from the neuroectoderm.

To introduce cell types outside of those available from directed differentiation, a promising strategy is to combine genetic overexpression techniques, which can override chemical cues, with directed differentiation. Recently, methods to vascularize cerebral organoids have been reported, in which PSCs were engineered to ectopically express human ETS variant 2 (*ETV2*)[70], a known driver of differentiation to endothelial cells from PSCs described by others and we[138,188,241,242]. These *ETV2*-expressing cells were reported to generate a vascular-like network in cerebral organoids. Although this method has exciting potential, it suffers from two limitations: one, *ETV2* overexpression induces only a low degree of vascularization in the organoids, likely due to it being

a weak inducer of vascularization in non-conducive media and niche settings; and two, it does not induce the full panoply of vascular lineages, such as smooth muscle cells, pericytes, and mesenchymal stem cells, all of which are critical for complete blood vessel development and function[243,244]. Thus, there is a need to explore alternative methods for vascularizing organoids.

Here, we overlay reprogramming methods with the directed differentiation of vascular organoids (VOs)[245] to derive cross-germ layer and cross-lineage organoids with complete vascular networks. Specifically, a recently described elegant VO differentiation approach[245] enables reproducible differentiation yielding complete blood vessel networks which include smooth muscle cells, mesenchymal stem cells, and endothelial cells, but these organoids lack organ specific parenchymal cell types limiting their utility for broader disease modelling and regenerative medicine applications. Here we introduce a parenchymal cell component into VOs via transcription factor (TF) overexpression and demonstrate this approach by building inducible neuro-vascular and myo-vascular organoids (**Figure 3.1**). This is done by the induced overexpression, in optimized media conditions, of neural transcription factor *NEUROD1* (iN) in VOs for neuro-vascular organoids (iN-VOs), and the induced overexpression of myogenic transcription factor *MYOD1* plus epigenetic regulator *BAF60c* (iMB) in VOs for myo-vascular organoids (iMB-VOs).

This yields a facile method for co-differentiation of lineage specific parenchymal cells and the entire blood vessel lineage from a single PSC line. After demonstrating the cross-tissue compatibility of this method via generating iN-VOs and iMB-VOs, more comprehensive optimization and characterization was conducted for long-term cultured iN-VOs. The resulting iN-VO is composed of a dense network of complete blood vessels, and a pure population of functional excitatory neurons. Our iN-VOs are perfusable when implanted *in vivo*, generate spontaneous and complex firing patterns, and can be sustained for at least 45 days *in vitro*. Lastly, we demonstrate a proof-of-concept for maturation and large-scale tissue creation by subjecting

aggregates of iMB-VOs to chronic electrical stimulation, showing enhanced skeletal muscle myosin and calcium handling gene expression. We present this approach as a proof-of-concept for the introduction of other parenchymal cell types of interest, via the overexpression of lineage-specifying TFs, in the context of a vascular organoid scaffold.

4.3 Materials and methods

4.3.1 Plasmid Construction

The piggyBac transposon plasmids for inducible overexpression of TFs were constructed using the backbone from 138-dCas9-Dnmt3a (Addgene #84570)[246]. 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 #120427), phDLX2_N174 (Addgene #60860)[247], EF1a_MYOD1_P2A_Hygro_Barcode (Addgene #120464), EF1a_ETV2_P2A_Hygro_Barcode (Addgene #120436), pBS-hBAF60c (Addgene #21036)[248], EF1a_mCherry_P2A_Hygro_Barcode (Addgene #120426), and pEGIP (Addgene #26777).

To construct a plasmid expressing the hyperactive piggyBac transposase[249], 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).

4.3.2 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).

4.3.3 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 a nucleofection program specific for hESCs was run on a 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.

4.3.4 2D Differentiation of Inducible Overexpression hESC lines

Clonal lines overexpressing *NEUROD1* were differentiated following a previously described protocol[122]. 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[45]. 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 *ETV2* were differentiated to endothelial cells using a previously demonstrated protocol[188]. 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, the medium was changed to EGM-2 (Lonza) containing 0.75 µg/ml puromycin and 1 µg/ml doxycycline. Cells were maintained in these conditions for 5 days.

Clonal lines overexpressing *MYOD* and *BAF60C* were differentiated to skeletal muscle following a process similar to a previously described protocol[250]. 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.

4.3.5 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 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 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 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 StemPro 34 Media + 15% FBS + 100 ng/ml VEGF + 100 ng/ml bFGF + 1 μg/ml dox until day 15.

For long term culture of iN-VOs, at day 15, medium was changed to StemPro 34 Media + 15% FBS + 100 ng/ml VEGF + 100 ng/ml bFGF + 20 ng/ml BDNF + 20 ng/ml NT3 + 1 μg/ml dox. Medium was replaced every 3-5 days.

4.3.6 Animals

Housing, husbandry and all procedures involving animals used in this study were performed in compliance with protocols 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.

4.3.7 In Vivo Perfusion

iN-VOs were cultured for 30 days before being implanted subcutaneously into $Rag2^{-/-};\gamma c^{-/-}$ 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.

4.3.8 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.

4.3.9 *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[251]. 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.

4.3.9 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 (MY-32, Sigma-Aldrich) 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.

4.3.10 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.

4.3.11 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).

4.3.12 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[196]. Expression matrices were filtered to remove any cells expressing less than 200 genes or expressing greater than 10% mitochondrial genes. DoubletFinder[197] 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[252] 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.

4.3.13 Gene Ontology Enrichment

Significant differentially expressed genes for each cell type were identified with thresholds of log fold change = 1, a minimum of 50% of cells in the cluster expressing the gene, and p-value < 0.05 by Wilcoxon rank sum test. Gene ontology enrichment analysis was then performed using topGO[253], with the elim method. Significant terms were selected as those with p-values less than 0.01.

4.4 Results

4.4.1 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 4.2.a**). To establish the utility of our PiggyBac overexpression platform, we first constructed a diverse array of cell lines and validated their ability to differentiate into functional tissue upon adding doxycycline. For neural tissue differentiation, we chose to overexpress *NEUROD1* to generate glutamatergic excitatory neurons[122,188], and *ASCL1+DLX2* for the differentiation of GABAergic inhibitory neurons[45]. For mesodermal tissue differentiation, we chose to overexpress established reprogramming factors *ETV2* for generation of endothelial cells, and *MYOD1+BAF60c*[250] for skeletal muscle differentiation.

Stable, dox-inducible *NEUROD1* (iN), *ASCL1+DLX2* (iAD), *ETV2* (iE), 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[249]. 48 hours after nucleofection, cells were maintained under puromycin selection, and clonal populations were picked and expanded. Clones for each cell line were individually validated by confirming dox-induced lineage-specific differentiation (**Figure 4.2.b**).

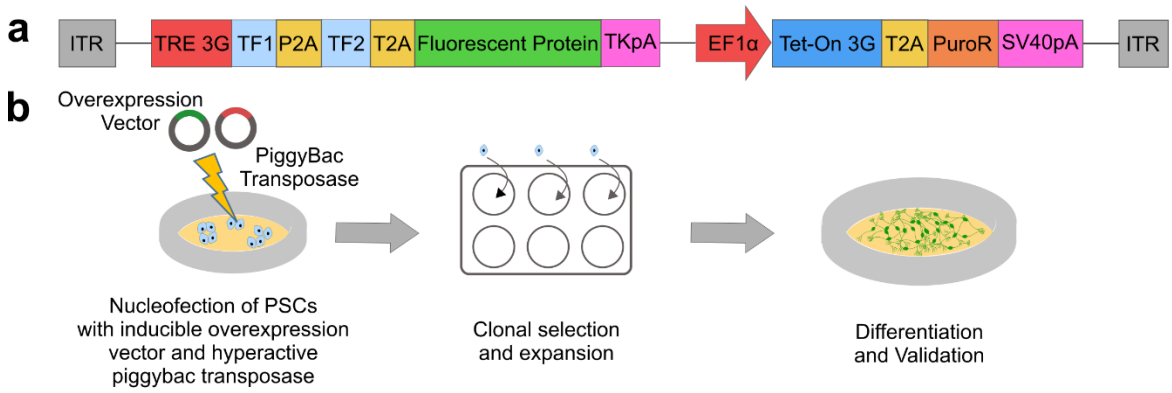


Figure 4.2: Schematic for construction of inducible cell lines

(a) Schematic of PiggyBac transposon based inducible overexpression vector. **(b)** Schematic of cell line generation and validation process.

For validation of neural tissue differentiation of iN and iAD lines, qRT-PCR analysis was conducted and demonstrated a strong upregulation of neuronal markers *MAP2* and *TUBB3* for both iN and iAD cells (**Figure 4.3.a, 4.3.b**) compared to undifferentiated hESCs. iN neurons were confirmed to be glutamatergic, demonstrated by upregulation of *vGLUT2* with no detectable expression of the inhibitory neuronal marker *vGAT* (**Figure 4.3.a**), while iAD neurons were confirmed to be GABAergic, demonstrated by upregulation of *vGAT* with minimal expression of the excitatory neuronal marker *vGLUT2* (**Figure 4.3.b**). *MAP2*-positive cells with classic neuronal morphology were also confirmed by immunofluorescence for both iN and iAD cells (**Figure 4.3.a, 4.3.b**). 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 4.3.a, 4.3.b**).

Differentiation of iE cells to endothelial-like cells was also confirmed by gene expression, immunofluorescence and functional assays (**Figure 4.3.c**). qRT-PCR analysis demonstrated a strong upregulation of endothelial markers *CD31* and *CDH5*. Endothelial identity was further confirmed by *CDH5*-immunostaining, where we observed characteristic endothelial-like

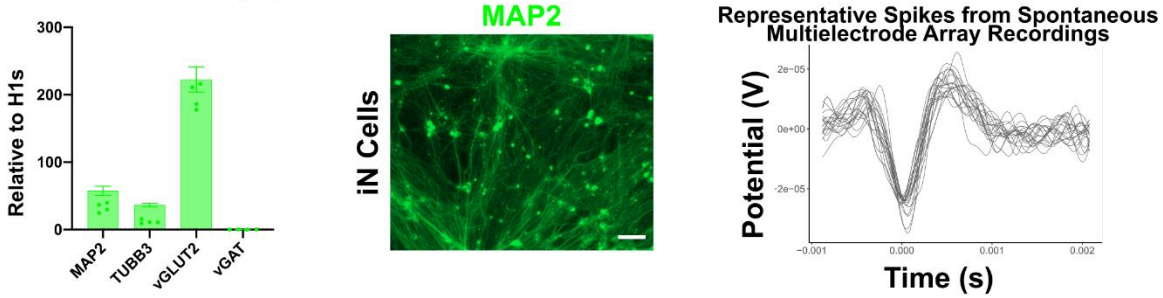
morphology. Lastly, endothelial function was confirmed via a tube formation assay by plating differentiated iE cells on Matrigel.

Finally, skeletal muscle differentiation of iMB cells was validated by gene expression and immunofluorescence assays (**Figure 4.3.d**). qRT-PCR analysis confirmed upregulation of skeletal muscle markers *MYH8*, *TNNC1* and *RYR*. Immunostaining confirmed the presence of characteristic spindle-shaped, *MYH*- and *SAA*-positive skeletal muscle morphology, along with *MYOG*-positive nuclei in iMB cells.

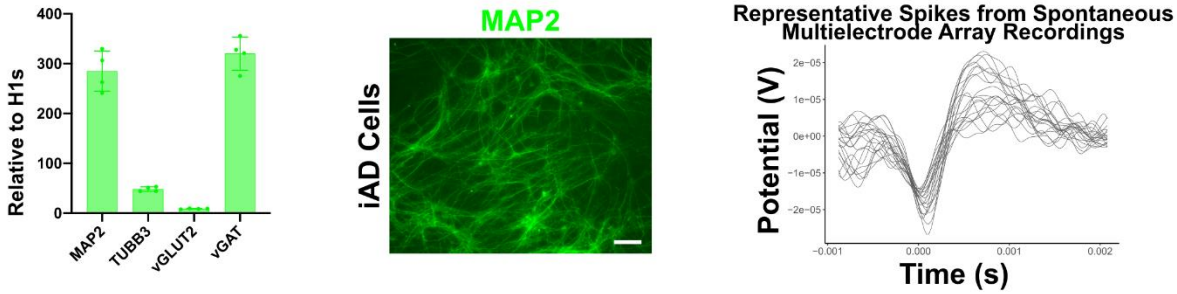
Figure 4.3: Validation of inducible cell lines

(a) Inducible *NEUROD1* (iN) cell line validation at 3 weeks post induction via: 1) qRT-PCR analysis of signature neuronal marker genes *MAP2*, *TUBB3*, *vGLUT2* and *vGAT*; Data represent the mean ± s.d. (n = 4 independent experiments) 2) immunofluorescence micrograph of *MAP2* labelled cells (Scale bars = 50 μm); and 3) representative spike plots from MEA measurements of spontaneously firing iN cells. **(b)** Inducible *ASCL1+DLX2* (iAD) cell line validation at 3 weeks post induction via: 1) qRT-PCR analysis of signature neuronal marker genes *MAP2*, *TUBB3*, *vGLUT2* and *vGAT*; Data represent the mean ± s.d. (n = 4 independent experiments) 2) immunofluorescence micrograph of *MAP2* labelled cells (Scale bars = 50 μm); and 3) representative spike plots from MEA measurements of spontaneously firing iAD cells. **(c)** Inducible *ETV2* (iE) cell line validation at 2 weeks post induction via: 1) qRT-PCR analysis of signature endothelial marker genes *CDH5* and *CD31*; Data represent the mean ± s.d. (n = 2 independent experiments) 2) immunofluorescence micrograph of *CDH5* labelled cells; and 3) matrigel tube formation assay done on iE cells (Scale bars = 25 μm). **(d)** Inducible *MYOD1+BAF60C* (iMB) cell line validation at 2 weeks post induction via: 1) qRT-PCR analysis of signature skeletal muscle marker genes *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)**.**

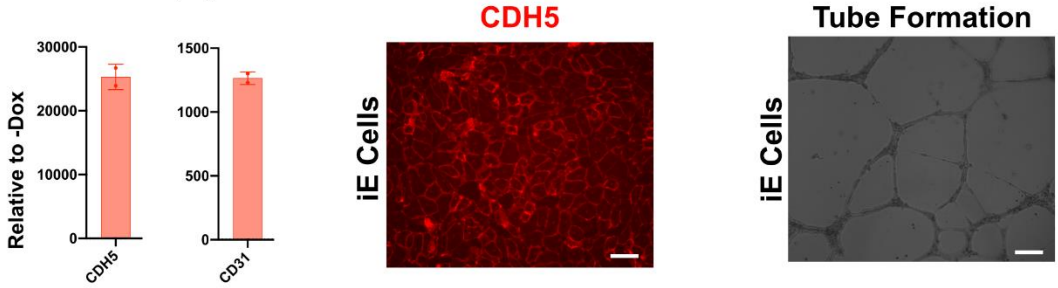
a Inducible *NEUROD1* (iN) Neuronal Cell Line Validation



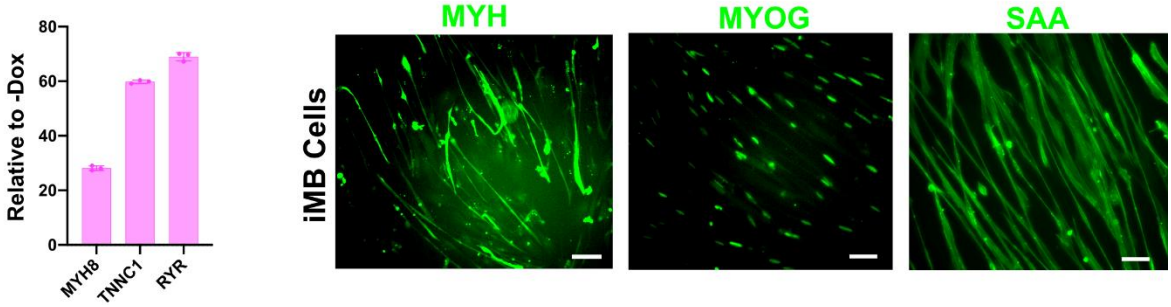
b Inducible *ASCL1* + *DLX2* (iAD) Neuronal Cell Line Validation



c Inducible *ETV2* (iE) Endothelial Cell Line Validation



d Inducible *MYOD1* + *BAF60C* (iMB) Skeletal Muscle Cell Line Validation



4.4.2 Constructing iN-VOs and iMB-VOs

After clonal hESC lines were established, we developed a modular, generalizable, TF overexpression based strategy to differentiate parenchymal cell types in VOs (**Figure 4.1**). We sought to assess first whether iN hESCs would differentiate to neurons when *NEUROD1* expression was induced by doxycycline under the culture conditions of VO differentiation (**Figure 4.4.a**). Since we did not know a priori if all iN cells would differentiate to neurons and disrupt formation of vascular networks in the organoid, we first tested induced neuro-vascular organoid (iN-VO) formation with two ratios of iN:Wild-Type (WT) cells, either 100% iN cells or a mix of 50% iN cells and 50% WT cells (**Figure 4.4.b**). The organoids were assayed at day 15 for expression of the neuronal marker *MAP2* and the endothelial marker *CDH5* by qRT-PCR. Induced organoids (+dox conditions) for both cell ratios showed a marked increase in *MAP2* expression, with expression levels 4- to 8-times higher compared to uninduced organoids (-dox conditions) (**Figure 4.4.c**), suggesting robust neuronal differentiation under doxycycline induction. Strikingly, induced organoids at both ratios showed comparable expression of *CDH5* as uninduced organoids (**Figure 4.4.c**), suggesting that endothelial networks are reliably formed and preserved even with induction of *NEUROD1* expression. Immunostaining was performed on both induced (+dox, iN-VOs) and uninduced (-dox, iN-VOs) organoids confirming *MAP2*-positive cells forming into bundle-like structures of neurons only in induced organoids, along with dense, interpenetrating *CDH5*-positive vascular networks in both conditions (**Figure 4.4.d**). More extensive gene expression analysis via qRT-PCR showed an upregulation of the excitatory neuron marker *vGLUT2,* along with cortical neuron markers *BRN2* and *FOXG1*[122] in day 15 iN-VOs (**Figure 4.4.e**). Moreover, we observed expression levels comparable to uninduced organoids of the endothelial markers *CDH5* and *VEPTP*, along with the smooth muscle marker *SMA* (**Figure 4.4.e**), confirming the maintenance of mature and complete blood vessel networks in day 15 iN-VOs. Given that *CDH5* expression remains high even in induced organoids formed from 100% iN

cells, we hypothesize that a subset of iN cells escapes neuronal differentiation due to stochastic silencing of the integrated PiggyBac cassette, leading them to differentiate to vascular lineages. Given the higher expression of *MAP2* and comparable expression of *CDH5*, subsequent experiments were conducted with iN-VOs formed from 100% iN cells.

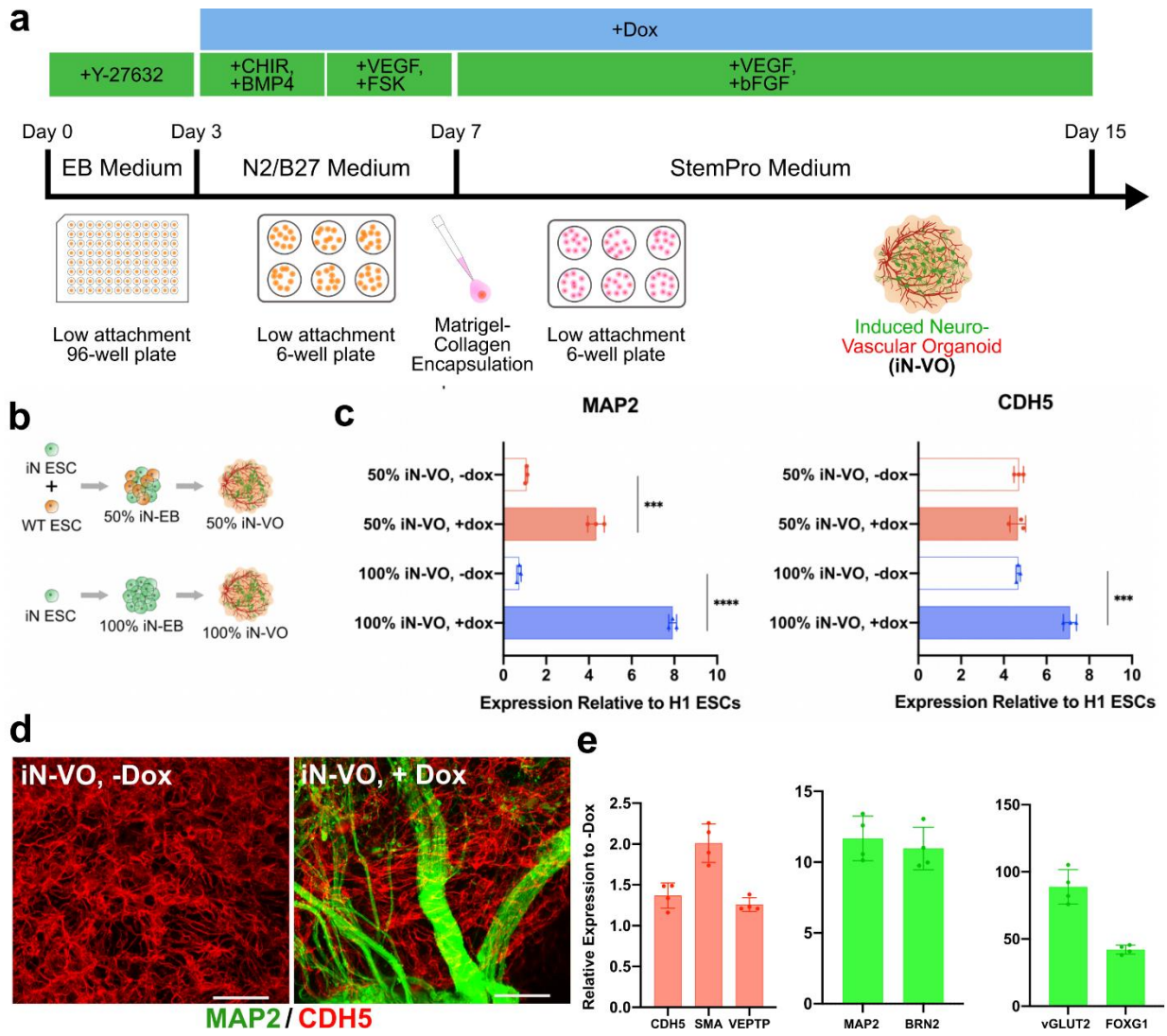


Figure 4.4: Generation of neuro-vascular organoids

(a) Culture protocol for the generation of induced neuro-vascular organoids. **(b)** Schematic for experiment to assess the optimal ratio of iN to WT cells for iN-VO formation. **(c)** 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. 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). **(d)** 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). **(e)** 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 = 4 organoids, from two independent experiments)

Once the formation of neuro-vascular iN-VOs was established, we then sought to demonstrate if our system was compatible with other germ-layer or cross-lineage tissue types. Using 100% iMB cells and the exact same VO media conditions used to generate iN-VOs, we grew induced myo-vascular organoids (iMB-VOs) implementing the same protocol used to generate iN-VOs (**Figure 4.5.a**). iMB-VOs had *MYH*-positive spindle-shaped morphologies (**Figure 4.5.b**), along with expression of skeletal muscle marker genes *MYOG*, *MYH8*, *TNNC1* and *RYR* (**Figure 4.5.c**), confirming that our platform could also be used to grow vascularized organ tissue of mesodermal origin, specifically skeletal muscle.

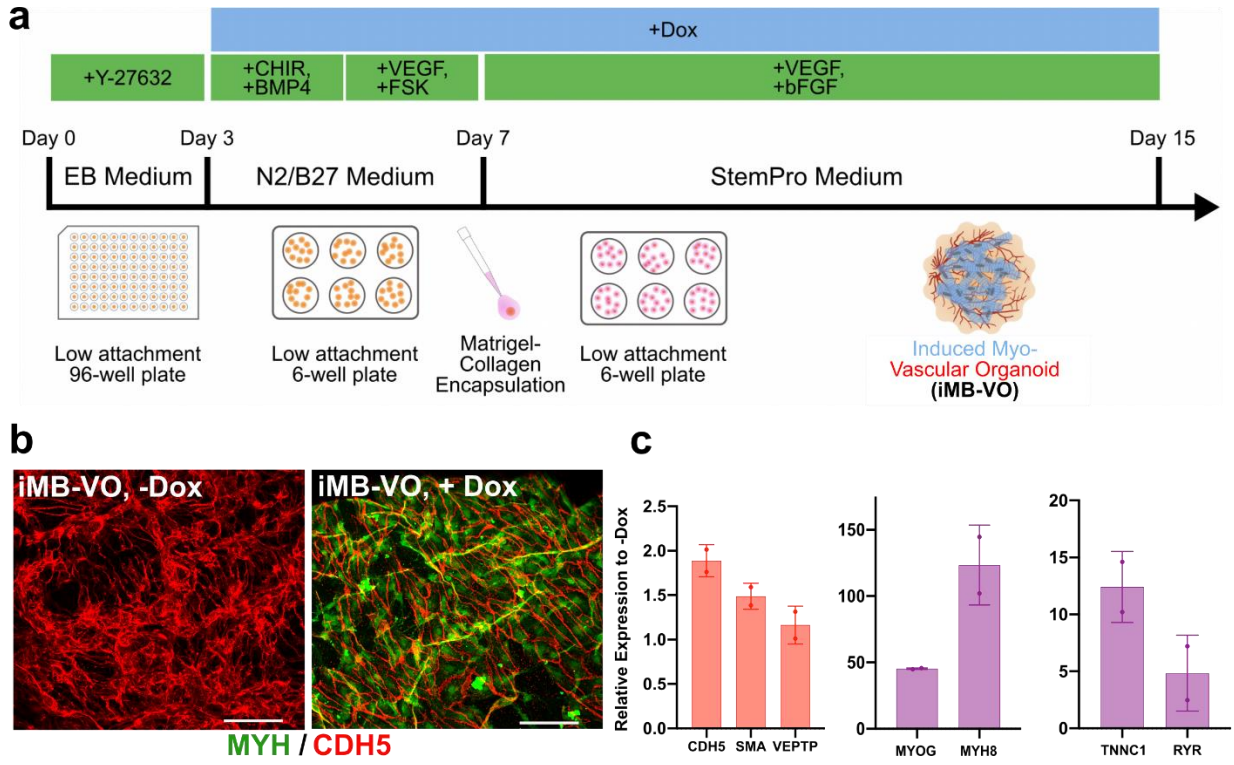


Figure 4.5: Generation of myo-vascular organoids

(a) Culture protocol for the generation of induced myo-vascular organoids. **(b)** 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). **(c)** 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 = 4 organoids, from two independent experiments).

4.4.3 Optimizing and characterizing long-term cultured iN-VOs

Upon optimizing our culture system to grow both neuro-vascular and myo-vascular organoids for 15 days, we then sought to enable long-term culture of the iN-VOs for beyond 15 days (**Figure 4.6.a**). When cultured in standard VO media conditions, neuronal survival was compromised and by day 30 there was no observable *MAP2* expression in induced organoids compared to uninduced organoids (**Figure 4.6.b**). To enhance the long term survival of neurons in the organoids, we tested supplementing standard VO media with the neurotrophic growth factors BDNF and NT-3, which are known to support neuronal survival *in vitro*[254,255], as well as the well-established neuronal cell culture supplement B-27. Supplements were added to the culture from day 15 onward, and organoids were assayed at day 30. In addition, we tested whether continued overexpression of *NEUROD1* beyond day 15 may lead to impaired neuronal survival by including conditions where doxycycline was removed from culture after day 15. All organoids were assayed at day 30 by qRT-PCR for *MAP2* and *CDH5* expression. We observed highest *MAP2* expression in organoids where doxycycline was retained in the medium till day 30, and the medium was supplemented with BDNF and NT-3, about 10-times higher than uninduced organoids (**Figure 4.6.b**). This was followed by organoids where medium was supplemented with BDNF, NT-3 and B-27 and doxycycline retained till day 30, where *MAP2* expression was about 5-times higher than uninduced organoids (**Figure 4.6.b**). On the other hand, only organoids subjected to medium containing BDNF and NT-3 were best able to maintain *CDH5* expression comparable to uninduced organoids (**Figure 4.6.b**). Thus, for subsequent long-term experiments we supplemented the culture medium with BDNF and NT-3 from day 15 onward to ensure long term survival of neurons with reliable maintenance of vascular lineages.

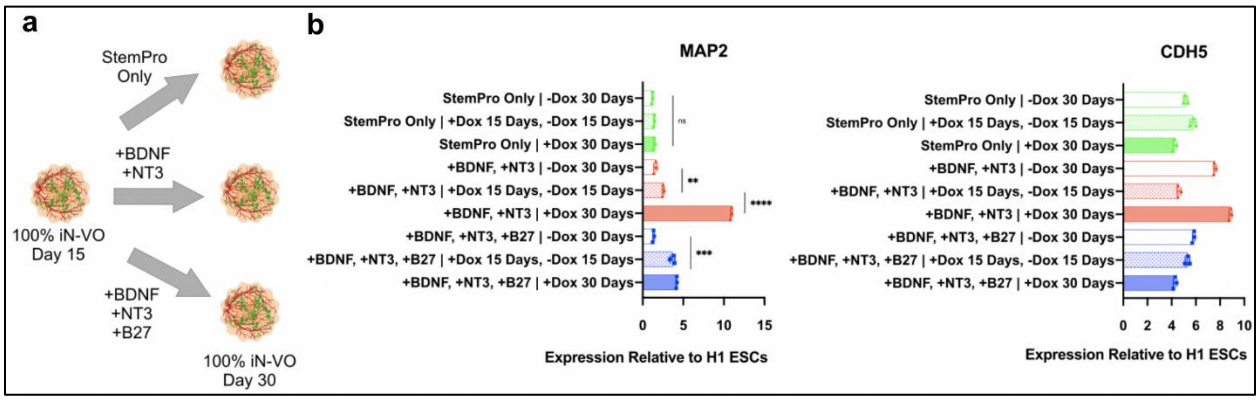


Figure 4.6: Long term culture of neuro-vascular organoids

(a) Schematic for experiment to assess optimal media supplements for long term iN-VO culture. **(b)** 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. 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).

Upon determining this optimized and repeatable protocol for long-term growth of iN-VOs, we comprehensively characterized these long-term cultured organoids (**Figure 4.7.a**).

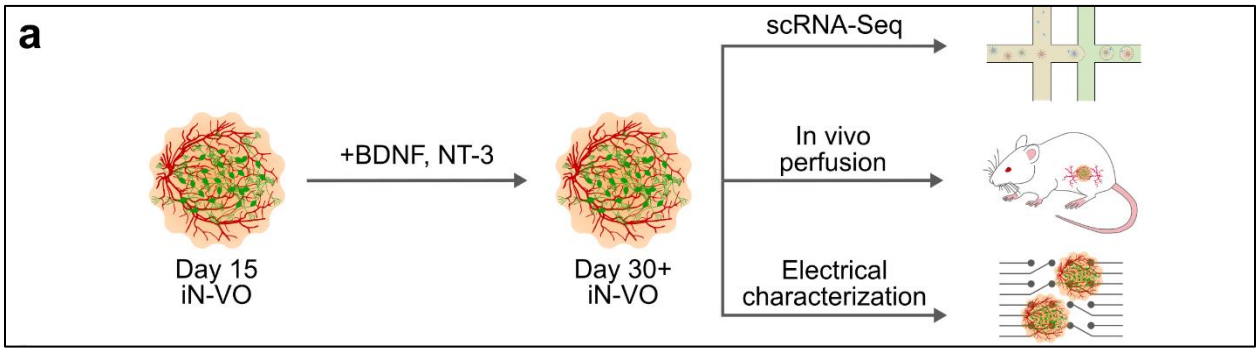


Figure 4.7: Characterization of long-term cultured neuro-vascular organoids

(a) Schematic for comprehensive characterization of long-term cultured neuro-vascular organoids.

To characterize the 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). We included organoids for which *NEUROD1* was

overexpressed for 15 days and then stopped, to test if we could obtain non-neuronal cell types or different neuronal subtypes by allowing for a degree of spontaneous differentiation after priming cells for a particular lineage. Organoids were dissociated using papain and sequencing libraries generated using droplet-based scRNA-seq on the 10X Chromium platform. Expression matrices were generated from the resultant data using 10X Genomics cellranger. We then used these expression matrices with the Seurat pipeline[196] to integrate data from all three types of organoids and remove any batch effects. The cells were then clustered and cell types were first coarsely classified using the Microwell-seq Mouse Cell Atlas as a reference[252], followed by a more refined classification using cell-type specific markers (**Figure 4.8.a**).

Using this pipeline, we identified that *PRRX1*+ mesenchymal progenitors[256], *MEOX2*+ differentiating pericytes[257] and cycling cells were present in all three types of organoids and constituted a majority of each of the organoids (**Figure 4.8.a-d**). 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 4.8.a-d**). We also found a population of *COL15A1*+ stromal cells – type XV collagen is an important component of microvessel basement membranes[258]. Interestingly, a population of *CD68*+/*CD74*+ macrophage-like cells was also present, suggesting the presence of immune cell types within VOs, and unexpectedly we also found a population of epithelial cells, which were *EPCAM*+ and included a *KRT17*+ subset as well as an *APOA1*+ subset which was enriched in lipid processing genes (**Figure 4.8.a-d**). Critically, *DCX*+ neurons were only present in organoids where *NEUROD1* expression was induced and constituted 4-6% of cells (**Figure 4.8.a-d**). Differentially expressed genes in the neuronal cluster also showed significant enrichment of neural-related Gene Ontology terms (**Figure 4.8.e**).

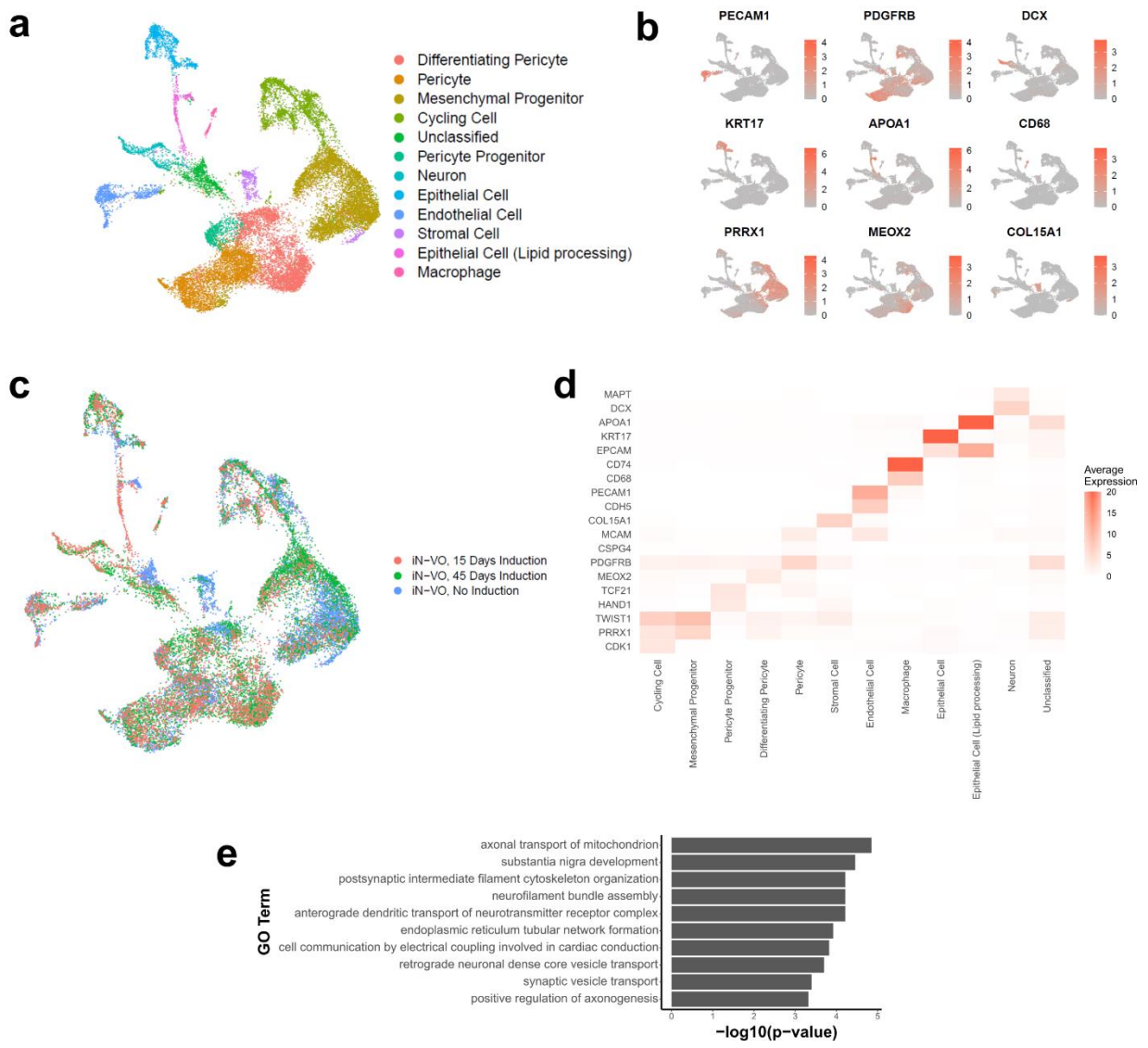


Figure 4.8: Single cell RNA-seq characterization of long-term cultured neuro-vascular organoids

(a) UMAP visualization of cell types from day 45 iN-VOs. Two independent induction conditions, along with one non-induction condition. **(b)** Cluster-specific expression of marker genes in day 45 iN-VOs. **(c)** UMAP visualization of iN-VO clusters annotated by sample type. Two independent induction conditions, along with one non-induction condition. **(d)** Expression of marker genes for each classified cell type. **(e)** Top 10 significantly enriched Gene Ontology terms for differentially expressed genes in the neuronal cluster.

Immunofluorescence imaging confirmed the presence of *MAP2*+ neurons displaying robust neurite growth, along with interpenetrating *CDH5*+ vascular networks through 30 days in culture (**Figure 4.9.a**). Bulk gene expression analysis via qRT-PCR showed an upregulation of the neuronal markers *MAP2*, *vGLUT2*, *BRN2* and *FOXG1* in day 30 iN-VOs (**Figure 4.9.b**). Similarly, we observed expression levels comparable to uninduced organoids of the endothelial

markers *CDH5* and *VEPTP*, along with the smooth muscle marker *SMA* (**Figure 4.9.b**), confirming the maintenance of mature and complete blood vessel networks with interspersed programmatically differentiated neurons in long-term cultured iN-VOs.

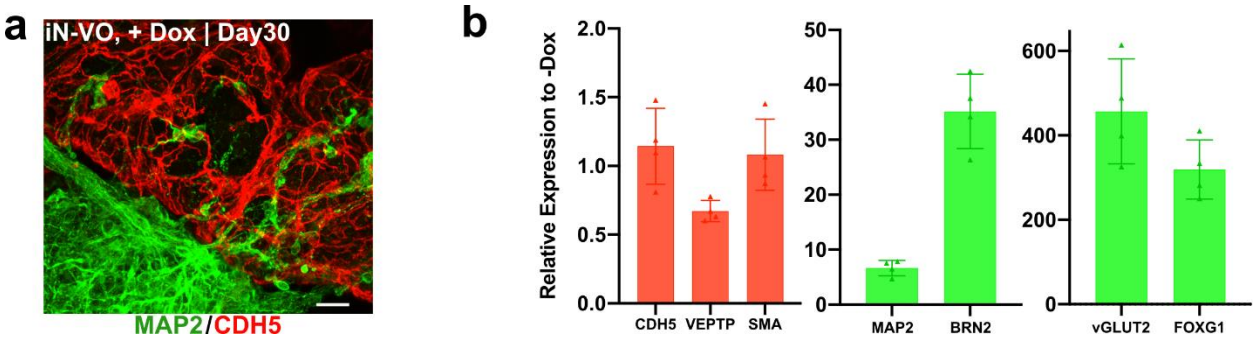


Figure 4.9: Marker gene expression in long-term cultured neuro-vascular organoids

(a) Immunofluorescence 100 µm z-stack, maximum projection, confocal micrographs of *MAP2*- and *CDH5*-labelled induced day 30 iN-VOs (Scale bars = 100 µm) **(b)** 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 30 of iN-VO culture. Data represent the mean ± s.d. (n = 4 organoids, from two independent experiments).

After molecular characteristics of the iN-VOs were assayed, we assessed the functionality of both neural and vascular cell-types in the iN-VO. To confirm functionality of the endothelium, we assayed the ability of the vascular networks to form perfusable blood vessels *in vivo*. iN-VOs were grown till day 30 *in vitro*, and then subcutaneously implanted in $Rag2^{-/-}$;$\gamma c^{-/-}$ immunodeficient mice. 30 days post-implantation, intravenous (IV) injection of a Texas Red dye followed by organoid extraction, fixation and immunofluorescent staining with *CDH5* showed co-localization of the dye with human *CDH5*+ vascular cells in the iN-VO (**Figure 4.10**). Thus demonstrating that iN-VOs successfully engrafted and connected with the host-vasculature.

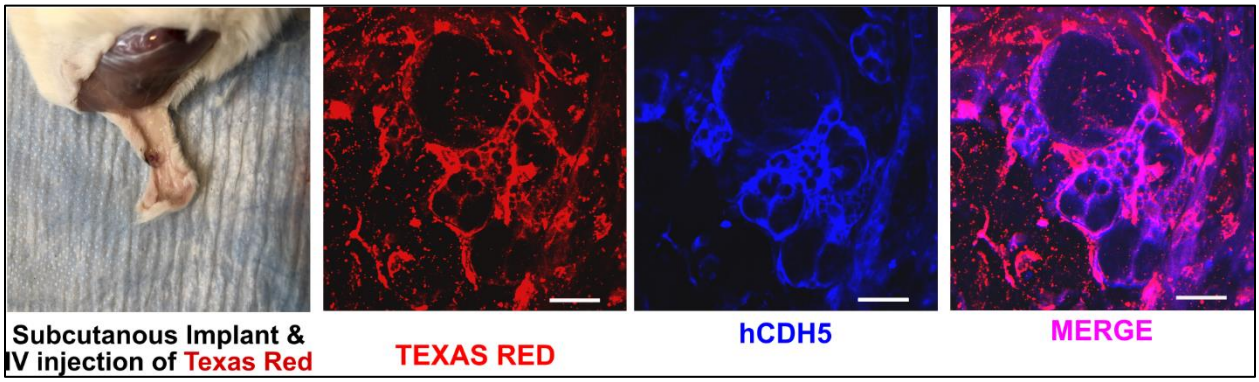


Figure 4.10: Perfusion testing of neuro-vascular organoids

Experimental validation of iN-VO perfusibility in vivo by subcutaneous implantation of iN-VO, showing immunofluorescence micrographs of Texas Red, *CDH5* and overlay (Scale bar = 25 μm). Representative image from two independent experiments.

To assess functionality of neurons in the iN-VOs, we assayed the organoids for spontaneous electrical activity using microelectrode arrays (MEAs) (**Figure 4.11**). Strikingly, while uninduced organoids displayed no spontaneous activity, spontaneous firing was repeatedly observed in iN-VOs (**Figure 4.11**), confirming the presence of functional neurons in the organoids. Interestingly, while neurons cultured in 2D displayed spontaneous activity only when co-cultured with glia (**Figure 4.3.a**), neurons which were differentiated and cultured in iN-VOs did not need the presence of glia for spontaneous firing. Taken together, we demonstrate the formation of functional neuro-vascular tissue with long term culture capability, using our method of combining directed differentiation with genetic overexpression in vascular organoids.

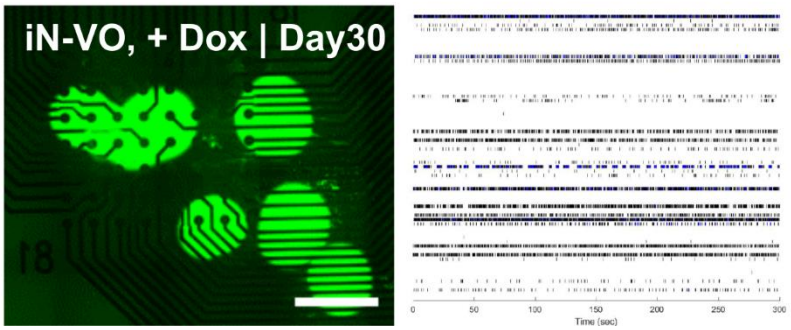


Figure 4.11: Electrophysiological testing of neuro-vascular organoids

Image of day 30 iN-VOs plated on microelectrode array (Scale bars = 500 μm) and raster plot of spontaneous activity from plated iN-VOs. Representative image and plot from two independent experiments.

4.4.4 *In vitro* maturation of iMB-VOs

Finally, while our iMB-VO platform was able to differentiate vascularized skeletal muscle, maturation of this lineage *in vitro* is a longstanding challenge. In order to further mature the differentiated skeletal muscle, we applied chronic electrical stimulation, which is an established strategy for skeletal muscle maturation[251,259,260].

To subject organoids to stimulation, we encapsulated the organoids in a hydrogel composed of fibrin and matrigel, and placed them in a custom designed chip between two graphite rods. A pulsed constant voltage stimulation was then applied to the organoids for a week after encapsulation to drive maturation (**Figure 4.12.a**). Organoids were then assayed for gene expression of muscle and calcium handling genes (**Figure 4.12.b**). We observed that embryonic skeletal muscle myosin (*MYH3*) was upregulated, while there was a small increase in expression of adult fast skeletal muscle myosin (*MYH2*) for stimulated organoids. Additionally, the calcium handling genes *CASQ2* and *SERCA2* were also highly upregulated. Lastly, we assayed both unstimulated and stimulated organoids by *MYH*-immunofluorescence imaging (**Figure 4.12.c**), confirming the presence of skeletal muscle fibers in both, indicating that while the skeletal muscle lineage is preserved in both conditions, stimulation drives maturation of the differentiated skeletal

muscle. This also confirmed that our iMB-VO organoids could be cultured for at least 21 days *in vitro* without any further optimization of media conditions, and opens the door for potentially longer-term culture for enhanced tissue maturation. In summary, our iMB-VO approach is a feasible method to generate mature, vascularized skeletal muscle tissue *in vitro*.

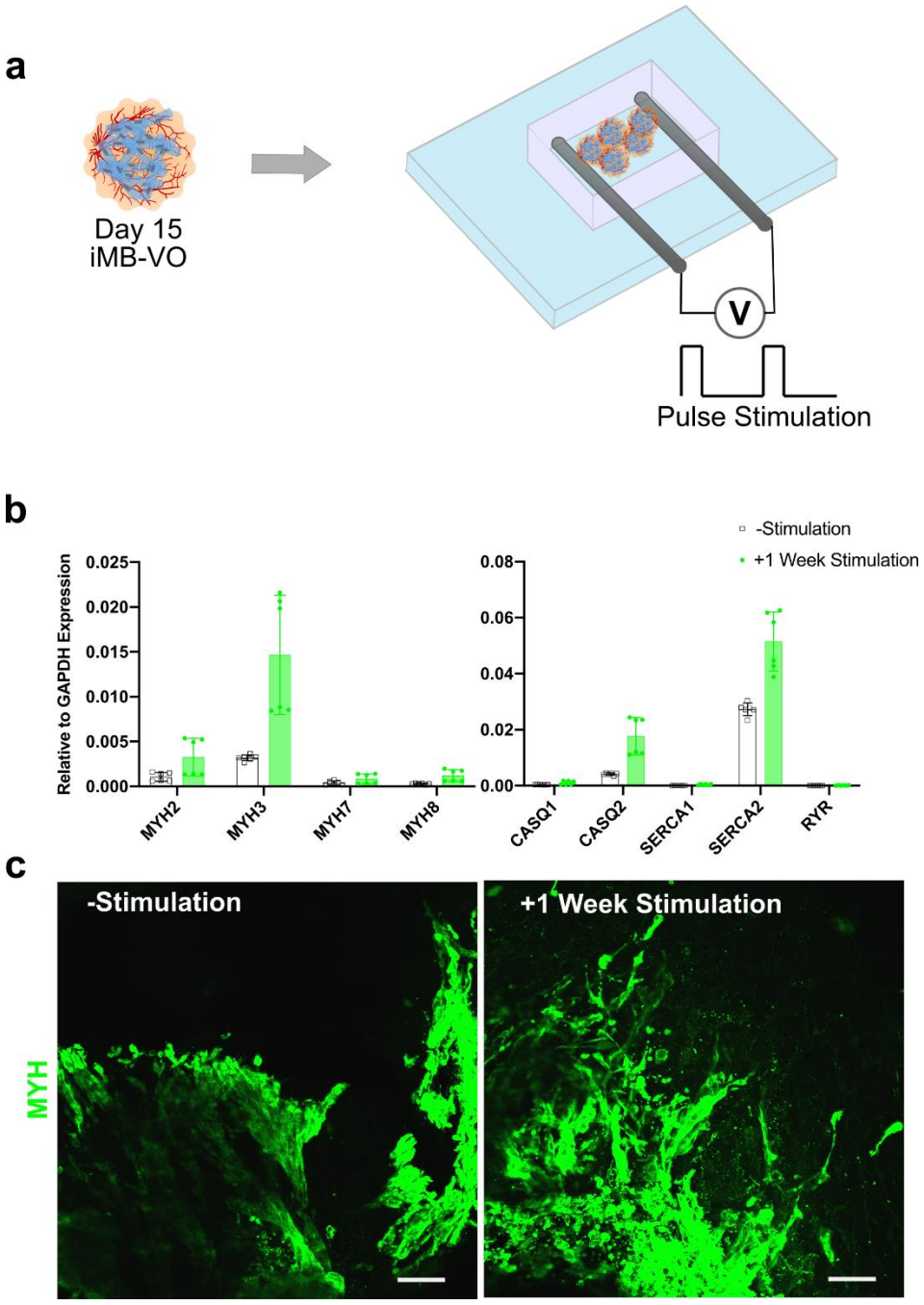


Figure 4.12: *In vitro* maturation of iMB-VOs

(a) Schematic of iMB-VO *in vitro* maturation by electrical stimulation. **(b)** 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) **(c)** Immunofluorescence 100 µm z-stack, maximum projection, confocal micrographs of *MYH*-labelled stimulated vs unstimulated iMB-VOs (Scale bars = 50 µm).

4.5 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. Here, we have built a piggyBac transposon-based overexpression system which offers the potential of large payload capacity for overexpression of multiple elements from a single vector. We used this system to introduce neurons into vascular organoids by the expression of *NEUROD1* in differentiating organoids, as well as to introduce skeletal muscle by the expression of *MYOD1* + *BAF60C* in differentiating 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 function, demonstrating that all vascular lineages along with neurons were present in induced organoids, and that both lineage branches, neuronal and vascular, were functional. 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. We envision that combining these modular building blocks may enable the fully *in vitro* creation of innervated and vascularized human skeletal muscle entirely from pluripotent stem cells, a major step toward regenerative therapy for volumetric muscle loss.

While providing a modular and powerful approach to engineer vascularized multi-lineage organoids, several challenges still remain to be solved. As demonstrated in our study, while we were able to differentiate neurons within the vascular organoid scaffold, maintaining their viability and function required careful and extensive optimization of culture conditions. Also, while we have

taken advantage of known 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[261] 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[262,263] or synthetic biology approaches[264].

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 many kinds of vascularized organoids, thus laying the foundation for completely *in vitro* derived, large-scale tissue systems for regenerative medicine purposes.

4.6 Acknowledgements

We would like to thank Yan Wu, Kian Kalhor, Reinner Wimmer, Cleber Trujillo and Ramin Dailamy for useful discussions and support throughout all stages of the study; and Jennifer Santini from UCSD School of Medicine Microscopy Core (Grant P30 NS047101). This work was generously supported by UCSD Institutional Funds and NIH grants (R01HG009285, RO1CA222826, RO1GM123313). This publication includes data generated at the UC San Diego IGM Genomics Center utilizing an Illumina NovaSeq 6000 that was purchased with funding from a National Institutes of Health SIG grant (#S10 OD026929).

Chapter 4 in part is a reprint of the material: Dailamy, A.*, Parekh, U.*, Katrekar, D., McDonald, D., Kumar, A., Moreno, A., Bagheri, P., Ng, T. N. & Mali, P. Programmatic Introduction of Parenchymal Cell Types in Blood Vessel Organoids (2021), currently under peer review (*equal contribution). The dissertation author is one of the two primary authors of the study.

Chapter 5: Flexible Printed Electrodes for Electrophysiology

5.1 Abstract

The measurement of bioelectronic signals for clinical and research applications requires the development of electrode systems which can facilitate the multi-scale measurement of signals at the organism, organ, network and single cell level. Here we present the development of printed sensors to measure electrophysiological signals *in vivo* and microelectrodes for the measurement of cellular scale field potential. To measure clinically significant signals for electroencephalography, electromyography, and electrocardiography, we developed stencil-printed multichannel concentric-ring electrodes on stretchable elastomers modified to improve adhesion to skin and minimize motion artifacts. These dry electrodes with a poly(3,4-ethylenedioxythiophene) polystyrene sulfonate (PEDOT:PSS) interface layer are optimized to show lower noise level than that of commercial electrodes, while the concentric ring geometry enables Laplacian filtering to pinpoint the bioelectric potential source with spatial resolution determined by the ring distance. We also present a preliminary demonstration of a novel method combining inkjet and electrohydrodynamic printing to fabricate flexible microelectrode arrays in a facile, reconfigurable manner.

5.2 Introduction

Measurement of electrophysiological signals is crucial for monitoring bodily states, facilitating clinical diagnosis of diseases and for interrogating electrically active cells and cellular networks. These measurements can be cumbersome and challenging due to the soft, deformable nature of tissue and arbitrary shapes which are not amenable to contact with rigid substrates. The development of flexible, conformal, rapidly prototyped systems which can be changed to accommodate various geometries, stiffness and motion are critical to advancing the ability to precisely measure electrical signals from tissue both *in vivo* and *ex vivo*.

Clinical *in vivo* electrophysiological recordings such as electroencephalography (EEG), electromyography (EMG), and electrocardiography (ECG) are routinely done using arrays of electrodes that measure local electrical activities[265–267]. Spatial information is important in EEG to localize active brain regions, and in EMG to measure neuromuscular activity. In ECG, the progression of cardiac depolarization signals can be tracked using twelve leads located on the torso to diagnose various cardiac abnormalities, such as arrhythmias and infarctions. However, the spatial resolution of an electrode array is affected by the blurring effects from different conductivities of the torso volume, and modified electrode designs[268,269] are shown to be more effective for increasing spatial resolution than merely increasing the number of electrodes in an array. In addition, conventional gel electrode arrays are limited by the complexity of electrode placement, which is laborious, time-consuming, and only transiently useful since performance degrades once the gel dries. Thus, to enhance the spatial resolution and reliability of electrophysiology studies while reducing setup complexity, we developed a dry concentric ring electrode design through an additive printing process[270–275] that allows rapid reconfiguration and integration of multi-channel electrodes.

We take advantage of the recent advances in the use of conducting polymer[276–279] as dry electrode interface. The conducting polymer poly(3,4-ethylenedioxythiophene) polystyrene

sulfonate (PEDOT:PSS) has enabled a better interface with skin, because of the mixed ionic/electronic conductivity[280] and biocompatibility[281] of PEDOT:PSS. Moreover, we also take advantage of developments in elastomers to pattern the dry electrodes on stretchable elastomers modified to improve adhesion to skin[282] and minimize motion artifacts.

Secondly, the advent of tissue engineering and organoid technology[48,49,283–285] has revolutionized the *ex vivo* modeling of human tissues and organs to study development and disease. While the engineering of organoid systems has progressed considerably to enable repeatability, complexity and precise control, tools for functional electrical interrogation of these 3-dimensional cell ensembles remain limited[286].

While microelectrode arrays (MEAs) remain a workhorse tool for functional characterization of 2-dimensional monolayer cell cultures, the typically rigid substrates are not ideal for measurements of soft, 3-dimensional organoids. In addition, arrays which can be easily adapted to the heterogenous geometries of organoids and other engineered tissue, would be needed for accurate mapping of signals. One promising path by which this may be addressed is through the use of optical electrophysiology tools[287,288], where voltage sensitive fluorescent indicators can be optically read out to record cellular electrical activity at high spatiotemporal resolution. But these approaches, while versatile and powerful, can be limited by the optical properties of the tissues under investigation. Another recent approach is to create flexible MEAs which can be embedded[289] or envelop the surface of organoids[92,290]. These approaches offer the combination of high coverage and potentially high spatiotemporal resolution and can also be compatible with floating culture common with organoid and engineered tissue.

While traditional fabrication techniques have been used for these promising techniques, printed electronics enables the ability to rapidly reconfigure architectures for the accommodation of any arbitrary geometry which is important not only for these self-assembled organoid systems but also for other engineered tissue in research and regenerative medicine applications. Such printed electrodes have been recently demonstrated via inkjet printing[291,292], but have two main

limitations. Inkjet printing typically results in minimum dimensions which are larger[291] than standard commercial MEAs which limits the resolution of the recorded signal, and to print such electrode arrays on soft, flexible materials careful treatment of the substrate is needed to adjust the wetting properties for printing of connected patterns[293].

We present here, a preliminary demonstration of a novel printing method which combines inkjet and electrohydrodynamic printing with a facile aqueous sacrificial etch process[294] to fabricate flexible microelectrodes with down to 10 μm diameter. Thus, potentially overcoming the drawbacks of other printed electrode fabrication processes.

5.3 Materials and Methods

5.3.1 Wearable Screen-Printed Electrode Fabrication

The substrate was comprised of polydimethylsiloxane (PDMS, Sylgard 184, Dow Corning) modified with PEIE (80% ethoxylated polyethylenimine solution, Sigma-Aldrich) to tune the mechanical compliance and adhesive properties of the PDMS elastomer. The PDMS base and cross-linking agent were mixed at 10:1 ratio by weight; for 10 g of the mixture, 35 μL of PEIE was added to induce a heterogeneous cross-linking network[282]. The mixture was de-gassed and poured into a mold to cure at 90°C for 3 hours to obtain a substrate with typical thickness of 2 mm. To form electrodes, conductive inks were spread with a wire-wound rod through a stencil mask (Metal Etch Services, San Marcos, CA) onto the cured substrate. Silver/silver chloride ink (E2414, Ecron) was mixed with Ecoflex-50 (Smooth On) at 94:6 weight ratio to increase the conductor stretchability[295]. The Ag/AgCl ink was annealed at 90°C for 5 minutes and the typical trace thickness was 30 μm. Subsequently, poly(3,4-ethylenedioxythiophene) polystyrene sulfonate (PEDOT:PSS for screen printing, Sigma-Aldrich) was printed through another stencil mask to cover the Ag/AgCl in the concentric ring region. The PEDOT:PSS formulation was 1 g of PEDOT:PSS with 0.4 ml of ethylene glycol, 5 μL of 4-dodecylbenzenesulfonic acid as a surfactant, and 40 mg 3-glycidoxypropyltrimethoxysilane as a crosslinker. By printing multiple passes, the

thickness of PEDOT:PSS was adjusted from 2.8 to 9.9 μm. In between printing passes, each layer was annealed at 100°C for 20 min. A layer of PDMS is used to encapsulate and isolate the Ag/AgCl interconnects from the skin. Currently thin wires were embedded in the elastomer substrate to make connections between the printed electrodes and the amplifier board, but in the future wireless communication chips may be embedded in the substrate to eliminate the dangling wires.

5.3.2 Wearable Screen-Printed Electrode Characterization

To measure the noise spectrum, each pair of test electrodes were placed side by side, with the active areas in contact with ~1 mm thick layer of conductive gel (Ten20, Weaver) connecting the two electrodes. Voltage noise between the electrode pair was measured by an OpenBCI board (Cyton biosensing board Kit, OpenBCI) set to an amplifier gain of 24. The reported noise was an average over 90 s at a sampling rate of 250 Hz. The background noise of the board amplifier was measured by shorting the two inputs. The time-domain voltage data was Fourier transformed with no additional signal filtering. The peaks at 30 and 60 Hz are due to power line noise.

The electrochemical impedance spectroscopy was taken with a potentiostat (Bio-Logic SP-200) in a two-electrode configuration supplying a sinusoidal signal of 10 mV without any dc bias. The electrode impedance was measured through a conductive gel, which was an uniform electrolyte sandwiched between the electrodes. Each electrode had an area of 0.5 cm^2. The Bio-Logic software EC-Lab was used for the equivalent circuit model fittings.

5.3.3 Wearable Screen-Printed Electrode In Vivo Recordings

A volunteer provided informed signed consent to participate in this study. The electrodes were connected to the OpenBCI board set to an amplifier gain of 24 sampled at 250 Hz. The sites for electrode attachment were swiped with 70% ethanol.

The EEG alpha rhythm oscillations were measured by placing the recording electrode on the left frontal area, at the F3 position in the 10-20 EEG position system. The reference and ground electrodes were attached at the two ear lobes, which were the A1 and A2 positions in the 10-20 EEG position system. The recording and reference electrodes were our dry electrodes directly adhered to the skin. For convenience with wiring, the ground electrode was a commercial gold electrode with conductive gel. To measure the alpha rhythm oscillations, the volunteer was asked to alternately open and close his eyes for 30 s periods. The EEG signals were digitally filtered in MATLAB by a sixth-order Butterworth filter with a range of 1 Hz to 40 Hz.

For EMG recordings, the volunteer was stationary and placed his left elbow on a desk. The reference electrode and the recording electrodes were our dry electrodes directly placed on the forearm. The volunteer was asked to clench his hand into a fist as EMG signals at E1 and E2 locations were simultaneously recorded by the OpenBCI board. The EMG signals were digitally filtered in MATLAB by a sixth-order Butterworth filter with a range of 5 Hz to 40 Hz.

For ECG recordings on the chest of a volunteer lying in supine position, the dry electrode was placed in the fourth intercostal space on the right chest near the sternum. The location is equivalent to the V_1 position in the standard 12-lead configuration. The Lead-I signal was measured using commercial 3M Ag/AgCl electrodes (T716) placed on the volunteer's left and right wrists, with the reference electrode on the left ankle. The ECG signals were digitally filtered in MATLAB by a 50th-order highpass filter with a cutoff at 0.5 Hz.

5.3.7 Substrate Preparation for Inkjet and Electrohydrodynamic Printing

Glass substrates were prepared by washing with detergent, followed by sequential cleaning with deionized water, acetone and isopropyl alcohol in an ultrasonicator. For flexible electrode fabrication, glass substrates were coated with a sacrificial layer of dextran as previously described[296]. Dextran (86 kDa, MP Biomedicals LLC) was prepared at 20% w/v and spin coated at 1500 rpm. Dextran films were then dried at 150 ºC on a hotplate.

5.3.8 Fabrication of Micropillar Electrodes

Silver nanoparticle ink, DryCure Ag-JB 1020B (C-Ink Co. Ltd) was used to inkjet print traces on the substrates with 10 μm drop spacing using a PSJet Printer fitted with a Dimatix printhead, with the stage heated to 55 ºC.

For printing of silver micropillars, a silver nanoparticle ink (Sigma) was loaded into a glass micropipette nozzle with internal diameter of 5-8 μm. An air pressure of 6-12 kPa was maintained after loading and the tip lowered to a distance of ~150 μm from the substrate. A printing voltage of 350-450 V was applied between the nozzle and conductive stage till ink was just prevented from dripping. A negative pulse of 100-300 V with rise time 100 μs, dwell time 300 μs and fall time of 0 μs was then applied between the substrate and nozzle to jet droplets. The stage was maintained at 45 ºC. Printed patterns were annealed at 300 ºC for 1 hr in a glove box.

After printed patterns were annealed, SU-8 2005 (MicroChem) or PDMS (Sylgard 184, Dow Corning) with a base catalyst ratio of 1:10 or a 8.3% solution of poly[4-vinylphenol] (PVP) in Propylene glycol monomethyl ether acetate (PGMEA) was spin coated at 3000 rpm. After spin coating SU-8, devices were soft baked at 95 ºC for 3 minutes followed by UV exposure up to a total exposure energy of 120 mJ/cm^2. Devices were then baked post-exposure at 95 ºC for 4 minutes and then hard baked at 150 ºC for 30 minutes. After spin-coating PDMS, devices were cured at 50 ºC overnight. After spin-coating with PVP, devices were cured at 200 ºC for 1 hour.

To release the flexible electrodes from the glass substrates, printed patterns coated with SU-8 or PDMS were placed in deionized water overnight to allow for dissolution of the dextran film.

5.3.9 Characterization of Micropillar Electrodes

Electrochemical impedance spectroscopy was taken with a potentiostat (Bio-Logic SP-200) in a two-electrode configuration supplying a sinusoidal signal of 10 mV without any DC bias. Two micropillar electrodes were connected by immersing in PBS.

5.4 Results

5.4.1 Fabrication of screen-printed hybrid, dry electrodes with concentric ring geometry

To fabricate these wearable electrodes, our choice of substrate was a moldable elastomer consisted of polydimethylsiloxane (PDMS) with a trace amount of ethoxylated polyethylenimine (PEIE), which modifies the crosslinking network to reduce the Young's modulus and roughen the substrate surface[282]. The substrate is semi-transparent, due to optical scattering by the micrometer-sized wrinkles on the surface. The micrometer-sized wrinkles on the PDMS-PEIE elastomer enhances van der Waals forces and increases the adhesion force to skin (**Figure 5.1.a**). The strong adhesion allows the electrodes printed on this substrate to be securely fastened to the skin and maintain good electrical contact without conductive gel.

The concentric ring electrode design (**Figure 5.1.a**) was patterned using a simple stencil-printing technique to directly deposit silver/silver chloride (Ag/AgCl) and PEDOT:PSS through a laser-cut stencil mask, and the geometry was based on the nine-point Laplacian estimation method[297] (**Figure 5.1.b**), which in our design was scaled to probe a dipole depth of ~1 cm beneath the electrode surface. The diameter of center disc was 4 mm. The inner diameter of the middle ring and the outermost ring was 6 mm and 14 mm, respectively, both with a linewidth of 1 mm. The Ag/AgCl traces were deposited to form highly conductive interconnects with the resistivity of 500 μΩ*cm. In the concentric ring region, the Ag/AgCl were covered with PEDOT:PSS to further lower the electrode impedance.

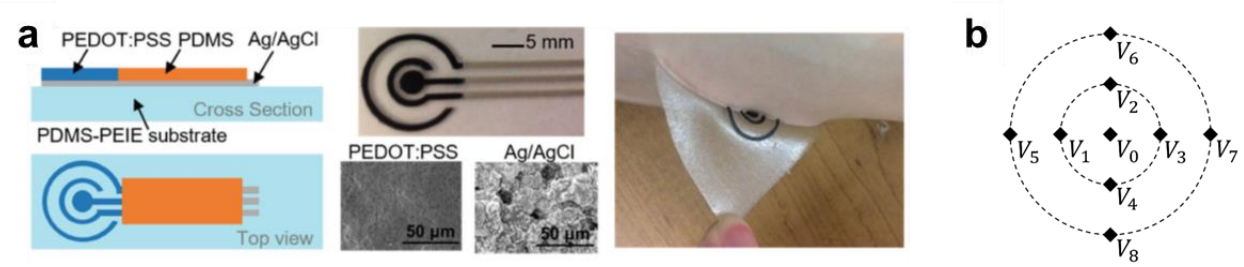


Figure 5.1: Concentric electrode design

(a) Schematic and photograph of a concentric ring electrode patterned by stencil printing. The scanning electron microscopy images show the surface morphology of a typical electrode covered with PEDOT:PSS or with bare Ag/AgCl. The printed dry electrode adheres well to skin as shown in the right photograph. **(b)** Schematics for the five-point and nine-point arrays

The concentric ring design is a physical approximation to the Laplacian filter, which is the second spatial derivative of the measured potentials and essentially assigns more weight to changes in bioelectric dipoles under measurement points, to enable better differentiation between concurrent, closely-spaced dipole sources[298,299]. While the Laplacian filter may be applied through computation on arrays consisting of discrete five-point or nine-point electrodes, the concentric ring approximation (**Figure 5.1.a)** simplifies orientation problems and significantly reduces the number of electrodes. Previously, concentric ring electrodes were demonstrated with silver or gold as the conductor interfacing with the human body. However, the metallic interfaces have high impedance and usually require a thin layer of conductive gel, which smears the spatial resolution and may also short the electrode rings if the gel bridges them. Thus practical implementation of the Laplacian filter requires a dry interface[300–302], which requires further materials and device development to match the signal-to-noise (SNR) level of standard gel electrodes.

5.4.2 Characterization of concentric ring electrodes

We then compared the voltage noise of our PEDOT:PSS-covered electrodes to commercially available Au and Ag/AgCl electrodes with hydrogel from 3M (**Figure 5.2.a**). The electrode impedance was measured through a conductive gel, which was an uniform electrolyte on top of the pair of electrodes. Since the power spectral density[303,304] S_N of voltage noise v_N is

inversely proportional to square root of the electrode area A by $S_N = v_N / A^{0.5}$, the noise spectra of the 3M electrode and the printed electrodes are scaled by area for comparison to the Au electrode with a diameter of 1 cm (**Figure 5.2.a**). The electrode diameters were 1.5 cm for the 3M electrode and 0.8 cm for the printed electrodes, which were made in disc shapes and not concentric rings to facilitate comparison. Among the various electrodes including commercial ones, our printed electrode with a 9.9 μm PEDOT:PSS layer shows the lowest noise voltage across the spectrum.

The excellent noise level of the printed electrodes is attributed to the low interfacial impedance of the organic PEDOT:PSS conductor. Noting the reduction in noise voltage with a thick PEDOT:PSS layer, we varied the PEDOT:PSS thickness and measured the electrode impedance by electrochemical impedance spectroscopy (**Figure 5.2.b**). The impedance of bare Ag/AgCl electrode is at least two times higher than the ones with PEDOT:PSS. As the PEDOT:PSS film thickness is increased from 2.8 μm to 9.9 μm, the impedance phase reveals a transition to higher capacitance at low frequency. The impedance magnitude for electrodes with 6.8 μm and 9.9 μm PEDOT:PSS are very similar for frequencies above 0.1 Hz, indicating that further increase beyond 9.9 μm is not necessary for impedance reduction and hence the printed electrodes for electrophysiological recordings were fabricated with PEDOT:PSS film with thicknesses ranging from7 μm to 9 μm.

The electrode impedance characteristics were analyzed by fitting to an equivalent circuit model[305,306] (**Figure 5.2.b**), where R_s is the series resistance, R_t is the charge transfer resistance, Z_{CPE} is the constant phase element, and Z_w is the semi-infinite Warburg diffusion element related to ion movement in PEDOT:PSS. The fit results are summarized in **Table 5.1**. The constant phase element[307] accounts for non-idealities in capacitance and is expressed by $Z_{CPE} = \frac{1}{Q}(j \cdot \omega)^{-n}$, where $j = \sqrt{-1}$, $0 < n < 1$, ω is radial frequency, and Q is a constant with the dimension of farads multiplied by seconds to the power of (n-1). If n=1, the expression is reduced to the case of an

ideal capacitor. The exponent *n* is affected by roughness or porosity, which explains the lower *n* value for the rough Ag/AgCl surface as seen in Fig. 1(a). The Warburg diffusion element is expressed by $Z_w = R_w/\sqrt{\omega} + R_w/(j \cdot \sqrt{\omega})$, where R_w is the Warburg resistance, ω_0 is a characteristic radial frequency. In our case Z_w is essentially negligible to electrophysiological measurement as it affects the impedance only at very low frequency range < 0.5 Hz.

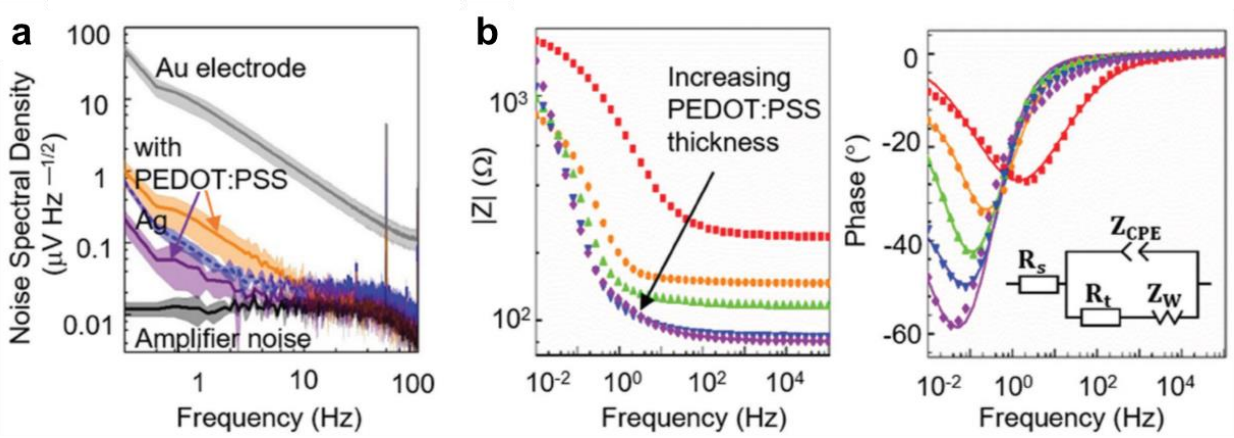


Figure 5.2: Noise and impedance characterization of electrodes

(a) Noise power spectral density of various electrodes. Error bars were indicated by the shaded area and calculated from averaging over a sampling period of 90 s. Gray line: commercial gold electrode. Orange line: printed electrode where the PEDOT:PSS thickness is 2.8 μm. Blue dashed line: commercial Ag/AgCl with hydrogel (3M part T716). Purple line: printed electrode where the PEDOT:PSS thickness is 9.9 μm. Black line: amplifier noise. **(b)** Impedance magnitude and phase angle of the printed electrodes. The thickness of PEDOT:PSS layer is 0 μm (red, bare Ag/AgCl), 2.8 μm (orange), 4.5 μm (green), 6.8 μm (blue), and 9.9 μm (purple). The error bars for the film thicknesses are provided in Table 1. The inset shows the equivalent circuit model used for parameter fittings (solid lines).

Table 5.1: Fit values to the equivalent circuit model for the printed electrodes

PEDOT:PSS thickness [μm]	R_S [Ω]	R_t [Ω]	Q [mF s^{n-1}]	n	R_W [Ω$s^{-1/2}$]
Bare Ag/AgCl	235	1636	0.47	0.60	-
2.8 ± 0.3	148	623	2.50	0.82	28
4.5 ± 0.5	119	916	4.13	0.84	33
6.7 ± 0.8	89	1000	5.16	0.85	117
9.9 ± 1.0	87	1788	6.06	0.88	151

With increasing PEDOT:PSS thickness, the electrode series resistance is reduced. The charge-transfer resistance, along with the constant phase element, increases and indicates capacitive characteristics becoming more dominant with thicker PEDOT:PSS films. The current measured by impedance spectroscopy is consisted of two components, where the Faradaic contribution is caused by electrochemical reaction on the electrode surface, and the non-Faradaic component is due to charging of the double-layer capacitance at the electrode-electrolyte interface. As the charge-transfer resistance increases, there is less Faradaic contribution. Meanwhile, the non-Faradaic component increases, because thicker PEDOT:PSS films are still permeable[280] to ions and larger surface areas enhance the electric double layer effect at low frequencies. Overall, the low interface impedance of the printed electrodes is suitable for bio-signal detection, since lower impedance would translate to higher SNR. In addition, the large capacitance is also advantageous if the electrodes are to be used for electrostimulation.

The change in conductor resistance was examined as the elastomeric substrate was stretched uniaxially along the length of printed lines with linewidth of 0.8 mm and thickness of 30 μm (**Figure 5.3.a**). The conductor samples were straight lines for the purpose of determining the stretchability limit of the printed Ag/AgCl blended with 6% elastomer. Higher tolerance to strain would be achieved if the conductor traces were designed with other patterns such as serpentines or fractals. The fabrication process here will be adaptable for implementing these geometries if an increase in strain tolerance is needed. Without any geometric optimization, the printed Ag/AgCl interconnect remained conductive up to 30% strain, which met the requirement for tolerating the chest volume expansion during breathing. The Ag/AgCl line was cracked and became open at 36.7% strain. The Ag/AgCl trace resistance R increased up to seven-fold from 0 to 30% strain (**Figure 5.3.b**), and the resistance recovered to the original value R_0 upon release (**Figure 5.3.c**). For the conductor line where 9 μm of PEDOT:PSS was coated on top of the Ag/AgCl, the resistance change with stretching was higher than the line without the coating, because the

PEDOT:PSS layer showed cracks by 15% strain. Therefore, we confined the PEDOT:PSS coating to the electrode areas for recordings and left the rest of the interconnect traces as Ag/AgCl to minimize resistance change with stretching.

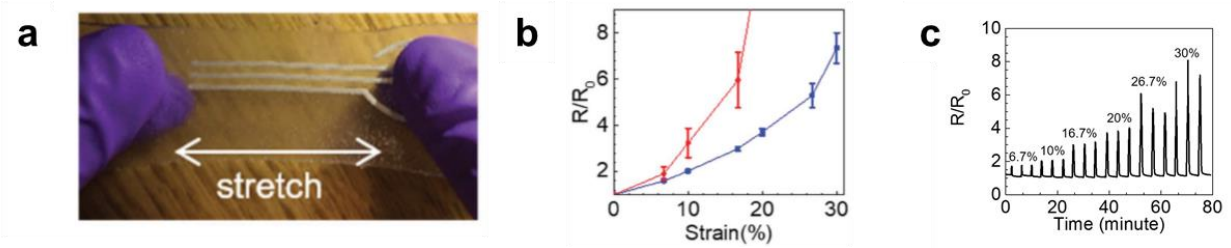


Figure 5.3: Stretchability of electrodes

(a) Photograph of printed lines being stretched linearly to 30% strain **(b)** Change in resistance of linear conductor traces under uniaxial stretching. Ag/AgCl blended with 6% elastomer, with 9 μm PEDOT:PSS coating (red) and without coating (blue). **(c)** Change in resistance of linear conductor traces under uniaxial stretching. The conductor line was Ag/AgCl blended with 6% elastomer and stretched at the indicated strain value for 3 repetitions.

The electrodes on PEIE-modified PDMS substrate adhered to skin by van der Waals force, allowing the electrodes to be deployed without conductive gel and adhesives that may be irritating for long-term wear. Stretching may initiate delamination at the edge of the substrate, but prior work[308] has shown that the delamination would be mitigated by decreasing the substrate thickness. The substrate stayed in place for several hours during light activities such as walking or carrying out EMG movement tasks.

5.4.3 Application of concentric ring electrodes for electrophysiological measurements

The dry electrodes can be integrated into one band to incorporate multiple measurement channels (**Figure 5.4.a**). Our substrate was connected to a stretchable textile to put the whole device under tension to keep the headband in place. We had placed a band of printed electrodes on a volunteer's forehead to measure alpha rhythm oscillations, which are prominent 10 Hz signals shown in the EEG spectrogram. The oscillations were detected when the person closed his eyes and gathered his concentration; otherwise, the signal disappeared, for example between 0-30 s and 60-90 s (**Figure 5.4.b**). The alpha rhythm oscillations occurred in the occipital lobe far

from forehead where the electrodes were located. Since the alpha rhythm signals were not localized at the forehead, there was no significant difference between the recordings of the left and right electrodes on the band. However, in EMG and ECG measurement below, the spatial resolution of the signals were enhanced and local signals were differentiated by the concentric ring electrodes.

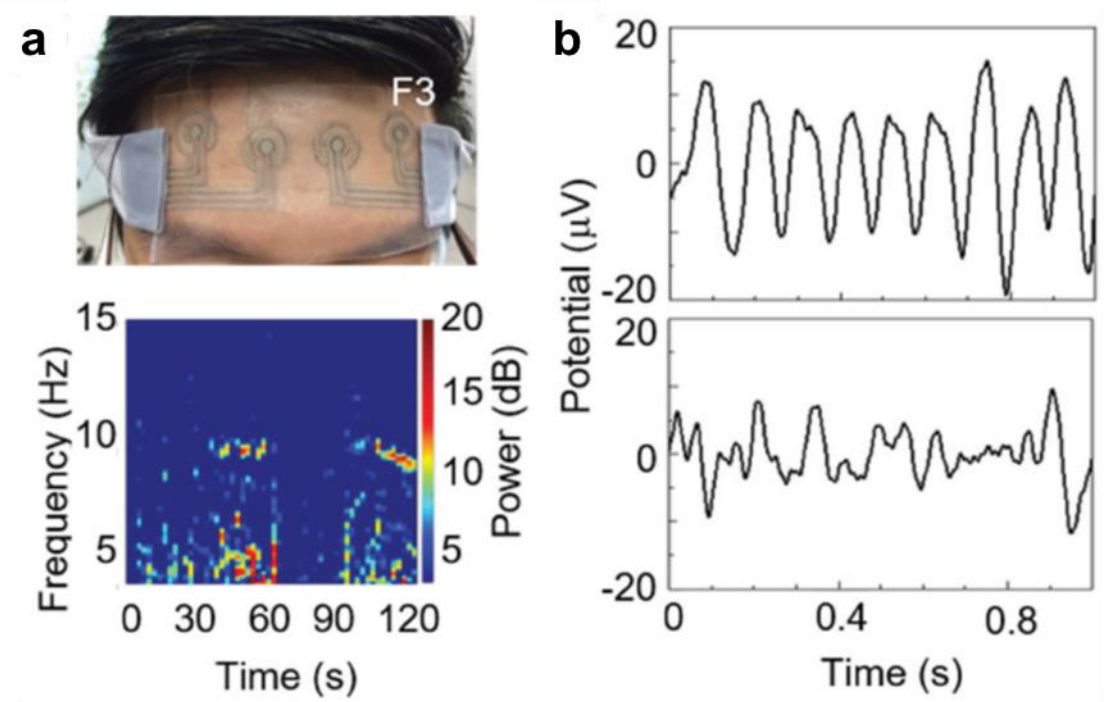


Figure 5.4: EEG measurement with dry electrodes

(a) Photograph of the printed dry electrodes incorporated into one headband. EEG spectrogram of alpha rhythm oscillations recorded by the dry electrode at F3 position according to the 10–20 system. **(b)** Time traces of the EEG recording during the period with (top) and without (bottom) alpha rhythm oscillations.

The concentric ring electrodes were configured to take differential measurement between the middle ring and the central electrode, and also between the outer ring and the central electrode. The same electrodes were used as conventional “disc” electrodes by taking measurement with only the central electrode, with respect to a reference electrode placed at locations at which negligible EMG activity is expected, for example at the bony part of the volunteer’s wrist for EMG.

With the electrode design, there were two options to apply the Laplacian filter, namely the bipolar and tripolar configurations[298]. The bipolar configuration used the five-point approximation commonly used for edge detection; the bipolar Laplacian potential was the difference between the outer ring and the central electrode, with respect to the area inside the outer ring. The tripolar configuration extended the filter to the nine-point approximation to increase localization, with a tradeoff in adding complexity with an additional middle ring (Supplemental Fig. S1). The approximation expressions[298] for the two cases are

Bipolar Laplacian potential: $L_{Bp} = \frac{4}{(2r)^2}(V_O - V_D)$

Tripolar Laplacian potential: $L_{Tp} = \frac{1}{3r^2}[16(V_M - V_D) - (V_O - V_D)]$

where V_D, V_M and V_O are the potential measured from the inner disc, the middle ring and the outer ring of an electrode, respectively, and *r* is the distance between the center of disc and middle ring. In our design the spacing between the center of disc and outer ring is 2*r*.

The EMG measurement (**Figure 5.5.a**) compares the signals obtained from two dry electrodes placed 45 mm apart on the forearm of a volunteer. As the volunteer clenched his hand into a fist, neuromuscular potential was detected by both electrodes. Using the conventional measurement technique, the two electrodes simultaneously recorded very similar signals (**Figure 5.5.b**), and as such implied that the spatial resolution was very low and the signals were practically indistinguishable at distance <45 mm. Upon changing to the Laplacian differential method, in which the signals at the concentric rings were measured with respect to the central disc, the EMG signal at the location of electrode E2 was shown to be stronger than at location E1, since E2 is closer to the neuromuscular potential source that activated the clenching movement (**Figure 5.5.a-b**). The Laplacian measurement distinguished the signal source location, offering better spatial resolution than the approach based on comparing potentials of discrete electrodes.

The signal-to-noise ratio for the bipolar and tripolar EMG signal was calculated by the equation $SNR = 20 * \log \sqrt{\sum_i Sf_i^2 / \sum_i (Sr_i - Sf_i)^2}$, where *Sr* is the raw signal and *Sf* is the filtered signal. The SNR from the tripolar configuration is higher than that of the bipolar one in (**Figure 5.5.c**) with respect to an equivalent area of $\pi(2r)^2$. If the bipolar Laplacian potential is measured with respect to πr^2, the SNR at the E2 location increases to 51.2 dB (**Figure 5.5.c**), because the signal is localized within the area within the middle ring. Overall the signals were distinguishable at the spatial resolution of *r*= 6 mm, corresponding to the distance between the middle ring and the disc center of the printed electrode. The signal from the concentric ring electrode is compared to the pair of conventional "disc" electrodes with the same sensing area. This spatial contrast of 6 mm from using concentric rings is much better than the previous limit from using disc electrodes, where signals at 45 mm apart were hardly distinguishable.

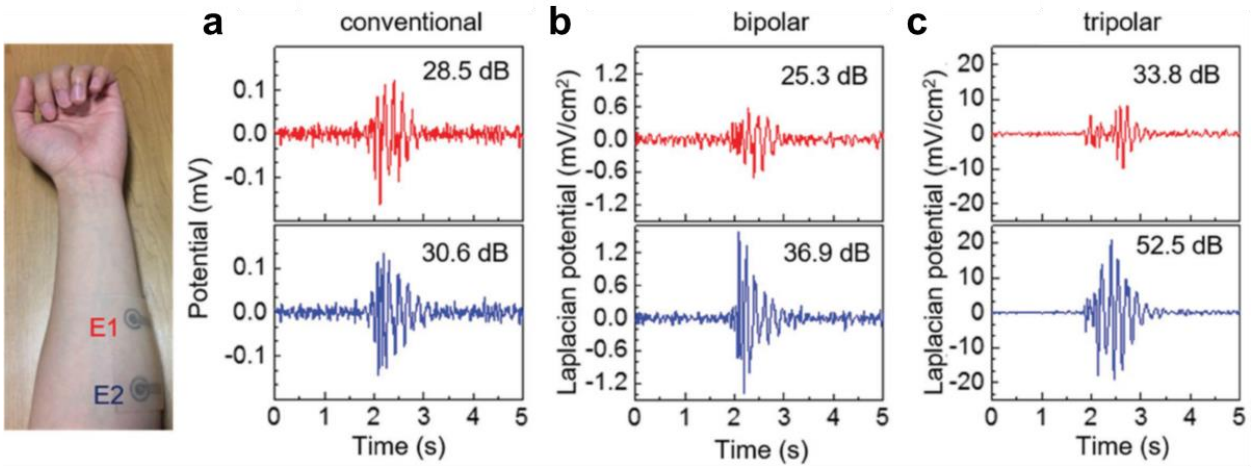


Figure 5.5: EMG measurement with dry electrodes

Photograph of the dry electrode locations in EMG recordings. The EMG signals measured **(a)** conventionally or with Laplacian filtering in **(b)** bipolar or **(c)** tripolar configuration. The SNR values are listed in the upper right corner of each plot.

For ECG, we used our dry concentric ring electrodes to measure the potential difference $V_O - V_D$ (in units of μV) for comparison with signals from the standard 12-leads measurement. There was no significant difference in signals of $V_O - V_D$ and $V_M - V_D$ (shown in Supplemental Fig.

S4), since the distance between the concentric rings used here is small (r= 6 mm) in relation to cardiac physiology. Here we did not compute Laplacian potentials (in units of $\mu V/cm^2$), because typical ECG analysis does not require spatial derivatives.

The visualization of the PQRST peaks is very important for clinical diagnosis. It is especially challenging to identify P peaks as they are the lowest of the ECG waves. P-waves provide clues to atrial activities, but the signals are weak because there are fewer cardiac cells involved in atrial activation than in other ventricular activities. The conventional Lead-I signal was measured by monopolar commercial gel electrodes connected to the limbs (**Figure 5.6.a**). Conventional Lead-I results typically show the highest P-waves when compared to measurement at V_1 position[299], but the limb positions are inconvenient for movement. With the compact dry electrode, the signal was taken at a convenient core position. This signal $V_O - V_D$ from concentric ring electrodes was small, because the distance between the outer rings and the central disc was much shorter than the distance across the arm span (> 1 m) between the Lead-I electrode and its reference electrode. Nevertheless, in terms of the P-wave amplitude with respect to the full ECG peak-to-peak amplitude, the signal provided similar P-wave contrast (**Figure 5.6.b**). Whereas the P-wave was at 13% for the Lead-I signal normalized by a peak-to-peak amplitude of 342 μV, the P-wave was at 18% for dry electrodes signal normalized by 34 μV (**Figure 5.6.b**). Hence, after amplitude normalization, the signals obtained by concentric ring electrodes maintain good contrast for identifying P-waves associated with atrial activity. The compact concentric ring configuration is reliable and comparable to the conventional limb leads and provides an additional advantage of eliminating dangling wire connections that impede mobile monitoring.

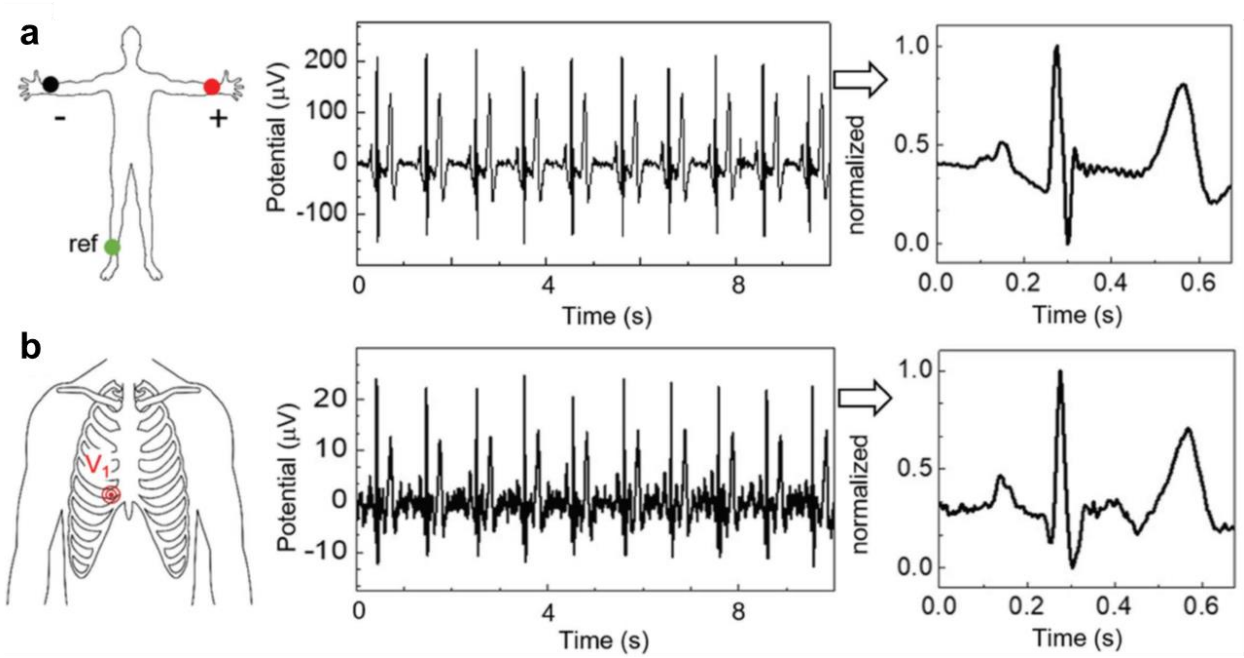


Figure 5.6: ECG measurement with dry electrodes

ECG electrode locations are indicated by the cartoon schematics. **(a)** Signal of Lead-I based on the standard 12-lead protocol. **(b)** Signal measured by dry concentric ring electrodes at the core V1 location.

5.4.6 Printed Flexible Microelectrodes

For measuring signals at the cellular level, we devised a novel method combining inkjet and electrohydrodynamic printing, and used a previously demonstrated sacrificial layer etch method[294] to transfer the printed patterns onto a flexible substrate. Rigid substrates were prepared for printing by cleaning and then spin coating with a layer of dextran, an aqueous-soluble polysaccharide. The solvent was then removed by drying the coated substrate at 150 ºC **(Figure 5.7.a)**. Dextran was chosen here as the sacrificial layer due to its water-soluble nature, which makes it compatible with elastomers by avoiding high-temperature processing steps and cell culture by avoiding the use of harsh solvents. Dextran coated substrates are also suitable for high fidelity deposition and release of patterns as compared to other water-soluble polymers[309,310].

Longer elements, namely traces for connection with external readout circuitry were patterned on the dextran coated substrates by inkjet printing of a colloidal silver ink **(Figure 5.7.a)**. While lower resolution than other methods, inkjet printing allows for high-speed, multi-pass printing yielding low impedance patterns. This was followed by the printing of micropillars at one end of the traces using electrohydrodynamic (EHD) printing **(Figure 5.7.b)**. Low frequency EHD printing allowed for the printing of high-resolution vertical structures. Once the patterns were printed, the elastomer or insulating polymer material was spin-coated on the printed pattern to leave only the pillars exposed **(Figure 5.7.c)**, and the insulating material was crosslinked. The fabricated patterns were then either used as rigid microelectrodes or released by dissolving the dextran in water, leaving the printed pattern embedded in the spin-coated elastomer layer **(Figure 5.7.d)** to form flexible microelectrode arrays.

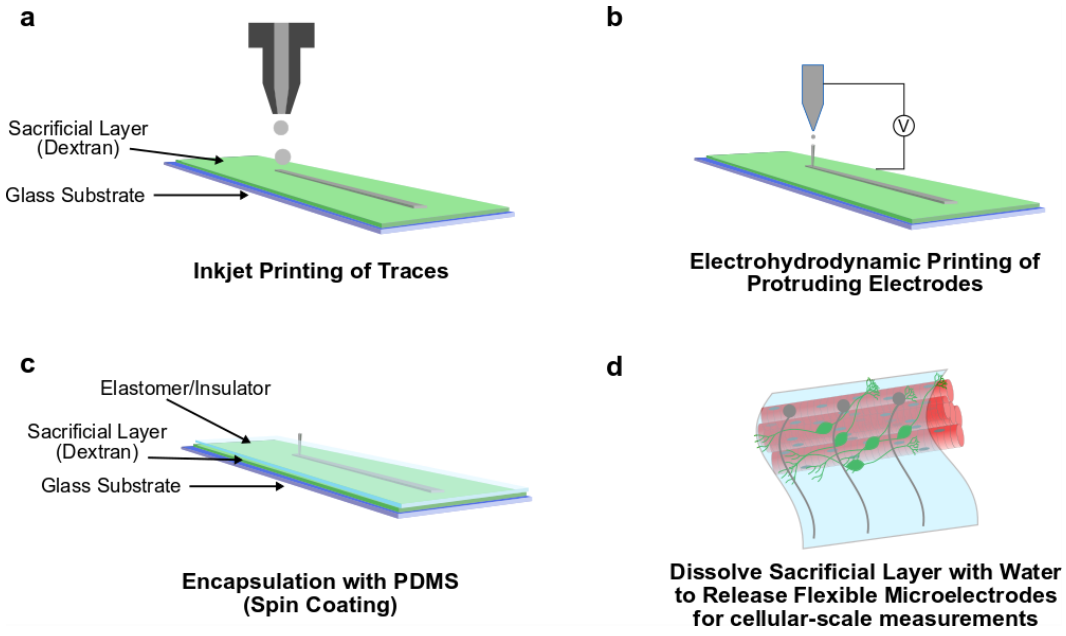


Figure 5.7: Schematic for fabrication of micropillar electrodes

(a) Traces were inkjet-printed using a colloidal silver ink on glass substrates coated with dextran. **(b)** Micropillar electrodes were printed at the end of the inkjet printed traces using electrohydrodynamic printing **(c)** A polymer or elastomer insulator layer was spin-coated on the electrodes such that only the micropillars were left exposed. **(d)** The sacrificial dextran layer was dissolved by immersing the fabricated device in deionized water overnight.

The printed electrodes **(Figure 5.8.a)** were then imaged using scanning electron microscopy, where we observed intact vertical micropillars covered by the elastomer **(Figure 5.8.b)**. We also completed a preliminary assessment of the electrodes using electrical impedance spectroscopy in the two electrode configuration, where the electrodes were immersed in a PBS **(Figure 5.8.c)**.

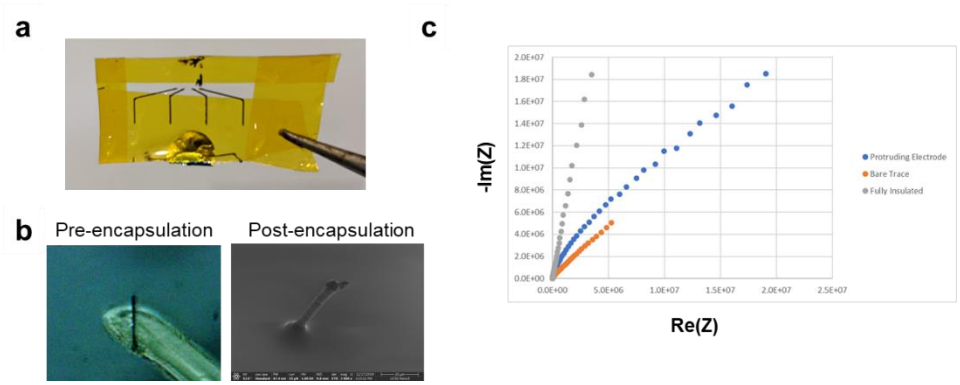


Figure 5.8: Characterization of micropillar electrodes

(a) Image of a representative flexible electrode array. **(b)** Optical image of a micropillar on an inkjet printed trace prior to elastomer encapsulation (left) SEM image of an encapsulated micropillar electrode (right). (c) Electrical impedance spectroscopy curve for micropillar electrode (blue), bare inkjet printed trace (orange) and fully insulated trace (grey)

For the insulating materials and elastomers, we tested polydimethylsiloxane (PDMS), SU-8 and poly[4-vinylphenol] (PVP). Each of them suffered from certain drawbacks and further testing and optimization is required to refine the choice and engineering of substrate materials. The Young's modulus of PDMS (0.57 MPa to 3.7 MPa)[311] did not confer sufficient rigidity to the released PDMS film and caused it to collapse without sufficient stiff backing such as polyimide **(Figure 5.8.a)**. Secondly, PDMS does not adhere well to the glass and we observed that the substrate as well can printed patterns can be damaged by mechanically active tissue like cardiomyocytes. On the other hand the Young's modulus of SU-8 (4.1 GPa)[311] is sufficiently high to confer structural rigidity to the released film, but the low wettability of SU-8 necessitates

dropping SU-8 on the whole substrate and the resist seems to not be fully released from the pillars, leading to insulator-like impedance. Finally, PVP has low viscosity and is able to insulate the inkjet printed traces while exposing the micropillars but the films are not thick enough to be released and the devices can only be used as rigid devices.

5.5 Discussion

In conclusion, we demonstrate printed dry electrodes on stretchable substrates that are moldable to conform and provide good adhesion to skin for electrophysiological recordings of EEG, EMG, and ECG, and for cutaneous stimulation during wound healing. These electrodes with a PEDOT:PSS interface layer are optimized to a thickness that shows lower noise level than that of commercial gel electrodes. The concentric ring geometry enables Laplacian filtering to pinpoint the bioelectric potential source with spatial resolution determined by the ring distance. The EMG and ECG recordings show enhanced localization of signal sources and the measurement using concentric ring electrodes provides significant advantage in being much more simple and easy to set up compared to using multiple discrete electrodes to extract Laplacian potentials. This work shows a new fabrication approach to integrate and create designs that enhance spatial resolution for high-quality electrophysiology monitoring devices at low cost.

We also present a preliminary demonstration of a novel method to fabricate flexible, printed microelectrodes. This method overcomes the challenges associated with resolution and wettability of elastomeric substrates associated with many printing technologies, by using electrohydrodynamic printing to define the minimum dimension and printing on a sacrificial layer on a rigid substrate. This method, with further development and optimization, could enable the production of microelectrode arrays in easily reconfigurable geometries to monitor and stimulate engineered tissues and organoids to assess function, drive maturation and study systems under electrical perturbation.

5.6 Acknowledgements

This work was partially funded by NextFlex grant (042299), NSF CMMI 1635729 and UC San Diego Medical Devices and Systems Initiative funded by the Electrical and Computer Engineering Department. T.P. was partially supported by the IBM PhD fellowship. We thank Zhenghui Wu for his experimental assistance and helpful discussions. Part of the work was performed at the San Diego Nanotechnology Infrastructure of UCSD, which is supported by NSF ECCS-1542148.

Chapter 5 in part is a reprint of the following materials: Wang, K., Parekh, U., Pailla, T., Garudadri, H., Gilja, V. & Ng, T.N. Stretchable Dry Electrodes with Concentric Ring Geometry for Enhancing Spatial Resolution in Electrophysiology. *Advanced Healthcare Materials* 6 (19), 1700552 (2017). The dissertation author is one of the two primary authors of the study.

Chapter 6: Flexible Printed Electrodes for Stimulation

6.1 Abstract

Wound healing is a complex process involving diverse changes in multiple cell types where the application of electric fields has been shown to accelerate wound closure. To define the efficacy of therapies based on electric fields, it would be valuable to have a platform to systematically study the effects of electrical stimulation upon the inflammation phase and the activation of signaling mediators. Here, we report on an *in vivo* electrical stimulation model in which flexible electrodes were applied to an animal model for monitoring inflammation in a wound. Subcutaneous implants of polyvinyl alcohol sponges elicited inflammation response as defined by the infiltration of leukocytes. The wound site was subjected to electric fields using two types of additively fabricated flexible electrode arrays. The sponges were then harvested for flow cytometry analysis to identify changes in phosphorylation state of intracellular targets. This platform enables studies of molecular mechanisms, as it shows that an application of low frequency electrical stimulation $\leq$0.5 Hz increased phosphorylation of Erk proteins in recruited leukocytes, identifying a signaling pathway that is activated during the healing process.

6.2 Introduction

Electrical stimulation has been found to be beneficial for the treatment of several disease and injury states of electrically active tissue, such as for muscular injury[312,313], stroke[313–318], spinal cord injury[319] and neurological disease[320]. Conformal and flexible electrodes with high ionic charge injection capacity are desirable for such applications as well. Such electrodes can match the stiffness of the contacted tissue, enable placement of electrodes at appropriate locations, reduce tissue damage, and enable sustained stimulation.

Acute and chronic wounds are a significant healthcare burden, and technologies for wound management[321–323] are highly desired. Since the endogenous electric field at wound sites is a critical driver of the healing process[324,325], electrical stimulation (ES) of wound sites has been proposed as a means to accelerate the wound healing process and supplement the standard of care[326–329]. ES is thought to reduce the chances of infection, increase angiogenesis, and accelerate cell migration. The effects of ES on cutaneous wound healing in animal[330] as well as clinical studies[326,328,331] have demonstrated that there is an unmet need for standardized platforms to assess the efficacy of specific parameters in wound healing. Here we focus on the inflammation phase. While *in vitro* studies have demonstrated the effects of electric fields on cell migration and dissected parts of the molecular machinery driving these effects[332–334], *in vivo* studies have focused primarily on large scale changes like wound area and angiogenesis[327,328]. *In vivo* studies to gain insights into the molecular mechanisms of these effects have been limited, and a platform for uncovering the complex interactions driven by ES *in vivo* would be a valuable complementary tool to gain further mechanistic insights into the inflammation component of the wound healing process.

Here we report a platform that enables sensitive *in vivo* studies of changes in protein levels upon electrical stimulation at wound sites on mice. The platform consisted of a polyvinyl alcohol (PVA) sponge which was implanted subcutaneously at the wound site. Instead of tissue

extraction, this PVA sponge enabled the capture and investigation of infiltrating immune cells[335–338]. Using subcutaneous implants of PVA sponges, we were able to establish data in animal models that can be measured in a laboratory setting. Such PVA sponges could be used as subcutaneous implants in humans with appropriate controls in a clinical setting[339]. The additively printed[340,341] ES electrodes used in this study were flexible with enhanced adhesion to skin, allowed precise control over stimulation locations, and were optimized for improved charge injection capacity[342,343]. As prior meta-studies[326,328,331] show that there is ongoing debate over the types of ES parameters necessary for wound treatment, the magnitude and frequency of ES waveforms were varied here to examine their effects on immune cells in mice wound models.

To assess the effects of ES, the cells captured in the PVA sponge were analyzed via flow cytometry, which enabled high-throughput measurement of cell populations and intra-cellular protein levels at the single cell level[344]. The multi-parametric flow cytometry analysis allowed us to isolate signals from specific sub-populations of cells, to rapidly distinguish different cell types in particular immune cells and investigate cell-type specific signal transduction pathways.

We demonstrate the versatility of the platform by investigating markers for two different signaling pathways that mediate immune cell response to ES, specifically phosphorylation-driven signaling events. Phosphorylation plays a key role in regulating protein activities, and plays an important role in driving signaling cascades[345] which guide cellular response to external stimuli. We focus on the Mapk (mitogen activated protein kinases)/Erk (extracellular signal-regulated kinases), and PI3k (phosphatidylinositol 3-kinase)/Akt (protein kinase B) pathways, which are known to play roles in immune cell activation and development[346–351], and also known to play a role in cell migration during wound healing[333,350].

6.3 Materials and Methods

6.3.1 Wearable Screen-Printed Electrode Fabrication

Electrodes were fabricated as detailed previously[340]. The substrate was polydimethylsiloxane (PDMS from Sylgard 184, Dow Corning) modified with a solution of 80% ethoxylated polyethylenimine (PEIE, Sigma-Aldrich). The PDMS base and curing agent were mixed in a 10:1 ratio by weight, and 35 µl of PEIE was added per 10 g of the PDMS mixture to reduce the Young's modulus and roughen the surface to increase adhesion by Van der Waal force.[282] This mixture was degassed and poured into a mold to cure at 90 ºC for 3 hr.

To pattern electrodes, conductive pastes was applied through a laser-etched stencil mask (Metal Etch Services, San Marcos, CA) using a blade coating applicator (Gardco). Each electrode had a diameter of 3 mm, while the center-to-center distance of the inner to outer electrodes was 9 mm. The interconnects and electrodes were patterned via screen printing through metal etched stencil masks. The silver/silver chloride ink (E2414, Ecron) was mixed with Ecoflex-50 (Smooth On) at 94:6 weight ratio to increase electrode stretchability. This ink was annealed at 90 ºC for 5 min. The poly(3,4-ethylenedioxythiophene) polystyrene sulfonate ink was formulated with 1 g of screen-printable PEDOT:PSS mixed with 0.4 ml of ethylene glycol to increase electrode conductivity, 5 µl of 4-dodecylbenzenesulfonic acid as a surfactant, and 40 mg 3-glycidoxypropyltrimethoxysilane as a crosslinker. This PEDOT:PSS ink was deposited through a second stencil mask to cover the ends of the Ag/AgCl traces and cured at 120 ºC for 1 hr. A layer of PDMS was deposited on top of the Ag/AgCl traces to encapsulate and isolate them from the skin. Thin wires were embedded in the encapsulation layer to connect the printed electrodes to the power supply and measurement equipment for stimulation.

6.3.2 Wearable Inkjet-printed Electrode Fabrication

Inkjet-printed electrodes were fabricated as detailed previously[341]. Harima gold nanopaste ink was printed on polyethylene naphthalate substrates with 30 μm drop spacing using a Dimatix Materials Printer (DMP-2800) to form the gold interconnects. Printed gold nanopaste ink was cured with a slow ramp annealing step (30–230 ºC with 0.7 ºC/min ramp), followed by a constant temperature bake at 230 ºC for 1 hr. Vias were then laser-cut in a Teflon tape and the tape was placed on top of the printed gold electrodes to act as the encapsulation layer.

To test charge injection capacity, we compare electrodes with or without PEDOT:PSS coated on top of the electrode ends. And because of this test, we use Au for the exposed electrodes instead of Ag, since Ag would be oxidized in the presence of aqueous solutions or hydrogels when ES was applied. Nonetheless, if the Ag was covered by PEDOT:PSS, the oxidization issue is negligible and not observed. To coat PEDOT:PSS on the electrodes, a Kapton tape was laser cut to be used as a mask for stencil printing, and placed on the electrode array with openings aligned to the electrode ends. PEDOT:PSS was then stencil-printed through this mask and annealed at 100 ºC for 10 minutes. The Kapton tape was then gently peeled off from the Teflon surface, leaving the intact electrode array and encapsulation layer. To increase the adhesion of the arrays on wound sites, a double sided tape was attached to the electrodes. Vias to the electrodes were formed by a laser cutter.

For impedance mapping, low impedance hydrogels were made with 10% (w/v) gelatin dissolved in phosphate buffered saline (PBS), and high impedance hydrogels were made with 10% (w/v) gelatin dissolved in distilled water. Gels were warmed to 37 ºC and deposited side by side on a glass slide for impedance measurements.

6.3.3 PVA Sponge Implant and Harvest

Implantation of PVA sponges (PVA Unlimited, Warsaw, IN) was performed in 6-8 week old male C57BL/6 mice as described in Refs. [335] and [352]. PVA sponges were hydrated in

phosphate buffered solution, autoclaved and then aseptically implanted subcutaneously in the back of the mice. Briefly, hair was shaved from the back of the mouse, an incision made to lift the skin, and the PVA sponge placed subcutaneously, after which the incision was sutured closed.[352] Three days after sponge implantation, mice were electrically stimulated with constant-current square-wave pulses for 1 hr. All mouse procedures were approved by the UCSD Institutional Animal Care and Use Committee. Sponges were removed after stimulation experiments and processed to wash out cells for flow cytometry analysis by collecting sponges in 1 mL sterile PBS, the sponge compressed gently to release cells for subsequent fixation, permeabilization, and flow cytometry analysis.

6.3.4 Electrical Stimulation

Each individual electrode was coated with a drop of conductive hydrogel (SignaGel, Parker Laboratories Inc.) with a pipette tip immediately before use. The gel layer ensured electrical contact between the mice skin and the electrodes. Mice with implanted PVA sponges were then anesthetized using isofluorane. Once the animals were not responsive, they were placed in nose cones for the stimulation procedure. The skin around the wound site was cleaned with isopropanol wipes. The electrode array was placed on the wound site, such that the central electrode was aligned on top of the implanted PVA sponge, while the outer ring of electrodes surrounded the wound area (Fig. 1C). The electrode array was secured via surgical tape, with even pressure applied over the wound site to keep the array in place during the course of the stimulation. Sham mice underwent the same procedure but were not electrically stimulated, as the electrode arrays were disconnected from the power supply. For mice subjected to electrical stimulation, the electrodes were connected to a Keithley 2400 Source Meter (Tektronix). Constant-current square-wave pulses of the desired frequency were programmed via custom software implemented in LabView (National Instruments). The central electrode was connected to ground, while the outer electrodes were all connected to the positive terminal of the power

supply. Stimulation was applied for 1 hr before the PVA sponge was harvested for further processing. The applied current was limited to a maximum of 600 μA, and the voltage was limited to below 5 V. Thus the applied electric field was up to 1 V/mm.

6.3.5 Immunostaining and Flow Cytometry

Cells harvested from the PVA sponges were analyzed via flow cytometry. Harvested cells were fixed and permeabilized cells, then labelled with CD45-VioGreen (130-110-803; Miltenyi Biotec, Bergisch Gladbach, Germany), CD11b-APC-Vio770 (130-109-288; Miltenyi Biotec), Ly6C-APC (130-102-341; Miltenyi Biotec) and either P-Erk (Cell Signaling, #14095) or P-Akt (Cell Signaling, #5315). For analysis of viability, a subset of the harvested cells was labelled with propidium iodide (130-093-233; Miltenyi Biotec) without fixing and permeabilization and analyzed via flow cytometry.

6.3.6 Flow Cytometry Analysis

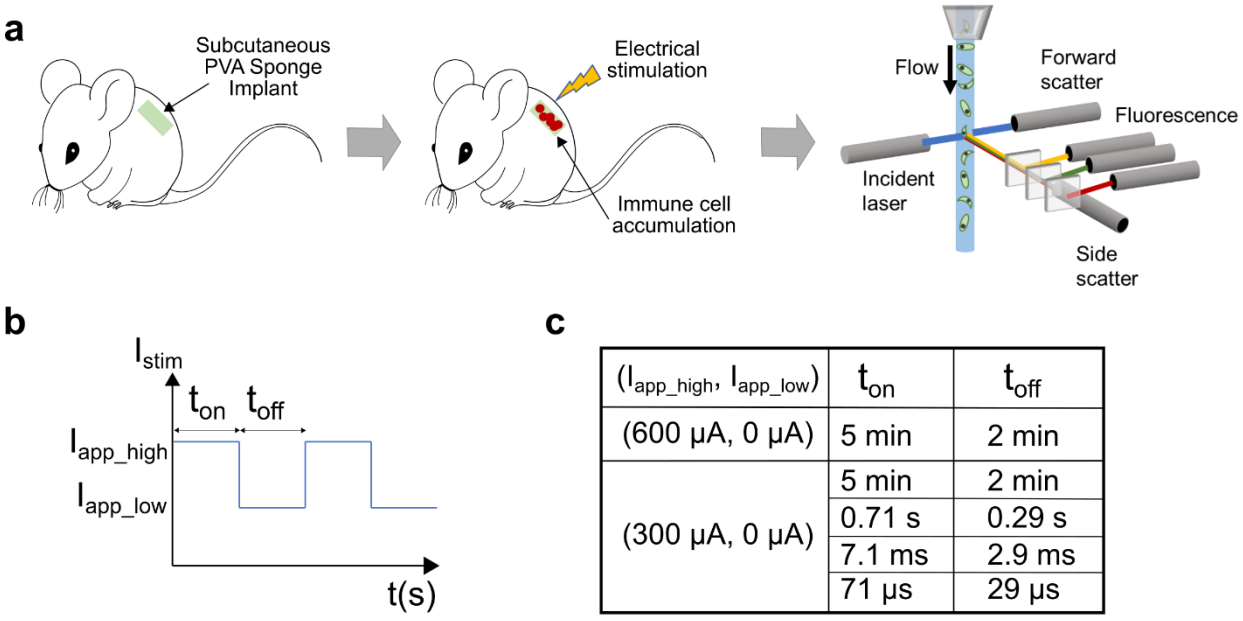
Flow cytometry data was analyzed using FlowJo software (FlowJo LLC, Ashland, OR) and statistical analysis was performed in R. Statistical significance was determined by two-sided Mann-Whitney U-test with $p \leq 0.05$. The mean fluorescent intensity of each data point was normalized to the average of the sham data points for each particular experiment. The figures follow the convention for plotting flow cytometry data, for which the numerical values are not included on the axes.

6.4 Results

6.4.1 Printed electrodes for electrical stimulation of cutaneous wounds

In addition to their performance for electrophysiological measurements, we also used these dry electrodes for in vivo electrical stimulation in the context of wound healing.

To study the effects of stimulation a wound model was created on the dorsal region of a 6-8 weeks old male C57BL/6 mouse (**Figure 6.1.a**). Following the subcutaneous implantation of PVA sponges into the mice, they were maintained in an unstimulated state, without electrodes attached, to enable the normal leukocyte infiltration for 3 days after sponge implantation as previously described.[335] Afterwards, mice were electrically stimulated with constant-current square-wave pulses for 1 hr at varying currents and frequencies (**Figure 6.1.b-c**). The stimulation current range was chosen based on previous studies applying electrical stimulation to wound areas in animal models and patients.[326,330] We chose to apply a monophasic current stimulation, hypothesizing that applying an exogenous electric field in the same direction as the wound field[324] would maximize the effect of stimulation, while the pulsed nature of the applied field would minimize heating effects and any electrochemical effects in the vicinity of the electrodes[353–355].



c

(I_{app_high}, I_{app_low})	t_{on}	t_{off}
(600 µA, 0 µA)	5 min	2 min
(300 µA, 0 µA)	5 min	2 min
	0.71 s	0.29 s
	7.1 ms	2.9 ms
	71 µs	29 µs

Figure 6.1: Experimental framework to study the effect of electrical stimulation during wound healing

(a) Schematic of the experimental workflow. Cells infiltrated in the sponge are harvested for analysis by flow cytometry. **(b)** Schematic of the stimulation waveform. **(c)** Table listing the stimulation parameters.

The ES electrodes used in this study were additively printed to integrate the wires for easy placement and precise control of stimulation over different wound regions. Two types of electrodes were fabricated and tested, to demonstrate applicability across different electrode geometries and substrate characteristics, as well as to demonstrate applicability with high density electrodes which can be used for spatial mapping. Specifically, we used a low density, stretchable array fabricated on a substrate with strong adhesion characteristics via facile screen printing (**Figure 6.2.a-b**), and a high density array fabricated via inkjet printing, suitable for impedance mapping (**Figure 6.2.c**).

The screen-printed array was fabricated on a stretchable substrate made of the silicone elastomer polydimethylsiloxane (PDMS), into which a small amount of ethoxylated polyethylenimine (PEIE) was added. This addition of PEIE has been shown to modify the PDMS cross-linking network to increase stretchability and Van der Waals adhesion by introducing microscale wrinkles on the elastomer surface,[282,340] thus serving as a good substrate for devices which need to adhere to skin. The electrodes consisted of a central electrode and four equally spaced outer electrodes (**Figure 6.2.a-b**). During stimulation the central electrode was placed on the wound site while the outer electrodes were on unwounded tissue. This configuration allowed an electric field to be directed from inside to outside the wound area or vice versa. The interconnects were patterned by blade-coating Ag/AgCl ink mixed with a stretchable elastomer (Ecoflex),[340] yielding stretchable traces with high conductivity of 2000 S*cm to enable low-impedance connections between the electrodes and controller hardware.

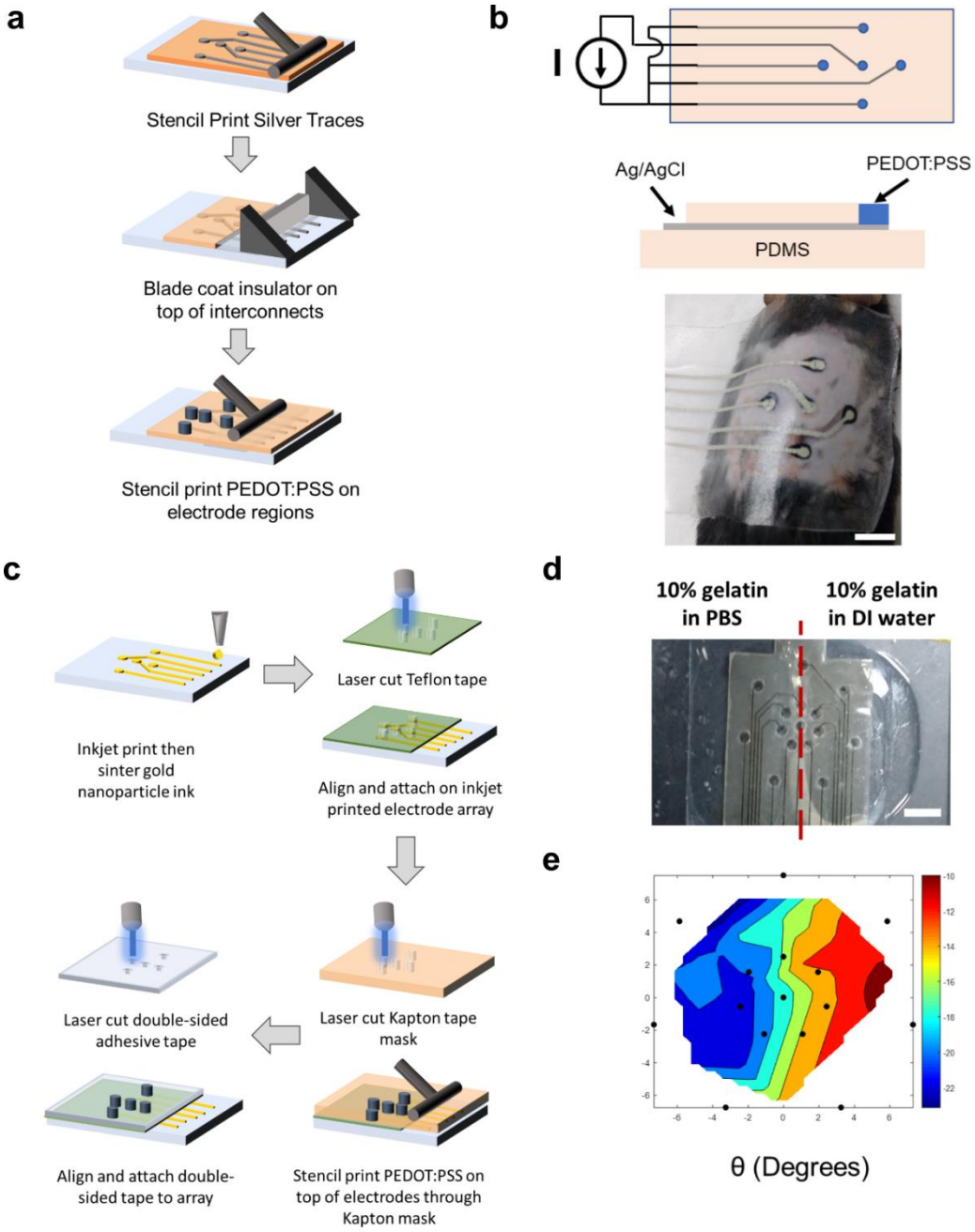
The inkjet-printed array in (**Figure 6.2.c**) was fabricated on a flexible polyethylene naphthalate substrate by printing gold interconnects using a gold nanoparticle ink.[341] The traces were insulated by attaching a Teflon tape on top, with vias in the tape having been cut by a laser-cutter. The finer patterning capability of the inkjet printing process allows for rapid prototyping and fabrication of high-density arrays. We demonstrated the capability of these arrays to map areas

of varying impedance, as found in wound areas which would have lower impedance in regions with damaged skin compared to undamaged skin. We used hydrogel phantoms to create areas of low and high impedance (**Figure 6.2.d**) and measured the impedance magnitude and phase angle between pairs of neighboring electrodes. Measurements of the impedance angle distinguish the regions of low and high conductivity (**Figure 6.2.e**) as discussed previously[323].

Both screen-printed and inkjet-printed electrodes were covered with a conducting polymer poly(3,4-ethylenedioxythiophene) polystyrene sulfonate (PEDOT:PSS) to provide a stable, biocompatible interface,[281,340,356–358] further lower the impedance at the electrode-skin interface, and provide a greater charge injection capacity for ES than metal surfaces.[342,343] As shown (**Figure 4.9.f**), the electrodes without PEDOT:PSS coated on the ends exhibited a reduction in current supplied over time, reducing from 300 μA to less than 10 μA current after 40 minutes, while the PEDOT:PSS coated electrodes maintained the current magnitude for the full 1 hr stimulation time. The electrode arrays with PEDOT:PSS were then used in our ES studies comparing the effects of applied current amplitudes and frequencies (**Figure 6.1.c**). The ES power supply was in constant-current mode, and voltage was monitored to limit the voltage to a maximum of 5 V.

Figure 6.2: Printed electrode design and testing

(a) Fabrication process of screen-printed electrodes **(b)** Schematic of the stimulation setup and photograph of electrodes mounted on top of the wound site, where the implanted sponge was beneath the center electrode. Scale bar: 9 mm. **(c)** Fabrication process for inkjet-printed electrodes. **(d)** Demonstration of impedance mapping using inkjet-printed electrodes on hydrogel phantoms. Scale bar: 5 mm **(e)** Spatial maps of impedance magnitude and angle between pairs of neighboring electrodes. Black dots indicate electrode location. (**f**) Current vs time during stimulation for electrodes with and without PEDOT:PSS coating. For stimulation tests on both sets of electrodes, a 300 μA square wave with duty cycle of 71.4% was applied, with voltage limited to a maximum of 5 V.


a
Stencil Print Silver Traces
Blade coat insulator on top of interconnects
Stencil print PEDOT:PSS on electrode regions
b
I
Ag/AgCl
PEDOT:PSS
PDMS
c
Inkjet print then sinter gold nanoparticle ink
Laser cut Teflon tape
Align and attach on inkjet printed electrode array
Laser cut Kapton tape mask
Stencil print PEDOT:PSS on top of electrodes through Kapton mask
Align and attach double-sided tape to array
Laser cut double-sided adhesive tape
d
10% gelatin in PBS
10% gelatin in DI water
e
θ (Degrees)
f
With PEDOT:PSS Coating
Without PEDOT:PSS Coating
Current (μA)
Time (s)

6.4.2 Assaying the effects of stimulation on wound-infiltrating leukocytes

To determine whether ES would have any deleterious effect on cell viability, we assessed the fraction of live cells at the applied current magnitudes of 300 μA and 600 μA. The applied square waveform frequency was at 2.3 mHz (a period of 7 min, with 5 min on to 2 min off for each duty cycle as detailed in Fig. 1C), and only current magnitudes were varied. After stimulation, the PVA sponges were extracted from the wound site, and cells were harvested from the sponge. A subset of the harvested cells were treated with the stain propidium iodide, which selectively enters and labels only dead cells due to their greater membrane permeability. The treated sample was analyzed using flow cytometry. The single-cell population was obtained by gating the single-cell region of forward- and side-scatter signals (**Figure 6.3.a**). In this single-cell subset, viable cells were ones not permeated by the propidium iodide dye and were counted within the indicated region of interest. The percentage of viable cells were calculated (**Figure 6.3.b**), with no significant differences in viability observed between ES applied at 300 μA, 600 μA, and sham.

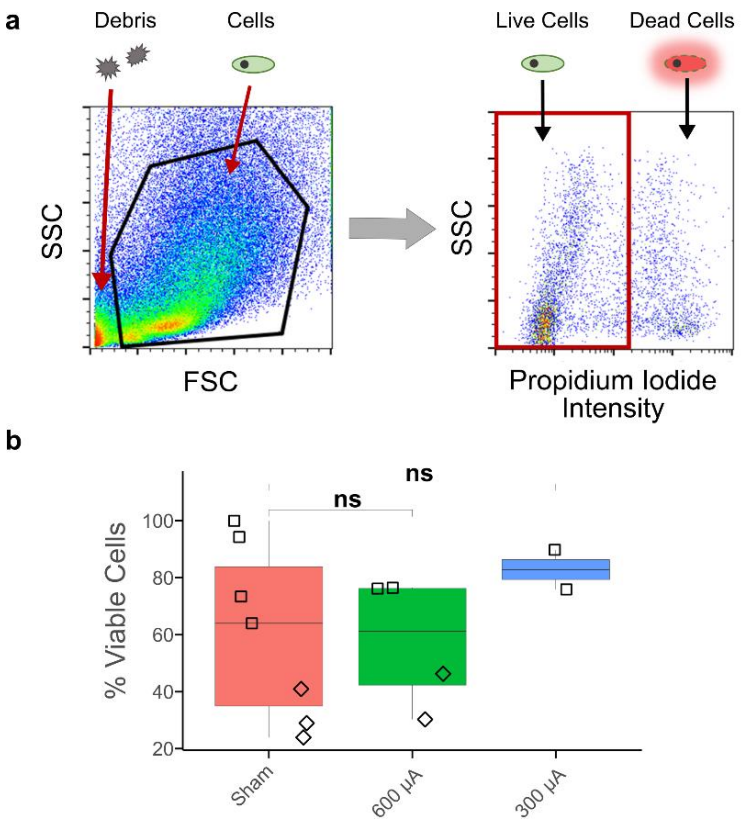


Figure 6.3: Effect of stimulation current magnitudes on cell viability

(a) Flow cytometry analysis workflow. Gating region for scattered light signals from single cells (left), and for fluorescence signal from dead cells by propidium iodide viability stain (right). SSC is the signal from side-scattered light and FSC is the signal from forward-scattered light. **(b)** Comparison of viability for different applied current magnitudes vs sham. The stimulation was at 2.3 mHz for 1 hour, with □ representing one-day experiments, and ◊representing three consecutive days of applying stimulation or sham condition. "ns" indicates not significant p-value > 0.05.

To demonstrate the applicability of the platform to study cell types that are characteristic of the inflammatory phase of wound healing, and more physiologically relevant compared to *in vitro* systems, we chose to focus on $CD45^+$ infiltrating leukocytes cells. As the viability was determined to be unaffected by ES, a subset of the harvested cells were fixed, permeabilized and labelled with fluorescently-tagged antibodies against the pan-leukocyte surface marker CD45, phosphorylated Erk or phosphorylated Akt to assess which pathways were activated in infiltrating immune cells upon application of ES (**Figure 6.4.a**). The phosphorylation of Erk was determined based on the role of the Mapk/Erk signaling cascade regulating the response to extracellular stimuli such as growth factors and extracellular matrix that mediate cellular responses including

proliferation, activation, differentiation and survival[348]. Phosphorylation of Akt was determined based on its role in the PI3k/Akt signaling network in regulating cell survival, angiogenesis, chemotaxis, activation and maturation[349].

Using the same waveform as the above viability evaluation, we examined the effect of varying current magnitudes on phosphorylation of Erk and Akt using site specific antibodies to each target in permeabilized $CD45^+$ cells. Following flow cytometry, single cells were gated, and within this subset, the $CD45^+$ population gated for analysis in. Here the stain CD45 was used to identify inflammatory leukocytes, since it is a pan-leukocyte marker,[359] and thus capturing the Erk or Akt signal across multiple $CD45^+$ cell types. $CD45^+$ cells were assessed for phosphorylated Erk (P-Erk) or Akt (P-Akt) fluorescence intensity, as shown in exemplar (**Figure 6.4.a**). To confirm that ES effects were not driven by factors specific to the electrode geometry or materials, we conducted ES experiments with both, the screen-printed and inkjet-printed electrodes. With the high-density inkjet-printed array, the electric field between the wound site and the intact tissue were adjusted to be at a similar magnitude as the experiments with low-density electrodes.

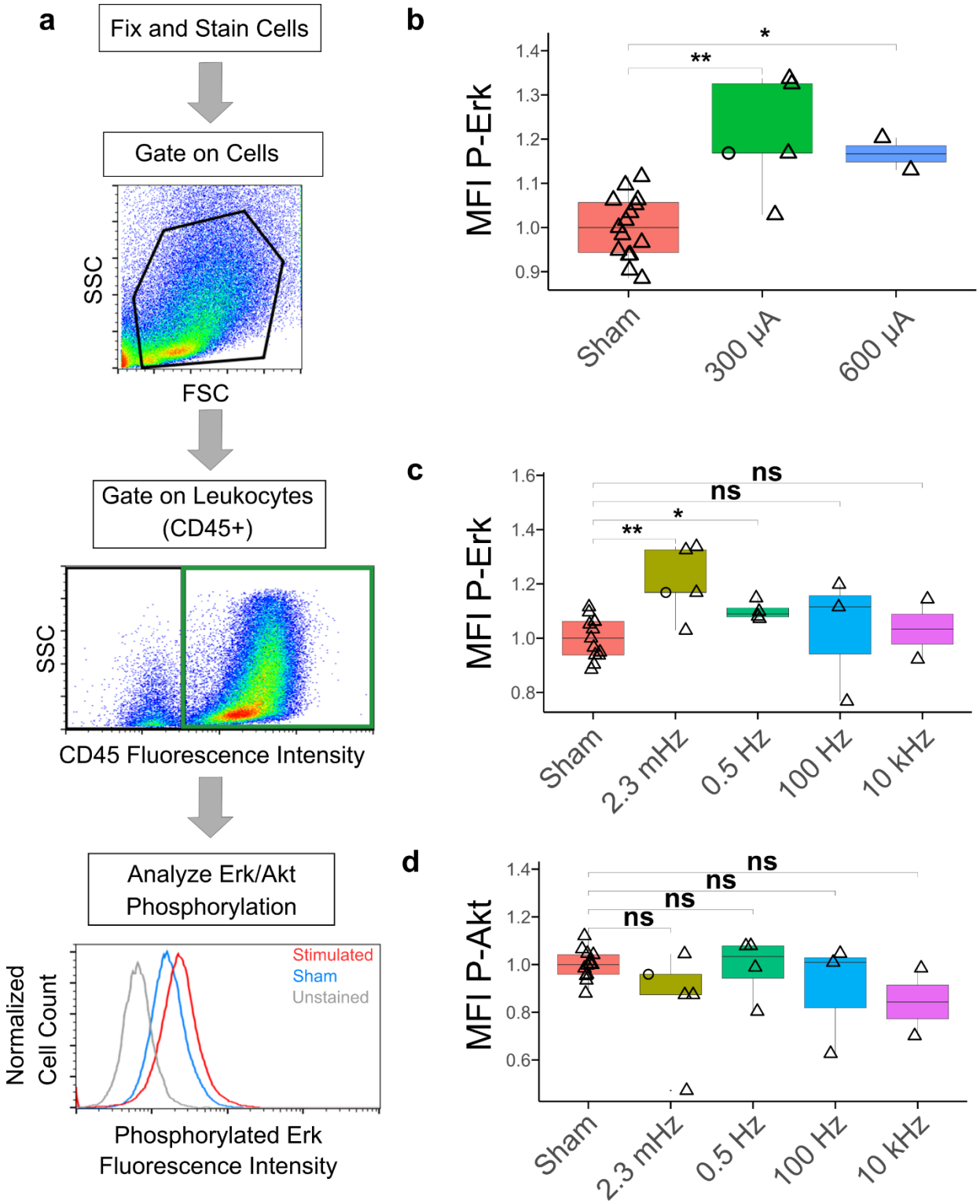
We observed that ES at both 300 μA and 600 μA increased P-Erk compared to the sham controls (**Figure 6.4.b**). Thus we used the 300 μA current to be safely within the viability and effective ES range in the subsequent experiments. As such, we also found that the increase in P-Erk level was comparable regardless of electrode density.

Next, we tested whether the frequency of stimulation could modulate the response of the Erk pathway, since the frequency dependence of signaling pathway activation has been observed previously *in vitro.*[360,361] The frequency of the square wave pulse was varied over 6 orders of magnitude to discern the frequency dependence of Erk or Akt phosphorylation. P-Erk levels were observed to be increased over sham at low ES frequencies, 2.3 mHz and 0.5 Hz, while no significant effect was observed for ES at the higher frequencies tested (**Figure 6.4.c**). In contrast,

P-Akt levels were unaffected by the application of ES across all the frequency range tested (**Figure 6.4.d**), thus indicating that the Erk pathway is among the earliest signaling pathways involved in the transduction of electrical signals in immune cells. The increase in P-Erk levels at 2.3 mHz low frequency stimulation was consistent across both electrode densities tested. The results show that Erk pathway activation is preferred over Akt after electrical stimulation and demonstrates feasibility to study intracellular changes based on our platform.

Figure 6.4: Electrical stimulation-driven activation in immune cells

(a) Flow cytometry gating and signal measurement workflow. Gating regions for scattered light signal from cells, CD45+ stained cells, and representative histograms of phosphorylated Erk signals. **(b)** Mean fluorescence intensity (MFI) of Erk phosphorylation in CD45+ cells, for different applied current ≤ 0.5 Hz. **(c)** MFI of Erk phosphorylation and **(d)** MFI of Akt phosphorylation in CD45+ cells. **(c)** and **(d)** are carried out under an applied current of 300 uA amplitude. Data points represented by ○ are for experiments conducted with inkjet printed electrodes; data points represented by Δ are for experiments conducted with screen-printed electrodes. * indicating p-value ≤ 0.05, and ** indicating: p-value ≤ 0.01. “ns” indicates not significant p-value > 0.05.


a
Fix and Stain Cells
Gate on Cells
SSC
FSC
Gate on Leukocytes (CD45+)
SSC
CD45 Fluorescence Intensity
Analyze Erk/Akt Phosphorylation
Normalized Cell Count
Stimulated
Sham
Unstained
Phosphorylated Erk Fluorescence Intensity
b
MFI P-Erk
1.4
1.3
1.2
1.1
1.0
0.9
*
**
Sham
300 µA
600 µA
c
MFI P-Erk
1.6
1.4
1.2
1.0
0.8
ns
ns
*
**
Sham
2.3 mHz
0.5 Hz
100 Hz
10 kHz
d
MFI P-Akt
1.4
1.2
1.0
0.8
0.6
ns
ns
ns
ns
Sham
2.3 mHz
0.5 Hz
100 Hz
10 kHz

6.5 Discussion

The PVA sponge implant model to define the inflammation component of wound healing, combined with stretchable electrode arrays, offers a versatile platform to examine the effects of ES parameters on molecular mechanisms that occur during the wound healing process. This approach provides a molecular readouts relevant in leukocytes that can be readily expandable to other pathways using the range of available and well-characterized phosphorylation state-specific antibodies. We observed that the Erk pathway is activated in inflammatory leukocytes when low frequency electrical stimulation ≤0.5 Hz is applied at the wound site. In contrast, the Akt pathway was not affected by ES application. Future studies may be directed towards identifying the different signaling pathways in sub-populations of cells that are activated during the healing process, thus helping to build a comprehensive map of cell-type specific signal transduction networks driving wound healing. Such a map could then be harnessed to develop advanced solutions to accelerate wound healing.

6.6 Acknowledgements

This work was supported by NextFlex Award #042299 and UC San Diego Medical Devices and Systems Initiative funded by the Electrical and Computer Engineering Department. Part of the work was performed at the San Diego Nanotechnology Infrastructure of UCSD, which is supported by NSF ECCS-1542148.

Chapter 6 in part is a reprint of the following materials: Wang, K.*, Parekh, U.*, Ting, J. K., Yamamoto, N. A. D., Zhu, J., Costantini, T., Arias, A. C., Eliceiri, B. P. & Ng, T. N. A Platform to Study the Effects of Electrical Stimulation on Immune Cell Activation During Wound Healing, *Advanced Biosystems* 3 (10), 1900106 (2019) (* equal contribution). The dissertation author is one of the two primary authors of the study.

Chapter 6: Conclusions and outlook

Since the discovery of pluripotent stem cells and especially of iPSC technology, they have been employed to make incredible progress in the modeling of development and disease, drug discovery and for regenerative medicine. Still, considerable progress remains to be made on utilizing PSCs to assemble functional, complex human tissue.

6.1 Stem cell differentiation

While there exists a dizzying array of methods to differentiate various cells present in the human body from PSCs, many of these – especially those which utilize chemically directed differentiation – suffer from the drawbacks of complexity, length and low yield. On the other hand, those employing reprogramming methods are promising but tedious to discover.

In chapter 2, we explored a solution to this conundrum by developing a high-throughput screen with an information-rich transcriptomic and fitness readout using gene overexpression combined with single cell RNA-sequencing to discover reprogramming recipes and study the effects of overexpression. This approach enabled the rapid construction of gene regulatory networks and the elucidation of the effects of overexpression, such as the induction of fates toward the epithelial and mesenchymal direction by *KLF4* and *SNAI2* respectively. As well as the direct discovery of reprogramming factors, such as *ETV2* which we found differentiates PSCs to an endothelial-like fate. This method also has the potential to explore combinatorial effects of overexpression since it enables the precise mapping of perturbagens to each cell.

6.2 Modeling disease states

The transition of cells from healthy to diseased states is influenced not only by cell intrinsic, but also extrinsic factors such as the microenvironment. While animal models provide an avenue to model cells in their native environment they do not capture human-specific biology, thus enabling the study of human cells in settings similar to their in vivo niche is a critical need.

Specifically, to study neoplastic transformation it is necessary to follow cell state changes over time, as healthy cells gain a transformed phenotype. Such studies require access to healthy, wild type cells, a means to perturb them with oncogenic drivers, and a means to assay their transformation potential and proliferative advantage.

In chapter 3, we developed a highly multiplexed screening method which leveraged the diverse cell types generated in spontaneously differentiated teratomas in combination with the scRNA-seq based screening method developed in chapter 2, and serial re-injection of proliferating tumors to study the tissue-specific effectors of oncogenic transformation across genes, variants and lineages. This platform allows for the spatiotemporal profiling of human tumor development in an in vivo setting, and allows for the longitudinal study of neoplastic transformation and cancer progression. We demonstrated the potential of the platform by profiling over 20 lineages and over 50 drivers in a single multiplex assay, finding that MYC alone or in combination with myristoylated AKT1 drove the transformation of neural lineages, while $MEK1^{S218D/S222D}$ drove proliferative advantage in mesenchymal, fibroblast-like lineages.

6.3 Vascularization of Engineered Tissue

While *in vivo* implantation is a viable and promising avenue for disease modeling, it generates chimeric tissue which may not be applicable for regenerative medicine and may miss some aspects of human-specific biology since elements like the vasculature are often generated from the host system. On the other hand, organoids have made tremendous progress as stem cell-derived models of organotypic tissue but they suffer from a few challenges, among them: the inability to derive cross-germ layer cell types and lack of vascularization which prevents large scale growth.

In chapter 4, we developed a method to introduce parenchymal cell types into blood vessel organoids by combining gene overexpression with chemically directed organoid differentiation.

Overexpression of the lineage-specifying transcription factors *NEUROD1* and *MYOD1* overrode chemical cues to drive neuronal and skeletal muscle identities, while the directed differentiation generated functional, complete blood vessel networks integrated with the parenchymal cells. This method may serve to create organoid building blocks to derive large scale tissues for eventual regenerative medicine applications.

6.4 Flexible Electronics for Electrophysiology and Stimulation

To fully enable the use of engineered tissue, major efforts are being directed toward devices for monitoring and perturbation of tissue *in vitro* and *in vivo*. Traditional electronic devices, composed of inorganic conducting and semiconducting materials and fabricated on rigid substrates, are sub-optimal for monitoring signals from the soft, deformable and curved surfaces of the body. Flexible electronic devices fabricated from both inorganic and organic materials offer a growing body of viable alternative devices. Flexible devices offer superior performance through conformal contact, deformability and enable aspects like long term wearable monitoring and stimulation devices.

In chapter 5, we enable the measurement of electrophysiological signals with high spatial resolution through the development of stretchable concentric electrodes fabricated as a hybrid of inorganic and organic materials. Silver was used to deposit traces for interfacing with external input/output circuitry, while PEDOT:PSS was used as an interface layer between the silver and skin, resulting in low noise measurements of clinically relevant electrical signals without the use of conductive gels. We used a highly stretchable and adhesive substrate made of PDMS modified by PEIE to enable wearability and conformal contact. In chapter 6, we also applied these and inkjet-printed electrodes toward electrical stimulation during wound healing, creating a platform technology which allowed the study of of the immune response during wound healing.

Finally, in a preliminary demonstration, we combined inkjet and electrohydrodynamic printing to fabricate flexible micropillar electrodes to potentially enable a facile method to create flexible microelectrode arrays for the measurement of local field potential in vivo, from organoids and other engineered tissue.

6.5 Future scope

Taken together, we have demonstrated a multi-faceted toolkit for the engineering and monitoring of human, organotypic tissue in health and disease. Looking ahead there are several aspects of these systems which can be optimized, enhanced or reimagined.

There are estimated to be approximately 1600 TFs in the human genome, and while searching for reprogramming recipes involving single TFs has been demonstrated across all human TFs[362], most recipes consist of combinations of TFs and screening combinations experimentally quickly becomes unfeasible as the number of combinations is increased. Thus, computational methods[363–365] to rationally design screening libraries is a must. Rationally chosen libraries can be combined with reporter assays and single cell screens to enable rapid discovery.

Further, while the introduction of other cell types into organoids has been demonstrated, there is no spatial control over how these cells are induced or arranged, and they remain in immature phenotypes not reaching the maturity of adult tissue. Several methods could be adapted to drive spatial organization, including synthetic biology techniques which can induce specific gene expression upon cell-cell contact[264] and programmable morphogen gradients[366], as well as optogenetic methods[262,263] which can control gene expression and differentiation using light-inducible dimerization. Beyond this, organoids can be combined with methods like bioprinting[65] to enable the creation of tissues of translationally relevant sizes.

Finally, much progress remains to be made in the development of devices for long term monitoring and stimulation of tissue. Intrinsically stretchable and geometrically designed stretchable electronics can both be utilized for these aims. The development of electrode-embedded organoids[289] and embedding high resolution electrode arrays *in vivo*[367,368], to spatiotemporally map electrical and chemical signals, as well as to precisely perturb tissues via stimulation and controlled chemical release will help unravel biology and enable the creation of novel therapies.

Appendix

DNA sequences from Chapter 2

EF1α + Kozak Sequence + mCherry + P2A + Hygromycin Resistance + Barcode

cgtgaggctccggtgcccgtcagtgggcagagcgcacatcgcccacagtccccgagaagttgggggggagggtcggc

aattgaaccggtgcctagagaaggtggcgcggggtaaactgggaaagtgatgtcgtgtactggctccgccttttcc

cgagggtgggggagaaccgtatataagtgcagtagtcgccgtgaacgttcttttcgcaacgggtttgccgccagaa

cacaggtaagtgccgtgtgtggttcccgcgggcctggcctctttacgggttatggcccttgcgtgccttgaattact

tccacctggctgcagtacgtgattcttgatcccgagcttcgggttggaagtgggtgggagagttcgaggccttgcgc

ttaaggagccccttcgcctcgtgcttgagttgaggcctggcctgggcgctggggccgccgcgtgcgaatctggtggc

accttcgcgcctgtctcgctgctttcgataagtctctagccatttaaaatttttgatgacctgctgcgacgcttttt

ttctggcaagatagtcttgtaaatgcgggccaagatctgcacactggtatttcggttttggggccgcgggcggcga

cggggcccgtgcgtcccagcgcacatgttcggcgaggcggggcctgcgagcgcggccaccgagaatcggacgggggt

agtctcaagctggccggcctgctctggtgcctggcctcgcgccgccgtgtatcgccccgccctgggcggcaaggctg

gcccggtcggcaccagttgcgtgagcggaaagatggccgcttcccggccctgctgcagggagctcaaaatggaggac

gcggcgctcgggagagcgggcgggtgagtcacccacacaaaggaaaagggcctttccgtcctcagccgtcgcttcat

gtgactccacggagtaccgggcgccgtccaggcacctcgattagttctcgagcttttggagtacgtcgtctttaggt

tgggggggaggggttttatgcgatggagtttccccacactgagtgggtggagactgaagttaggccagcttggcactt

gatgtaattctccttggaatttgccctttttgagtttggatcttggttcattctcaagcctcagacagtggttcaaa

gtttttttcttccatttcaggtgtcgtgaggaattagcttggtactaatacgactcactatagggagacccaagctg

gctagttaagcttgatatcgaattcctgcagcccgggggatccaccggtcgccaccatggtgagcaagggcgaggag

gataacatggccatcatcaaggagttcatgcgcttcaaggtgcacatggagggctccgtgaacggccacgagttcga

gatcgagggcgagggcgagggccgcccctacgagggcacccagaccgccaagctgaaggtgaccaagggtggccccc

tgcccttcgcctgggacatcctgtcccctcagttcatgtacggctccaaggcctacgtgaagcaccccgccgacatc

cccgactacttgaagctgtccttccccgagggcttcaagtgggagcgcgtgatgaacttcgaggacggcggcgtggt

gaccgtgacccaggactcctccctgcaggacggcgagttcatctacaaggtgaagctgcgcggcaccaacttcccct

ccgacggccccgtaatgcagaagaagaccatgggctgggaggcctcctccgagcggatgtaccccgaggacggcgcc

ctgaagggcgagatcaagcagaggctgaagctgaaggacggcggccactacgacgctgaggtcaagaccacctacaa

ggccaagaagcccgtgcagctgcccggcgcctacaacgtcaacatcaagttggacatcacctcccacaacgaggact

acaccatcgtggaacagtacgaacgcgccgagggccgccactccaccggcggcatggacgagctgtacaagGGATCC

GGCGCAACAAACTTCTCTCTGCTGAAACAAGCCGGAGATGTCGAAGAGAATCCTGGACCGATGAAAAAGCCTGAACT

CACCGCGACGTCTGTCGAGAAGTTTCTGATCGAAAAGTTCGACAGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCG

AAGAATCTCGTGCTTTCAGCTTCGATGTAGGAGGGCGTGGATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTTC

TACAAAGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCGCTCCCGATTCCGGAAGTGCTTGACATTGGGGAATT

CAGCGAGAGCCTGACCTATTGCATCTCCCGCCGTGCACAGGGTGTCACGTTGCAAGACCTGCCTGAAACCGAACTGC

CCGCTGTTCTGCAGCCGGTCGCGGAGGCCATGGATGCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGGC

CCATTCGGACCGCAAGGAATCGGTCAATACACTACATGGCGTGATTTCATATGCGCGATTGCTGATCCCCATGTGTA

TCACTGGCAAACTGTGATGGACGACACCGTCAGTGCGTCCGTCGCGCAGGCTCTCGATGAGCTGATGCTTTGGGCCG

AGGACTGCCCCGAAGTCCGGCACCTCGTGCACGCGGATTTCGGCTCCAACAATGTCCTGACGGACAATGGCCGCATA

ACAGCGGTCATTGACTGGAGCGAGGCGATGTTCGGGGATTCCCAATACGAGGTCGCCAACATCTTCTTCTGGAGGCC

GTGGTTGGCTTGTATGGAGCAGCAGACGCGCTACTTCGAGCGGAGGCATCCGGAGCTTGCAGGATCGCCGCGGCTCC

GGGCGTATATGCTCCGCATTGGTCTTGACCAACTCTATCAGAGCTTGGTTGACGGCAATTTCGATGATGCAGCTTGG

GCGCAGGGTCGATGCGACGCAATCGTCCGATCCGGAGCCGGGACTGTCGGGCGTACACAAATCGCCCGCAGAAGCGC

GGCCGTCTGGACCGATGGCTGTGTAGAAGTACTCGCCGATAGTGGAAACCGACGCCCCAGCACTCGTCCGAGGGCAA

AGGAATAAGTTAGAACTATTTCCTGGCTGTTACGCGNNNNNNNNNNNNNNNNNNNNCTACTGACGGGCTTGTTAAAC

AGTG

ASCL1

ATGGAGTCTTCTGCTAAAATGGAGTCCGGAGGCGCGGGACAACAACCACAACCGCAACCACAACAACCCTTCCTGCC

GCCGGCCGCATGTTTTTTCGCGACCGCTGCTGCTGCTGCAGCGGCGGCGGCTGCTGCCGCCGCGCAATCCGCCCAAC

AGCAACAACAACAACAGCAGCAGCAGCAACAAGCGCCTCAACTTCGACCCGCTGCAGACGGGCAGCCCTCAGGGGGA

GGGCACAAGAGCGCTCCGAAGCAGGTTAAAAGGCAGAGGAGCAGTAGTCCCGAACTGATGCGATGTAAGAGGCGCCT

CAATTTTAGCGGTTTTGGTTACTCTTTGCCCCAGCAGCAGCCGGCTGCCGTAGCTCGCCGAAATGAGCGGGAAAGGA

ACCGCGTTAAACTTGTGAATCTCGGTTTCGCGACACTTCGAGAGCACGTACCAAATGGGGCAGCTAACAAGAAAATG

AGTAAAGTTGAGACACTGCGGTCTGCAGTGGAGTATATTAGAGCTCTTCAACAATTGCTTGACGAGCACGATGCCGT

ATCAGCCGCATTTCAAGCCGGGGTGCTGTCCCCAACAATATCTCCGAACTACAGCAATGATCTTAATAGCATGGCGG

GAAGTCCCGTTTCCTCCTACTCCTCTGATGAGGGCAGCTACGACCCTCTCAGTCCCGAGGAGCAAGAGCTTCTTGAC
TTCACTAACTGGTTC

ASCL3

ATGATGGACAACAGAGGCAACTCTAGTCTACCTGACAAACTTCCTATCTTCCCTGATTCTGCCCGCTTGCCACTTAC
CAGGTCCTTCTATCTGGAGCCCATGGTCACTTTCCACGTGCACCCAGAGGCCCCGGTGTCATCTCCTTACTCTGAGG
AGCTGCCACGGCTGCCTTTTCCCAGCGACTCTCTTATCCTGGGAAATTACAGTGAACCCTGCCCCTTCTCTTTCCCG
ATGCCTTATCCAAATTACAGAGGGTGCGAGTACTCCTACGGGCCAGCCTTCACCCGGAAAAGGAATGAGCGGGAAAG
GCAGCGGGTGAAATGTGTCAATGAAGGCTACGCCCAGCTCCGACATCATCTGCCAGAGGAGTATTTGGAGAAGCGAC
TCAGCAAAGTGGAAACCCTCAGAGCTGCGATCAAGTACATTAACTACCTGCAGTCTCTTCTGTACCCTGATAAAGCT
GAGACAAAGAATAACCCTGGAAAAGTTTCCTCCATGATAGCAACCACCAGCCACCATGCTGACCCTATGTTCAGAAT
TGTTTGCCCAACTTTCTTGTACAAAGTTGTCCCC

ASCL4

ATGGAGACGCGTAAACCGGCGGAACGGCTGGCCTTGCCATACTCGCTGCGCACCGCGCCCCTGGGCGTTCCGGGGAC
CCTGCCCGGACTCCCGCGGAGGGACCCCCTCAGGGTCGCCCTGCGTCTGGACGCCGCGTGCTGGGAGTGGGCGCGCA
GCGGCTGCGCACGGGGATGGCAGTACTTGCCCGTGCCGCTGGACAGCGCCTTCGAGCCCGCCTTCCTCCGCAAGCGC
AACGAGCGCGAGCGGCAGCGGGTGCGCTGCGTGAACGAGGGCTATGCGCGCCTCCGAGACCACCTGCCCCGGGAGCT
GGCAGACAAGCGCCTCAGCAAAGTGGAGACGCTCCGCGCTGCCATCGACTACATCAAGCACCTGCAGGAGCTGCTGG
AGCGCCAGGCCTGGGGGCTCGAGGGCGCGGCCGGCGCCGTCCCCCAGCGCAGGGCGGAATGCAACAGCGACGGGGAG
TCCAAGGCCTCTTCGGCGCCTTCGCCCAGCAGCGAGCCCGAGGAGGGGGGCAGC

ASCL5

ATGCCGATGGGGGCAGCAGAAAGAGGTGCTGGGCCCCAATCATCTGCAGCACCATGGGCTGGTTCAGAAAAGGCGGC
AAAGAGAGGGCCATCAAAAAGCTGGTACCCAAGAGCTGCTGCATCTGATGTCACGTGCCCGACTGGTGGTGATGGAG
CTGACCCAAAACCTGGACCTTTTGGAGGTGGTTTAGCTTTAGGGCCTGCGCCCAGAGGAACAATGAATAATAATTTC
TGCAGGGCCCTTGTTGACAGAAGGCCTTTAGGACCCCCTTCATGTATGCAATTAGGTGTAATGCCACCGCCAAGACA
AGCGCCCCTCCCGCCGGCTGAACCCCTTGGAAATGTACCTTTCCTCCTATACCCTGGCCCAGCTGAACCACCATATT

ATGATGCATATGCTGGTGTTTTCCCATATGTGCCTTTCCCTGGTGCTTTTGGTGTATATGAATACCCTTTTGAGCCG

GCTTTTATCCAAAAGAGGAATGAAAGAGAGAGACAGAGAGTGAAGTGTGTGAATGAAGGATACGCCAGATTGAGAGG

CCATTTGCCTGGTGCCCTGGCAGAAAAGAGATTATCAAAAGTTGAAACCCTGAGGGCGGCAATCAGATATATAAAAT

ACCTCCAAGAACTCCTTTCATCAGCACCTGATGGATCGACACCACCGGCTTCAAGAGGTTTACCTGGAACTGGACCA

TGCCCTGCACCGCCTGCTACACCAAGGCCAGACAGACCTGGAGATGGAGAAGCAAGAGCACCTTCTTCCCTTGTCCC

TGAATCTTCTGAATCATCATGTTTTTCGCCTTCCCCTTTTTTAGAAAGTGAAGAATCCTGGCA

ATF7

ATGGGAGACGACAGACCGTTTGTGTGCAATGCCCCGGGCTGTGGACAGAGATTTACAAACGAGGACCACCTGGCAGT

TCATAAACACAAGCATGAGATGACATTGAAATTTGGCCCAGCCCGAACTGACTCAGTCATCATTGCAGATCAAACGC

CTACTCCAACTAGATTCCTGAAGAACTGTGAGGAGGTGGGACTCTTCAATGAACTAGCTAGCTCCTTTGAACATGAA

TTCAAGAAAGCTGCAGATGAGGATGAGAAAAAGGCAAGAAGCAGGACTGTTGCCAAAAAACTGGTGGCTGCTGCTGG

GCCCCTTGACATGTCTCTGCCTTCCACACCAGACATCAAAATCAAAGAAGAAGAGCCAGTGGAGGTAGACTCATCCC

CACCTGATAGCCCTGCCTCTAGTCCCTGTTCCCCACCACTGAAGGAGAAGGAGGTTACCCCAAAGCCTGTTCTGATC

TCTACCCCCACACCCACCATTGTACGTCCTGGCTCCCTGCCTCTCCACTTGGGCTATGATCCACTTCATCCAACCCT

TCCCTCCCCAACCTCTGTCATCACACAGGCTCCACCATCCAACAGGCAAATGGGGTCTCCCACTGGCTCCCTCCCTC

TTGTCATGCATCTTGCTAATGGACAGACCATGCCTGTGTTGCCAGGGCCTCCAGTACAGATGCCGTCTGTTATATCG

CTGGCCAGACCTGTGTCCATGGTGCCCAACATTCCTGGTATCCCTGGCCCACCAGTTAACAGTAGTGGCTCCATTTC

TCCCTCTGGCCACCCTATACCATCAGAAGCCAAGATGAGACTGAAAGCCACCCTAACTCACCAAGTCTCCTCAATCA

ATGGTGGTTGTGGAATGGTGGTGGGTACTGCCAGCACCATGGTGACAGCCCGCCCAGAGCAGAGCCAGATTCTCATC

CAGCACCCTGATGCCCCATCCCCTGCCCAGCCACAGGTCTCACCAGCTCAGCCCACCCCTAGTACTGGGGGGCGACG

GCGGCGCACAGTAGATGAAGATCCAGATGAGCGACGGCAGCGCTTTCTGGAGCGCAACCGGGCTGCAGCCTCCCGCT

GCCGCCAAAAGCGAAAGCTGTGGGTGTCCTCCCTAGAGAAGAAGGCCGAAGAACTCACTTCTCAGAACATTCAGCTG

AGTAATGAAGTCACATTACTACGCAATGAGGTGGCCCAGTTGAAACAGCTACTGTTAGCTCATAAAGACTGCCCAGT

CACTGCACTACAGAAAAAGACTCAAGGCTATTTAGAAAGCCCCAAGGAAAGCTCAGAGCCAACGGGTTCTCCAGCCC

CTGTGATTCAGCACAGCTCAGCAACAGCCCCTAGCAATGGCCTCAGTGTTCGCTCTGCAGCTGAAGCTGTGGCCACC

TCGGTCCTCACTCAGATGGCCAGCCAAAGGACAGAACTGAGCATGCCGATACAATCGCATGTAATCATGACCCCACA

GTCCCAGTCTGCGGGCAGA

CDX2

ATGTACGTGAGCTACCTCCTGGACAAGGACGTGAGCATGTACCCTAGCTCCGTGCGCCACTCTGGCGGCCTCAACCT
GGCGCCGCAGAACTTCGTCAGCCCCCCGCAGTACCCGGACTACGGCGGTTACCACGTGGCGGCCGCAGCTGCAGCGG
CAGCGAACTTGGACAGCGCGCAGTCCCCGGGGCCATCCTGGCCGGCAGCGTATGGCGCCCCACTCCGGGAGGACTGG
AATGGCTACGCGCCCGGAGGCGCCGCGGCCGCCGCCAACGCCGTGGCTCACGGCCTCAACGGTGGCTCCCCGGCCGC
AGCCATGGGCTACAGCAGCCCCGCAGACTACCATCCGCACCACCACCCGCATCACCACCCGCACCACCCGGCCGCCG
CGCCTTCCTGCGCTTCTGGGCTGCTGCAAACGCTCAACCCCGGCCCTCCTGGGCCCGCCGCCACCGCTGCCGCCGAG
CAGCTGTCTCCCGGCGGCCAGCGGCGGAACCTGTGCGAGTGGATGCGGAAGCCGGCGCAGCAGTCCCTCGGCAGCCA
AGTGAAAACCAGGACGAAAGACAAATATCGAGTGGTGTACACGGACCACCAGCGGCTGGAGCTGGAGAAGGAGTTTC
ACTACAGTCGCTACATCACCATCCGGAGGAAAGCCGAGCTAGCCGCCACGCTGGGGCTCTCTGAGAGGCAGGTTAAA
ATCTGGTTTCAGAACCGCAGAGCAAAGGAGAGGAAAATCAACAAGAAGAAGTTGCAGCAGCAACAGCAGCAGCAGCC
ACCACAGCCGCCTCCGCCGCCACCACAGCCTCCCCAGCCTCAGCCAGGTCCTCTGAGAAGTGTCCCAGAGCCCTTGA
GTCCGGTGTCTTCCCTGCAAGCCTCAGTGTCTGGCTCTGTCCCTGGGGTTCTGGGGCCAACTGGGGGGGTGCTAAAC
CCCACCGTCACCCAG

CRX

ATGATGGCGTATATGAACCCGGGGCCCCACTATTCTGTCAACGCCTTGGCCCTAAGTGGCCCCAGTGTGGATCTGAT
GCACCAGGCTGTGCCCTACCCAAGCGCCCCCAGGAAGCAGCGGCGGGAGCGCACCACCTTCACCCGGAGCCAACTGG
AGGAGCTGGAGGCACTGTTTGCCAAGACCCAGTACCCAGACGTCTATGCCCGTGAGGAGGTGGCTCTGAAGATCAAT
CTGCCTGAGTCCAGGGTTCAGGTTTGGTTCAAGAACCGGAGGGCTAAATGCAGGCAGCAGCGACAGCAGCAGAAACA
GCAGCAGCAGCCCCCAGGGGGCCAGGCCAAGGCCCGGCCTGCCAAGAGGAAGGCGGGCACGTCCCCAAGACCCTCCA
CAGATGTGTGTCCAGACCCTCTGGGCATCTCAGATTCCTACAGTCCCCCTCTGCCCGGCCCCTCAGGCTCCCCAACC
ACGGCAGTGGCCACTGTGTCCATCTGGAGCCCAGCCTCAGAGTCCCCTTTGCCTGAGGCGCAGCGGGCTGGGCTGGT
GGCCTCAGGGCCGTCTCTGACCTCCGCCCCCTATGCCATGACCTACGCCCCGGCCTCCGCTTTCTGCTCTTCCCCCT
CCGCCTATGGGTCTCCGAGCTCCTATTTCAGCGGCCTAGACCCCTACCTTTCTCCCATGGTGCCCCAGCTAGGGGGC
CCGGCTCTTAGCCCCCTCTCTGGCCCCTCCGTGGGACCTTCCCTGGCCCAGTCCCCCACCTCCCTATCAGGCCAGAG
CTATGGCGCCTACAGCCCCGTGGATAGCTTGGAATTCAAGGACCCCACGGGCACCTGGAAATTCACCTACAATCCCA
TGGACCCTCTGGACTACAAGGATCAGAGTGCCTGGAAGTTTCAGATCTTG

ERG

ATGGCCAGCACTATTAAGGAAGCCTTATCAGTTGTGAGTGAGGACCAGTCGTTGTTTGAGTGTGCCTACGGAACGCC
ACACCTGGCTAAGACAGAGATGACCGCGTCCTCCTCCAGCGACTATGGACAGACTTCCAAGATGAGCCCACGCGTCC
CTCAGCAGGATTGGCTGTCTCAACCCCCAGCCAGGGTCACCATCAAAATGGAATGTAACCCTAGCCAGGTGAATGGC
TCAAGGAACTCTCCTGATGAATGCAGTGTGGCCAAAGGCGGGAAGATGGTGGGCAGCCCAGACACCGTTGGGATGAA
CTACGGCAGCTACATGGAGGAGAAGCACATGCCACCCCCAAACATGACCACGAACGAGCGCAGAGTTATCGTGCCAG
CAGATCCTACGCTATGGAGTACAGACCATGTGCGGCAGTGGCTGGAGTGGGCGGTGAAAGAATATGGCCTTCCAGAC
GTCAACATCTTGTTATTCCAGAACATCGATGGGAAGGAACTGTGCAAGATGACCAAGGACGACTTCCAGAGGCTCAC
CCCCAGCTACAATGCCGACATCCTTCTCTCACATCTCCACTACCTCAGAGAGACTCCTCTTCCACATTTGACTTCAG
ATGATGTTGATAAAGCCTTACAAAACTCTCCACGGTTAATGCATGCTAGAAACACAGGGGGTGCAGCTTTTATTTTC
CCAAATACTTCAGTATATCCTGAAGCTACGCAAAGAATTACAACTAGGCCAGATTTACCATATGAGCCCCCCAGGAG
ATCAGCCTGGACCGGTCACGGCCACCCCACGCCCCAGTCGAAAGCTGCTCAACCATCTCCTTCCACAGTGCCCAAAA
CTGAAGACCAGCGTCCTCAGTTAGATCCTTATCAGATTCTTGGACCAACAAGTAGCCGCCTTGCAAATCCAGGCAGT
GGCCAGATCCAGCTTTGGCAGTTCCTCCTGGAGCTCCTGTCGGACAGCTCCAACTCCAGCTGCATCACCTGGGAAGG
CACCAACGGGGAGTTCAAGATGACGGATCCCGACGAGGTGGCCCGGCGCTGGGGAGAGCGGAAGAGCAAACCCAACA
TGAACTACGATAAGCTCAGCCGCGCCCTCCGTTACTACTATGACAAGAACATCATGACCAAGGTCCATGGGAAGCGC
TACGCCTACAAGTTCGACTTCCACGGGATCGCCCAGGCCCTCCAGCCCCACCCCCCGGAGTCATCTCTGTACAAGTA
CCCCTCAGACCTCCCGTACATGGGCTCCTATCACGCCCACCCACAGAAGATGAACTTTGTGGCGCCCCACCCTCCAG
CCCTCCCCGTGACATCTTCCAGTTTTTTTGCTGCCCCAAACCCATACTGGAATTCACCAACTGGGGGTATATACCCC
AACACTAGGCTCCCCACCAGCCATATGCCTTCTCATCTGGGCACTTACTAC

ESRRG

ATGTCAAACAAAGATCGACACATTGATTCCAGCTGTTCGTCCTTCATCAAGACGGAACCTTCCAGCCCAGCCTCCCT
GACGGACAGCGTCAACCACCACAGCCCTGGTGGCTCTTCAGACGCCAGTGGGAGCTACAGTTCAACCATGAATGGCC
ATCAGAACGGACTTGACTCGCCACCTCTCTACCCTTCTGCTCCTATCCTGGGAGGTAGTGGGCCTGTCAGGAAACTG
TATGATGACTGCTCCAGCACCATTGTTGAAGATCCCCAGACCAAGTGTGAATACATGCTCAACTCGATGCCCAAGAG
ACTGTGTTTAGTGTGTGGTGACATCGCTTCTGGGTACCACTATGGGGTAGCATCATGTGAAGCCTGCAAGGCATTCT
TCAAGAGGACAATTCAAGGCAATATAGAATACAGCTGCCCTGCCACGAATGAATGTGAAATCACAAAGCGCAGACGT

AAATCCTGCCAGGCTTGCCGCTTCATGAAGTGTTTAAAAGTGGGCATGCTGAAAGAAGGGGTGCGTCTTGACAGAGT

ACGTGGAGGTCGGCAGAAGTACAAGCGCAGGATAGATGCGGAGAACAGCCCATACCTGAACCCTCAGCTGGTTCAGC

CAGCCAAAAAGCCATTGCTCTGGTCTGATCCTGCAGATAACAAGATTGTCTCACATTTGTTGGTGGCTGAACCGGAG

AAGATCTATGCCATGCCTGACCCTACTGTCCCCGACAGTGACATCAAAGCCCTCACTACACTGTGTGACTTGGCCGA

CCGAGAGTTGGTGGTTATCATTGGATGGGCGAAGCATATTCCAGGCTTCTCCACGCTGTCCCTGGCGGACCAGATGA

GCCTTCTGCAGAGTGCTTGGATGGAAATTTTGATCCTTGGTGTCGTATACCGGTCTCTTTCGTTTGAGGATGAACTT

GTCTATGCAGACGATTATATAATGGACGAAGACCAGTCCAAATTAGCAGGCCTTCTTGATCTAAATAATGCTATCCT

GCAGCTGGTAAAGAAATACAAGAGCATGAAGCTGGAAAAAGAAGAATTTGTCACCCTCAAAGCTATAGCTCTTGCTA

ATTCAGACTCCATGCACATAGAAGATGTTGAAGCCGTTCAGAAGCTTCAGGATGTCTTACATGAAGCGCTGCAGGAT

TATGAAGCTGGCCAGCACATGGAAGACCCTCGTCGAGCTGGCAAGATGCTGATGACACTGCCACTCCTGAGGCAGAC

CTCTACCAAGGCCGTGCAGCATTTCTACAACATCAAACTAGAAGGCAAAGTCCCAATGCACAAACTTTTTTTGGAAA

TGTTGGAGGCCAAGGTC

ETV2

ATGGATCTTTGGAACTGGGATGAAGCTTCCCCTCAAGAAGTTCCCCCCGGAAATAAACTCGCGGGGCTTGGAAGACT

CCCTCGCCTTCCGCAACGCGTCTGGGGCGGATGCCCTGGTGGAGCCTCAGCGGACCCAAACCCTTTGTCTCCAGCGG

AGGGGGCAAAGTTGGGTTTCTGCTTCCCGGATCTTGCTTTGCAAGGCGATACTCCAACGGCGACGGCAGAGACCTGT

TGGAAAGGCACCAGTAGCTCCCTGGCCAGCTTTCCGCAGCTCGATTGGGGGTCAGCCCTTCTCCATCCCGAAGTTCC

CTGGGGGGCGGAACCCGACTCCCAAGCCCTTCCCTGGAGTGGTGATTGGACAGATATGGCATGCACAGCCTGGGACA

GTTGGTCCGGGGCGTCACAGACATTGGGACCAGCCCCACTTGGACCGGGGCCTATCCCCGCAGCAGGAAGCGAAGGA

GCTGCTGGTCAGAACTGTGTGCCCGTGGCTGGTGAGGCTACCAGTTGGTCCAGGGCCCAGGCAGCAGGCAGTAACAC

CAGCTGGGATTGCTCAGTGGGGCCTGACGGGGATACTTATTGGGGCTCTGGTCTTGGTGGAGAACCGAGAACGGACT

GTACGATAAGTTGGGGCGGTCCAGCTGGGCCTGATTGTACTACGTCATGGAATCCTGGCTTGCACGCCGGCGGCACG

ACAAGCCTTAAGAGATATCAAAGTTCAGCCCTTACAGTTTGCTCAGAACCTTCCCCGCAAAGTGACCGAGCGTCACT

GGCGCGATGTCCTAAAACTAATCATCGAGGGCCGATCCAGTTGTGGCAGTTTTTGCTTGAACTCCTTCACGATGGCG

CGAGGAGCAGTTGCATCAGATGGACCGGTAACAGCAGGGAGTTCCAATTGTGTGACCCCAAGGAAGTGGCTCGACTG

TGGGGTGAGCGCAAACGGAAGCCTGGTATGAATTACGAAAAGTTGAGTAGGGGTTTGCGATATTACTATAGGCGCGA

CATCGTTCGAAAGTCCGGTGGTCGAAAGTACACATACAGATTCGGCGGTCGCGTACCATCTCTTGCATACCCTGATT
GCGCAGGCGGGGGTAGGGGTGCGGAAACACAA

FLI1

ATGGACGGGACTATTAAGGAGGCTCTGTCGGTGGTGAGCGACGACCAGTCCCTCTTTGACTCAGCGTACGGAGCGGC
AGCCCATCTCCCCAAGGCCGACATGACTGCCTCGGGGAGTCCTGACTACGGGCAGCCCCACAAGATCAACCCCCTCC
CACCACAGCAGGAGTGGATCAATCAGCCAGTGAGGGTCAACGTCAAGCGGGAGTATGACCACATGAATGGATCCAGG
GAGTCTCCGGTGGACTGCAGCGTTAGCAAATGCAGCAAGCTGGTGGGCGGAGGCGAGTCCAACCCCATGAACTACAA
CAGCTATATGGACGAGAAGAATGGCCCCCCTCCTCCCAACATGACCACCAACGAGAGGAGAGTCATCGTCCCCGCAG
ACCCCACACTGTGGACACAGGAGCATGTGAGGCAATGGCTGGAGTGGGCCATAAAGGAGTACAGCTTGATGGAGATC
GACACATCCTTTTTCCAGAACATGGATGGCAAGGAACTGTGTAAAATGAACAAGGAGGACTTCCTCCGCGCCACCAC
CCTCTACAACACGGAAGTGCTGTTGTCACACCTCAGTTACCTCAGGGAAAGTTCACTGCTGGCCTATAATACAACCT
CCCACACCGACCAATCCTCACGATTGAGTGTCAAAGAAGACCCTTCTTATGACTCAGTCAGAAGAGGAGCTTGGGGC
AATAACATGAATTCTGGCCTCAACAAAAGTCCTCCCCTTGGAGGGGCACAAACGATCAGTAAGAATACAGAGCAACG
GCCCCAGCCAGATCCGTATCAGATCCTGGGCCCGACCAGCAGTCGCCTAGCCAACCCTGGAAGCGGGCAGATCCAGC
TGTGGCAATTCCTCCTGGAGCTGCTCTCCGACAGCGCCAACGCCAGCTGTATCACCTGGGAGGGGACCAACGGGGAG
TTCAAAATGACGGACCCCGATGAGGTGGCCAGGCGCTGGGGCGAGCGGAAAAGCAAGCCCAACATGAATTACGACAA
GCTGAGCCGGGCCCTCCGTTATTACTATGATAAAAACATTATGACCAAAGTGCACGGCAAAAGATATGCTTACAAAT
TTGACTTCCACGGCATTGCCCAGGCTCTGCAGCCACATCCGACCGAGTCGTCCATGTACAAGTACCCTTCTGACATC
TCCTACATGCCTTCCTACCATGCCCACCAGCAGAAGGTGAACTTTGTCCCTCCCCATCCATCCTCCATGCCTGTCAC
TTCCTCCAGCTTCTTTGGAGCCGCATCACAATACTGGACCTCCCCCACGGGGGGAATCTACCCCAACCCCAACGTCC
CCCGCCATCCTAACACCCACGTGCCTTCACACTTAGGCAGCTACTAC

FOXA1

ATGTTGGGCACCGTGAAGATGGAGGGGCATGAGACAAGCGACTGGAATTCCTACTACGCGGATACCCAAGAAGCGTA
TTCTTCAGTTCCCGTAAGCAATATGAACTCCGGATTGGGGAGCATGAATAGTATGAACACGTATATGACAATGAATA
CGATGACCACCAGCGGCAACATGACACCGGCCTCCTTTAATATGTCATATGCGAACCCTGGTCTTGGCGCTGGCCTC
TCACCAGGTGCGGTCGCTGGAATGCCCGGGGGGAGCGCCGGAGCGATGAACTCCATGACCGCTGCGGGCGTGACGGC

CATGGGTACGGCCCTGTCACCCAGTGGAATGGGAGCTATGGGGGCCCAGCAAGCCGCCTCAATGAATGGATTGGGGC

CCTATGCCGCGGCGATGAATCCCTGCATGTCCCCTATGGCTTATGCCCCCAGCAATTTGGGTCGCAGTAGAGCCGGC

GGTGGTGGCGATGCCAAAACCTTCAAGCGAAGTTATCCTCATGCGAAGCCTCCTTATTCATATATATCCTTGATTAC

GATGGCGATACAGCAGGCCCCGTCTAAGATGCTGACTCTGAGTGAGATATACCAGTGGATCATGGACCTTTTTCCTT

ACTACCGGCAAAACCAACAGAGATGGCAAAACTCAATACGCCATAGCCTTTCCTTCAATGATTGCTTTGTCAAAGTC

GCTCGGAGCCCTGACAAGCCCGGTAAAGGGTCCTATTGGACCCTTCATCCAGATAGCGGCAATATGTTCGAGAATGG

TTGTTATCTTAGACGGCAGAAACGATTCAAATGTGAGAAACAGCCAGGTGCCGGCGGTGGTGGCGGCAGCGGTTCAG

GCGGAAGTGGTGCCAAGGGTGGGCCTGAGTCTAGAAAAGACCCCAGCGGAGCAAGCAATCCAAGCGCGGACTCTCCC

CTGCACCGCGGTGTTCATGGTAAGACAGGTCAGCTTGAGGGGGCGCCTGCTCCAGGCCCGGCTGCGTCACCGCAAAC

ACTGGACCATAGTGGAGCTACAGCGACCGGAGGTGCTTCAGAACTCAAGACGCCTGCGTCCTCCACTGCGCCTCCGA

TCTCCAGTGGTCCCGGTGCACTTGCCTCTGTTCCTGCATCTCATCCAGCACACGGACTCGCGCCGCACGAGTCCCAG

CTCCATTTGAAAGGGGACCCACACTACAGCTTTAACCACCCATTCTCTATTAACAATTTGATGTCATCCTCAGAACA

GCAGCATAAACTCGACTTCAAAGCCTATGAACAGGCCCTGCAGTATTCTCCATATGGCTCTACACTTCCTGCTTCTC

TTCCATTGGGGTCTGCAAGTGTGACAACGCGCTCCCCAATCGAGCCAAGTGCCCTCGAGCCTGCTTATTATCAAGGA

GTATATTCCCGACCAGTTTTGAATACAAGT

FOXA2

ATGCTGGGAGCGGTGAAGATGGAAGGGCACGAGCCGTCCGACTGGAGCAGCTACTATGCAGAGCCCGAGGGCTACTC

CTCCGTGAGCAACATGAACGCCGGCCTGGGGATGAACGGCATGAACACGTACATGAGCATGTCGGCGGCCGCCATGG

GCAGCGGCTCGGGCAACATGAGCGCGGGCTCCATGAACATGTCGTCGTACGTGGGCGCTGGCATGAGCCCGTCCCTG

GCGGGGATGTCCCCCGGCGCGGGCGCCATGGCGGGCATGGGCGGCTCGGCCGGGGCGGCTGGCGTGGCGGGCATGGG

GCCGCACTTGAGTCCCAGCCTGAGCCCGCTCGGGGGGCAGGCGGCCGGGGCCATGGGCGGCCTGGCCCCCTACGCCA

ACATGAACTCCATGAGCCCCATGTACGGGCAGGCGGGCCTGAGCCGCGCCCGCGACCCCAAGACCTACAGGCGCAGC

TACACGCACGCAAAGCCGCCCTACTCGTACATCTCGCTCATCACCATGGCCATCCAGCAGAGCCCCAACAAGATGCT

GACGCTGAGCGAGATCTACCAGTGGATCATGGACCTCTTCCCCTTCTACCGGCAGAACCAGCAGCGCTGGCAGAACT

CCATCCGCCACTCGCTCTCCTTCAACGACTGTTTCCTGAAGGTGCCCCGCTCGCCCGACAAGCCCGGCAAGGGCTCC

TTCTGGACCCTGCACCCTGACTCGGGCAACATGTTCGAGAACGGCTGCTACCTGCGCCGCCAGAAGCGCTTCAAGTG

CGAGAAGCAGCTGGCGCTGAAGGAGGCCGCAGGCGCCGCCGGCAGCGGCAAGAAGGCGGCCGCCGGGGCCCAGGCCT

CACAGGCTCAACTCGGGGAGGCCGCCGGGCCGGCCTCCGAGACTCCGGCGGGCACCGAGTCGCCTCACTCGAGCGCC

TCCCCGTGCCAGGAGCACAAGCGAGGGGGCCTGGGAGAGCTGAAGGGGACGCCGGCTGCGGCGCTGAGCCCCCCAGA

GCCGGCGCCCTCTCCCGGGCAGCAGCAGCAGGCCGCGGCCCACCTGCTGGGCCCGCCCCACCACCCGGGCCTGCCGC

CTGAGGCCCACCTGAAGCCGGAACACCACTACGCCTTCAACCACCCGTTCTCCATCAACAACCTCATGTCCTCGGAG

CAGCAGCACCACCACAGCCACCACCACCACCAGCCCCACAAAATGGACCTCAAGGCCTACGAACAGGTGATGCACTA

CCCCGGCTACGGTTCCCCCATGCCTGGCAGCTTGGCCATGGGCCCGGTCACGAACAAAACGGGCCTGGACGCCTCGC

CCCTGGCCGCAGATACCTCCTACTACCAGGGGGTGTACTCCCGGCCCATTATGAACTCCTCTTTG

FOXA3

ATGCTGGGCTCAGTGAAGATGGAGGCCCATGACCTGGCCGAGTGGAGCTACTACCCGGAGGCGGGCGAGGTCTACTC

GCCGGTGACCCCAGTGCCCACCATGGCCCCCCTCAACTCCTACATGACCCTGAATCCTCTAAGCTCTCCCTATCCCC

CTGGGGGGCTCCCTGCCTCCCCACTGCCCTCAGGACCCCTGGCACCCCCAGCACCTGCAGCCCCCCTGGGGCCCACT

TTCCCAGGCCTGGGTGTCAGCGGTGGCAGCAGCAGCTCCGGGTACGGGGCCCCGGGTCCTGGGCTGGTGCACGGGAA

GGAGATGCCGAAGGGGTATCGGCGGCCCCTGGCACACGCCAAGCCACCGTATTCCTATATCTCACTCATCACCATGG

CCATCCAGCAGGCGCCGGGCAAGATGCTGACCTTGAGTGAAATCTACCAGTGGATCATGGACCTCTTCCCTTACTAC

CGGGAGAATCAGCAGCGCTGGCAGAACTCCATTCGCCACTCGCTGTCTTTCAACGACTGCTTCGTCAAGGTGGCGCG

TTCCCCAGACAAGCCTGGCAAGGGCTCCTACTGGGCCCTACACCCCAGCTCAGGGAACATGTTTGAGAATGGCTGCT

ACCTGCGCCGCCAGAAACGCTTCAAGCTGGAGGAGAAGGTGAAAAAAGGGGGCAGCGGGGCTGCCACCACCACCAGG

AACGGGACAGGGTCTGCTGCCTCGACCACCACCCCCGCGGCCACAGTCACCTCCCCGCCCCAGCCCCCGCCTCCAGC

CCCTGAGCCTGAGGCCCAGGGCGGGGAAGATGTGGGGGCTCTGGACTGTGGCTCACCCGCTTCCTCCACACCCTATT

TCACTGGCCTGGAGCTCCCAGGGGAGCTGAAGCTGGACGCGCCCTACAACTTCAACCACCCTTTCTCCATCAACAAC

CTAATGTCAGAACAGACACCAGCACCTCCCAAACTGGACGTGGGGTTTGGGGGCTACGGGGCTGAAGGTGGGGAGCC

TGGAGTCTACTACCAGGGCCTCTATTCCCGCTCTTTGCTTAATGCATCC

FOXP1

ATGATGCAAGAATCTGGGACTGAGACAAAAAGTAACGGTTCAGCCATCCAGAATGGGTCGGGCGGCAGCAACCACTT

ACTAGAGTGCGGCGGTCTTCGGGAGGGGCGGTCCAACGGAGAGACGCCGGCCGTGGACATCGGGGCAGCTGACCTCG

CCCACGCCCAGCAGCAGCAGCAACAGTGGCATCTCATAAACCATCAGCCCTCTAGGAGTCCCAGCAGTTGGCTTAAG

AGACTAATTTCAAGCCCTTGGGAGTTGGAAGTCCTGCAGGTCCCCTTGTGGGGAGCAGTTGCTGAGACGAAGATGAG

TGGACCTGTGTGTCAGCCTAACCCTTCCCCATTT

GATA1

ATGGAGTTCCCTGGCCTGGGGTCCCTGGGGACCTCAGAGCCCCTCCCCCAGTTTGTGGATCCTGCTCTGGTGTCCTC

CACACCAGAATCAGGGGTTTTCTTCCCCTCTGGGCCTGAGGGCTTGGATGCAGCAGCTTCCTCCACTGCCCCGAGCA

CAGCCACCGCTGCAGCTGCGGCACTGGCCTACTACAGGGACGCTGAGGCCTACAGACACTCCCCAGTCTTTCAGGTG

TACCCATTGCTCAACTGTATGGAGGGGATCCCAGGGGGCTCACCATATGCCGGCTGGGCCTACGGCAAGACGGGGCT

CTACCCTGCCTCAACTGTGTGTCCCACCCGCGAGGACTCTCCTCCCCAGGCCGTGGAAGATCTGGATGGAAAAGGCA

GCACCAGCTTCCTGGAGACTTTGAAGACAGAGCGGCTGAGCCCAGACCTCCTGACCCTGGGACCTGCACTGCCTTCA

TCACTCCCTGTCCCCAATAGTGCTTATGGGGGCCCTGACTTTTCCAGTACCTTCTTTTCTCCCACCGGGAGCCCCCT

CAATTCAGCAGCCTATTCCTCTCCCAAGCTTCGTGGAACTCTCCCCCTGCCTCCCTGTGAGGCCAGGGAGTGTGTGA

ACTGCGGAGCAACAGCCACTCCACTGTGGCGGAGGGACAGGACAGGCCACTACCTATGCAACGCCTGCGGCCTCTAT

CACAAGATGAATGGGCAGAACAGGCCCCTCATCCGGCCCAAGAAGCGCCTGATTGTCAGTAAACGGGCAGGTACTCA

GTGCACCAACTGCCAGACGACCACCACGACACTGTGGCGGAGAAATGCCAGTGGGGATCCCGTGTGCAATGCCTGCG

GCCTCTACTACAAGCTACACCACCAGCACTACTGTGGTGGCTCCGCTCAGCTCATGAGGGCACAGAGCATGGCCTCC

AGAGGAGGGGTGGTGTCCTTCTCCTCTTGTAGCCAGAATTCTGGACAACCCAAGTCTCTGGGCCCCAGGCACCCCCT

GGCT

GATA2

ATGGAGGTGGCGCCGGAGCAGCCGCGCTGGATGGCGCACCCGGCCGTGCTGAATGCGCAGCACCCCGACTCACACCA

CCCGGGCCTGGCGCACAACTACATGGAACCCGCGCAGCTGCTGCCTCCAGACGAGGTGGACGTCTTCTTCAATCACC

TCGACTCGCAGGGCAACCCCTACTATGCCAACCCCGCTCACGCGCGGGCGCGCGTCTCCTACAGCCCCGCGCACGCC

CGCCTGACCGGAGGCCAGATGTGCCGCCCACACTTGTTGCACAGCCCGGGTTTGCCCTGGCTGGACGGGGGCAAAGC

AGCCCTCTCTGCCGCTGCGGCCCACCACCACAACCCCTGGACCGTGAGCCCCTTCTCCAAGACGCCACTGCACCCCT

CAGCTGCTGGAGGCCCTGGAGGCCCACTCTCTGTGTACCCAGGGGCTGGGGGTGGGAGCGGGGGAGGCAGCGGGAGC

TCAGTGGCCTCCCTCACCCCTACAGCAACCCACTCTGGCTCCCACCTTTTCGGCTTCCCACCCACGCCACCCAAAGA

AGTGTCTCCTGACCCTAGCACCACGGGGGCTGCGTCTCCAGCCTCATCTTCCGCGGGGGGTAGTGCAGCCCGAGGAG

AGGACAAGGACGGCGTCAAGTACCAGGTGTCACTGACGGAGAGCATGAAGATGGAAAGTGGCAGTCCCCTGCGCCCA

GGCCTAGCTACTATGGGCACCCAGCCTGCTACACACCACCCCATCCCCACCTACCCCTCCTATGTGCCGGCGGCTGC

CCACGACTACAGCAGCGGACTCTTCCACCCCGGAGGCTTCCTGGGGGGACCGGCCTCCAGCTTCACCCCTAAGCAGC

GCAGCAAGGCTCGTTCCTGTTCAGAAGGCCGGGAGTGTGTCAACTGTGGGGCCACAGCCACCCCTCTCTGGCGGCGG

GACGGCACCGGCCACTACCTGTGCAATGCCTGTGGCCTCTACCACAAGATGAATGGGCAGAACCGACCACTCATCAA

GCCCAAGCGAAGACTGTCGGCCGCCAGAAGAGCCGGCACCTGTTGTGCAAATTGTCAGACGACAACCACCACCTTAT

GGCGCCGAAACGCCAACGGGGACCCTGTCTGCAACGCCTGTGGCCTCTACTACAAGCTGCACAATGTTAACAGGCCA

CTGACCATGAAGAAGGAAGGGATCCAGACTCGGAACCGGAAGATGTCCAACAAGTCCAAGAAGAGCAAGAAAGGGGC

GGAGTGCTTCGAGGAGCTGTCAAAGTGCATGCAGGAGAAGTCATCCCCCTTCAGTGCAGCTGCCCTGGCTGGACACA

TGGCACCTGTGGGCCACCTCCCGCCCTTCAGCCACTCCGGACACATCCTGCCCACTCCGACGCCCATCCACCCCTCC

TCCAGCCTCTCCTTCGGCCACCCCCACCCGTCCAGCATGGTGACCGCCATGGGC

GATA4

ATGTACCAGAGCCTGGCTATGGCTGCTAATCATGGACCTCCCCCTGGAGCCTATGAAGCCGGAGGACCTGGCGCTTT

TATGCATGGAGCTGGCGCCGCTTCTTCTCCCGTGTATGTGCCTACACCTAGAGTGCCCAGCAGCGTGCTGGGCCTTT

CTTATCTTCAGGGAGGAGGAGCAGGATCTGCTTCTGGCGGAGCTTCAGGCGGATCTTCTGGAGGCGCTGCTTCAGGT

GCTGGACCTGGAACTCAACAGGGATCTCCTGGATGGTCACAGGCAGGAGCTGATGGAGCCGCTTATACCCCTCCTCC

TGTGAGCCCCAGGTTTAGCTTTCCTGGCACAACAGGCTCTTTAGCTGCCGCTGCTGCTGCAGCCGCAGCTAGAGAAG

CAGCTGCATATTCTAGTGGCGGAGGAGCTGCTGGAGCCGGCTTAGCTGGAAGAGAGCAGTACGGAAGAGCCGGATTT

GCCGGAAGCTATAGCAGCCCTTACCCTGCCTATATGGCCGATGTTGGCGCATCTTGGGCAGCCGCCGCAGCAGCTTC

TGCAGGACCTTTTGACTCACCTGTGCTTCACTCTCTGCCTGGCAGAGCTAATCCTGCCGCCAGACATCCCAACCTGG

ACATGTTCGACGACTTCAGCGAGGGCAGAGAATGCGTGAACTGCGGAGCCATGAGCACCCCCCTTTGGAGAAGAGAC

GGCACCGGCCACTACCTTTGCAATGCCTGTGGCCTGTACCACAAGATGAACGGCATCAACAGACCCCTGATCAAGCC

CCAGAGAAGACTGAGCGCTAGCAGAAGAGTGGGCCTGTCCTGCGCCAATTGCCAGACCACAACCACCACACTGTGGA

GGAGAAATGCCGAGGGCGAGCCTGTGTGTAACGCCTGTGGACTGTACATGAAGCTGCACGGCGTGCCCAGACCTCTG

GCCATGAGAAAGGAGGGCATCCAGACCAGAAAGAGAAAGCCCAAGAACCTGAACAAGAGCAAGACCCCCGCTGCTCC

TTCTGGAAGCGAGAGCCTGCCTCCAGCCTCTGGAGCCAGCAGCAATAGCTCTAACGCCACCACATCTTCTTCTGAGG

AGATGAGGCCCATCAAAACCGAGCCAGGCCTGAGCAGCCACTACGGCCACAGCTCTAGCGTGAGCCAGACTTTTAGC

GTGTCTGCCATGTCAGGCCACGGACCTAGCATTCACCCTGTGCTGAGCGCCCTGAAGTTGAGCCCACAGGGCTATGC

TTCTCCTGTGTCTCAGAGCCCTCAGACCTCCAGCAAGCAGGACTCCTGGAATTCTCTGGTGCTGGCCGACAGCCACG

GCGATATCATCACCGCC

GATA6

ATGGCCCTGACCGACGGCGGATGGTGTCTCCCTAAAAGATTCGGCGCCGCTGGCGCTGATGCTTCTGACAGCAGAGC

CTTCCCCGCTAGGGAACCCAGCACACCACCTAGCCCCATCAGCAGCTCAAGCTCTAGCTGTAGCAGAGGCGGAGAGA

GAGGACCTGGAGGCGCTTCTAACTGCGGCACACCTCAGCTGGATACAGAAGCCGCCGCCGGACCACCAGCCAGATCT

CTTTTACTTAGCAGCTACGCCAGCCACCCTTTTGGCGCTCCTCATGGACCCTCTGCTCCTGGTGTGGCCGGACCTGG

CGGAAACCTGAGCTCTTGGGAGGACCTTCTGCTGTTTACCGACCTGGACCAGGCTGCCACCGCTAGCAAGCTTCTGT

GGAGCAGCAGGGGCGCTAAGCTGAGCCCTTTTGCCCCTGAGCAGCCCGAGGAGATGTACCAGACCCTGGCTGCTTTA

AGCTCTCAGGGACCTGCCGCTTATGACGGAGCCCCTGGTGGATTTGTTCACTCAGCGGCAGCAGCCGCAGCTGCTGC

AGCCGCTGCCAGCTCACCTGTGTATGTGCCTACCACAAGAGTGGGCAGCATGTTACCTGGACTTCCTTACCATCTGC

AGGGCAGCGGAAGCGGCCCTGCTAACCATGCCGGAGGAGCTGGAGCTCACCCCGGATGGCCTCAGGCTTCTGCAGAT

TCTCCTCCTTATGGATCTGGAGGAGGAGCAGCTGGAGGGGGAGCTGCAGGACCAGGTGGAGCCGGAAGCGCAGCAGC

ACATGTGTCTGCCAGATTTCCCTATAGCCCTAGCCCTCCTATGGCCAATGGCGCTGCTAGAGAACCCGGAGGATATG

CTGCGGCAGGCTCTGGCGGCGCTGGCGGAGTTTCTGGAGGTGGATCTTCACTGGCCGCTATGGGAGGAAGAGAGCCT

CAGTACTCTTCTCTGAGCGCCGCTAGACCACTGAACGGCACCTATCATCACCACCACCATCACCATCATCATCACCC

CAGCCCTTACTCCCCTTATGTGGGAGCCCCCCTTACACCCGCTTGGCCTGCCGGCCCTTTCGAGACACCTGTGCTGC

ACAGCCTTCAGTCTAGAGCTGGCGCACCTTTACCAGTGCCTAGAGGCCCCTCTGCCGACTTGCTGGAGGATCTGAGC

GAGAGCAGAGAGTGCGTGAACTGTGGCAGCATCCAGACACCCCTGTGGAGAAGAGACGGCACCGGCCACTACCTGTG

CAACGCTTGCGGCCTGTACAGCAAGATGAATGGGCTGAGCAGACCCCTGATCAAGCCCCAGAAGAGGGTGCCCAGCA

GCAGACGGCTGGGACTGAGCTGCGCCAACTGTCATACCACAACAACCACACTGTGGCGGAGAAACGCCGAGGGCGAG

CCCGTGTGTAACGCCTGCGGCCTTTACATGAAGCTGCACGGCGTGCCCAGACCTCTGGCCATGAAGAAGGAGGGAAT

CCAGACCAGAAAGAGAAAGCCCAAGAACATCAACAAGAGCAAGACCTGCAGCGGCAACAGCAACAACAGCATCCCCA

TGACCCCCACCAGCACATCTAGCAACAGCGACGACTGTAGCAAGAACACATCACCTACCACCCAGCCCACAGCTAGC

GGAGCCGGCGCCCCCGTGATGACAGGCGCCGGAGAGTCCACAAATCCCGAGAATAGCGAACTGAAGTACTCTGGACA

GGACGGACTGTATATCGGCGTGAGCCTGGCTTCTCCCGCCGAGGTGACCAGCTCTGTCAGACCTGACTCTTGGTGTG
CCCTCGCCCTGGCC

GLI1

ATGTTCAACTCGATGACCCCACCACCAATCAGTAGCTATGGCGAGCCCTGCTGTCTCCGGCCCCTCCCCAGTCAGGG
GGCCCCCAGTGTGGGGACAGAAGGACTGTCTGGCCCGCCCTTCTGCCACCAAGCTAACCTCATGTCCGGCCCCCACA
GTTATGGGCCAGCCAGAGAGACCAACAGCTGCACCGAGGGCCCACTCTTTTCTTCTCCCCGGAGTGCAGTCAAGTTG
ACCAAGAAGCGGGCACTGTCCATCTCACCTCTGTCGGATGCCAGCCTGGACCTGCAGACGGTTATCCGCACCTCACC
CAGCTCCCTCGTAGCTTTCATCAACTCGCGATGCACATCTCCAGGAGGCTCCTACGGTCATCTCTCCATTGGCACCA
TGAGCCCATCTCTGGGATTCCCAGCCCAGATGAATCACCAAAAAGGGCCCTCGCCTTCCTTTGGGGTCCAGCCTTGT
GGTCCCCATGACTCTGCCCGGGGTGGGATGATCCCACATCCTCAGTCCCGGGGACCCTTCCCAACTTGCCAGCTGAA
GTCTGAGCTGGACATGCTGGTTGGCAAGTGCCGGGAGGAACCCTTGGAAGGTGATATGTCCAGCCCCAACTCCACAG
GCATACAGGATCCCCTGTTGGGGATGCTGGATGGGCGGGAGGACCTCGAGAGAGAGGAGAAGCGTGAGCCTGAATCT
GTGTATGAAACTGACTGCCGTTGGGATGGCTGCAGCCAGGAATTTGACTCCCAAGAGCAGCTGGTGCACCACATCAA
CAGCGAGCACATCCACGGGGAGCGGAAGGAGTTCGTGTGCCACTGGGGGGGCTGCTCCAGGGAGCTGAGGCCCTTCA
AAGCCCAGTACATGCTGGTGGTTCACATGCGCAGACACACTGGCGAGAAGCCACACAAGTGCACGTTTGAAGGGTGC
CGGAAGTCATACTCACGCCTCGAAAACCTGAAGACGCACCTGCGGTCACACACGGGTGAGAAGCCATACATGTGTGA
GCACGAGGGCTGCAGTAAAGCCTTCAGCAATGCCAGTGACCGAGCCAAGCACCAGAATCGGACCCATTCCAATGAGA
AGCCGTATGTATGTAAGCTCCCTGGCTGCACCAAACGCTATACAGATCCTAGCTCGCTGCGAAAACATGTCAAGACA
GTGCATGGTCCTGACGCCCATGTGACCAAACGGCACCGTGGGGATGGCCCCCTGCCTCGGGCACCATCCATTTCTAC
AGTGGAGCCCAAGAGGGAGCGGGAAGGAGGTCCCATCAGGGAGGAAAGCAGACTGACTGTGCCAGAGGGTGCCATGA
AGCCACAGCCAAGCCCTGGGGCCCAGTCATCCTGCAGCAGTGACCACTCCCCGGCAGGGAGTGCAGCCAATACAGAC
AGTGGTGTGGAAATGACTGGCAATGCAGGGGGCAGCACTGAAGACCTCTCCAGCTTGGACGAGGGACCTTGCATTGC
TGGCACTGGTCTGTCCACTCTTCGCCGCCTTGAGAACCTCAGGCTGGACCAGCTACATCAACTCCGGCCAATAGGGA
CCCGGGGTCTCAAACTGCCCAGCTTGTCCCACACCGGTACCACTGTGTCCCGCCGCGTGGGCCCCCCAGTCTCTCTT
GAACGCCGCAGCAGCAGCTCCAGCAGCATCAGCTCTGCCTATACTGTCAGCCGCCGCTCCTCCCTGGCCTCTCCTTT
CCCCCCTGGCTCCCCACCAGAGAATGGAGCATCCTCCCTGCCTGGCCTTATGCCTGCCCAGCACTACCTGCTTCGGG
CAAGATATGCTTCAGCCAGAGGGGGTGGTACTTCGCCCACTGCAGCATCCAGCCTGGATCGGATAGGTGGTCTTCCC

ATGCCTCCTTGGAGAAGCCGAGCCGAGTATCCAGGATACAACCCCAATGCAGGGGTCACCCGGAGGGCCAGTGACCC

AGCCCAGGCTGCTGACCGTCCTGCTCCAGCTAGAGTCCAGAGGTTCAAGAGCCTGGGCTGTGTCCATACCCCACCCA

CTGTGGCAGGGGGAGGACAGAACTTTGATCCTTACCTCCCAACCTCTGTCTACTCACCACAGCCCCCCAGCATCACT

GAGAATGCTGCCATGGATGCTAGAGGGCTACAGGAAGAGCCAGAAGTTGGGACCTCCATGGTGGGCAGTGGTCTGAA

CCCCTATATGGACTTCCCACCTACTGATACTCTGGGATATGGGGGACCTGAAGGGGCAGCAGCTGAGCCTTATGGAG

CGAGGGGTCCAGGCTCTCTGCCTCTTGGGCCTGGTCCACCCACCAACTATGGCCCCAACCCCTGTCCCCAGCAGGCC

TCATATCCTGACCCCACCCAAGAAACATGGGGTGAGTTCCCTTCCCACTCTGGGCTGTACCCAGGCCCCAAGGCTCT

AGGTGGAACCTACAGCCAGTGTCCTCGACTTGAACATTATGGACAAGTGCAAGTCAAGCCAGAACAGGGGTGCCCAG

TGGGGTCTGACTCCACAGGACTGGCACCCTGCCTCAATGCCCACCCCAGTGAGGGGCCCCCACATCCACAGCCTCTC

TTTTCCCATTACCCCCAGCCCTCTCCTCCCCAATATCTCCAGTCAGGCCCCTATACCCAGCCACCCCCTGATTATCT

TCCTTCAGAACCCAGGCCTTGCCTGGACTTTGATTCCCCCACCCATTCCACAGGGCAGCTCAAGGCTCAGCTTGTGT

GTAATTATGTTCAATCTCAACAGGAGCTACTGTGGGAGGGTGGGGGCAGGGAAGATGCCCCCGCCCAGGAACCTTCC

TACCAGAGTCCCAAGTTTCTGGGGGGTTCCCAGGTTAGCCCAAGCCGTGCTAAAGCTCCAGTGAACACATATGGACC

TGGCTTTGGACCCAACTTGCCCAATCACAAGTCAGGTTCCTATCCCACCCCTTCACCATGCCATGAAAATTTTGTAG

TGGGGGCAAATAGGGCTTCACATAGGGCAGCAGCACCACCTCGACTTCTGCCCCCATTGCCCACTTGCTATGGGCCT

CTCAAAGTGGGAGGCACAAACCCCAGCTGTGGTCATCCTGAGGTGGGCAGGCTAGGAGGGGGTCCTGCCTTGTACCC

TCCTCCCGAAGGACAGGTATGTAACCCCCTGGACTCTCTTGATCTTGACAACACTCAGCTGGACTTTGTGGCTATTC

TGGATGAGCCCCAGGGGCTGAGTCCTCCTCCTTCCCATGATCAGCGGGGCAGCTCTGGACATACCCCACCTCCCTCT

GGGCCCCCCAACATGGCTGTGGGCAACATGAGTGTCTTACTGAGATCCCTACCTGGGGAAACAGAATTCCTCAACTC

TAGTGCC

HAND2

ATGAGTCTGGTAGGTGGTTTTCCCCACCACCCGGTGGTGCACCACGAGGGCTACCCGTTTGCCGCCGCCGCCGCCGC

CAGCCGCTGCAGCCATGAGGAGAACCCCTACTTCCATGGCTGGCTCATCGGCCACCCCGAGATGTCGCCCCCCGACT

ACAGCATGGCCCTGTCCTACAGCCCCGAGTATGCCAGCGGCACCGCCAACCGCAAGGAGCGGCGCAGGACTCAGAGC

ATCAACAGCGCCTTCGCCGAACTGCGCGAGTGCATCCCCAACGTACCCGCCGACACCAAACTCTCCAAAATCAAGAC

CCTGCGCCTGGCCACCAGCTACATCGCCTACCTCATGGACCTGCTGGCCAAGGACGACCAGAATGGCGAGGCGGAGG

CCTTCAAGGCAGAGATCAAGAAGACCGACGTGAAAGAGGAGAAGAGGAAGAAGGAGCTGAACGAAATCTTGAAAAGC

ACAGTGAGCAGCAACGACAAGAAAACCAAAGGCCGGACGGGCTGGCCGCAGCACGTCTGGGCCCTGGAGCTCAAGCA

G

HNF1A

ATGGTTTCTAAACTGAGCCAGCTGCAGACGGAGCTCCTGGCGGCCCTGCTGGAGTCAGGGCTGAGCAAAGAGGCACT

GCTCCAGGCACTGGGTGAGCCGGGGCCCTACCTCCTGGCTGGAGAAGGCCCCCTGGACAAGGGGGAGTCCTGCGGCG

GCGGTCGAGGGGAGCTGGCTGAGCTGCCCAATGGGCTGGGGGGAGACTCGGGGCTCCGAGGACGAGACGGACGACGAT

GGGGAAGACTTCACGCCACCCATCCTCAAAGAGCTGGAGAACCTCAGCCCTGAGGAGGCGGCCCACCAGAAAGCCGT

GGTGGAGACCCTTCTGCAGGAGGACCCGTGGCGTGTGGCGAAGATGGTCAAGTCCTACCTGCAGCAGCACAACATCC

CACAGCGGGAGGTGGTCGATACCACTGGCCTCAACCAGTCCCACCTGTCCCAACACCTCAACAAGGGCACTCCCATG

AAGACGCAGAAGCGGGCCGCCCTGTACACCTGGTACGTCCGCAAGCAGCGAGAGGTGGCGCAGCAGTTCACCCATGC

AGGGCAGGGAGGGCTGATTGAAGAGCCCACAGGTGATGAGCTACCAACCAAGAAGGGGCGGAGGAACCGTTTCAAGT

GGGGCCCAGCATCCCAGCAGATCCTGTTCCAGGCCTATGAGAGGCAGAAGAACCCTAGCAAGGAGGAGCGAGAGACG

CTAGTGGAGGAGTGCAATAGGGCGGAATGCATCCAGAGAGGGGTGTCCCCATCACAGGCACAGGGGCTGGGCTCCAA

CCTCGTCACGGAGGTGCGTGTCTACAACTGGTTTGCCAACCGGCGCAAAGAAGAAGCCTTCCGGCACAAGCTGGCCA

TGGACACGTACAGCGGGCCCCCCCCAGGGCCAGGCCCGGGACCTGCGCTGCCCGCTCACAGCTCCCCTGGCCTGCCT

CCACCTGCCCTCTCCCCCAGTAAGGTCCACGGTGTGCGCTATGGACAGCCTGCGACCAGTGAGACTGCAGAAGTACC

CTCAAGCAGCGGCGGTCCCTTAGTGACAGTGTCTACACCCCTCCACCAAGTGTCCCCCACGGGCCTGGAGCCCAGCC

ACAGCCTGCTGAGTACAGAAGCCAAGCTGGTCTCAGCAGCTGGGGGCCCCCTCCCCCCTGTCAGCACCCTGACAGCA

CTGCACAGCTTGGAGCAGACATCCCCAGGCCTCAACCAGCAGCCCCAGAACCTCATCATGGCCTCACTTCCTGGGGT

CATGACCATCGGGCCTGGTGAGCCTGCCTCCCTGGGTCCTACGTTCACCAACACAGGTGCCTCCACCCTGGTCATCG

GCCTGGCCTCCACGCAGGCACAGAGTGTGCCGGTCATCAACAGCATGGGCAGCAGCCTGACCACCCTGCAGCCCGTC

CAGTTCTCCCAGCCGCTGCACCCCTCCTACCAGCAGCCGCTCATGCCACCTGTGCAGAGCCATGTGACCCAGAGCCC

CTTCATGGCCACCATGGCTCAGCTGCAGAGCCCCCACGCCCTCTACAGCCACAAGCCCGAGGTGGCCCAGTACACCC

ACACAGGCCTGCTCCCGCAGACTATGCTCATCACCGACACCACCAACCTGAGCGCCCTGGCCAGCCTCACGCCCACC

AAGCAGGTCTTCACCTCAGACACTGAGGCCTCCAGTGAGTCCGGGCTTCACACGCCGGCATCTCAGGCCACCACCCT

CCACGTCCCCAGCCAGGACCCTGCCGGCATCCAGCACCTGCAGCCGGCCCACCGGCTCAGCGCCAGCCCCACAGTGT

CCTCCAGCAGCCTGGTGCTGTACCAGAGCTCAGACTCCAGCAATGGCCAGAGCCACCTGCTGCCATCCAACCACAGC

GTCATCGAGACCTTCATCTCCACCCAGATGGCCTCTTCCTCCCAGTTG

HNF1B

ATGGTTAGCAAACTGACATCCCTCCAGCAGGAACTTCTTTCTGCCCTCCTCTCCAGTGGGGTAACCAAAGAGGTACT

GGTCCAGGCTTTGGAGGAGTTGCTCCCCTCACCGAATTTTGGTGTAAAGTTGGAGACTCTCCCCCTCTCCCCTGGTT

CTGGAGCAGAGCCGGATACTAAACCGGTATTTCATACGCTTACAAACGGACACGCAAAGGGTCGGCTTTCAGGTGAC

GAAGGGTCTGAGGACGGCGATGATTATGACACCCCGCCCATCCTCAAAGAACTGCAGGCCCTTAATACAGAGGAAGC

GGCGGAGCAGCGAGCTGAAGTTGACAGAATGCTCTCAGAAGATCCGTGGAGAGCTGCGAAAATGATTAAGGGATATA

TGCAGCAACATAACATTCCCCAGAGAGAGGTAGTTGATGTTACCGGCCTTAACCAGAGCCACCTGTCTCAGCATCTC

AATAAGGGTACTCCTATGAAAACACAGAAGCGAGCGGCCCTTTACACATGGTACGTGCGGAAGCAACGAGAAATTCT

CCGACAGTTCAATCAGACAGTACAATCTTCAGGGAACATGACGGATAAAAGCTCACAGGATCAGCTCTTGTTTCTCT

TCCCCGAGTTCAGCCAACAGTCCCACGGTCCAGGTCAATCTGATGATGCTTGCAGTGAACCTACAAACAAAAAAATG

AGGAGGAACAGGTTTAAATGGGGACCGGCCTCTCAGCAGATACTGTACCAAGCGTACGATCGGCAGAAAAACCCAAG

CAAAGAGGAGCGCGAGGCATTGGTCGAGGAGTGTAATCGGGCCGAGTGCTTGCAACGGGGTGTAAGTCCTAGCAAAG

CCCATGGTCTCGGCTCAAACTTGGTCACGGAGGTGAGGGTATATAATTGGTTTGCCAACAGGCGGAAGGAGGAAGCA

TTCCGGCAAAAGCTGGCGATGGATGCCTACTCAAGCAACCAGACACATAGCCTCAACCCTCTGTTGTCACACGGGTC

CCCTCATCACCAACCTTCTTCCTCTCCACCCAACAAACTTTCTGGTGTCCGATATTCCCAGCAGGGGAACAACGAGA

TAACATCTTCCTCTACTATAAGTCATCACGGAAATTCTGCAATGGTAACGTCACAGAGTGTGTTGCAACAGGTATCA

CCCGCGTCTCTTGATCCAGGCCACAATCTGTTGAGCCCTGACGGAAAGATGATCTCTGTTTCTGGTGGCGGACTCCC

GCCGGTCTCCACACTTACCAACATACATAGTCTCAGTCATCATAATCCTCAGCAGAGCCAAAACCTGATTATGACTC

CTCTTAGCGGAGTGATGGCTATTGCGCAATCTTTGAACACCTCACAAGCACAATCTGTACCCGTCATAAACAGCGTA

GCGGGCTCATTGGCGGCGCTCCAACCAGTGCAGTTCTCCCAGCAGCTCCATTCACCCCATCAACAGCCTCTGATGCA

GCAGAGCCCTGGTAGTCACATGGCTCAACAGCCGTTCATGGCAGCTGTCACTCAGCTCCAGAACTCCCATATGTATG

CCCACAAGCAAGAACCACCACAATACAGTCACACATCAAGATTCCCCAGTGCTATGGTTGTTACTGACACATCCTCT

ATCTCAACTCTGACGAACATGTCCAGTAGTAAACAATGTCCTCTGCAAGCATGG

HNF4A

ATGCGACTCTCCAAAACCCTCGTCGACATGGACATGGCCGACTACAGTGCTGCACTGGACCCAGCCTACACCACCCT

GGAATTTGAGAATGTGCAGGTGTTGACGATGGGCAATGACACGTCCCCATCAGAAGGCACCAACCTCAACGCGCCCA

ACAGCCTGGGTGTCAGCGCCCTGTGTGCCATCTGCGGGGACCGGGCCACGGGCAAACACTACGGTGCCTCGAGCTGT

GACGGCTGCAAGGGCTTCTTCCGGAGGAGCGTGCGGAAGAACCACATGTACTCCTGCAGATTTAGCCGGCAGTGCGT

GGTGGACAAAGACAAGAGGAACCAGTGCCGCTACTGCAGGCTCAAGAAATGCTTCCGGGCTGGCATGAAGAAGGAAG

CCGTCCAGAATGAGCGGGACCGGATCAGCACTCGAAGGTCAAGCTATGAGGACAGCAGCCTGCCCTCCATCAATGCG

CTCCTGCAGGCGGAGGTCCTGTCCCGACAGATCACCTCCCCCGTCTCCGGGATCAACGGCGACATTCGGGCGAAGAA

GATTGCCAGCATCGCAGATGTGTGTGAGTCCATGAAGGAGCAGCTGCTGGTTCTCGTTGAGTGGGCCAAGTACATCC

CAGCTTTCTGCGAGCTCCCCCTGGACGACCAGGTGGCCCTGCTCAGAGCCCATGCTGGCGAGCACCTGCTGCTCGGA

GCCACCAAGAGATCCATGGTGTTCAAGGACGTGCTGCTCCTAGGCAATGACTACATTGTCCCTCGGCACTGCCCGGA

GCTGGCGGAGATGAGCCGGGTGTCCATACGCATCCTTGACGAGCTGGTGCTGCCCTTCCAGGAGCTGCAGATCGATG

ACAATGAGTATGCCTACCTCAAAGCCATCATCTTCTTTGACCCAGATGCCAAGGGGCTGAGCGATCCAGGGAAGATC

AAGCGGCTGCGTTCCCAGGTGCAGGTGAGCTTGGAGGACTACATCAACGACCGCCAGTATGACTCGCGTGGCCGCTT

TGGAGAGCTGCTGCTGCTGCTGCCCACCTTGCAGAGCATCACCTGGCAGATGATCGAGCAGATCCAGTTCATCAAGC

TCTTCGGCATGGCCAAGATTGACAACCTGTTGCAGGAGATGCTGCTGGGAGGTCCGTGCCAAGCCCAGGAGGGGCGG

GGTTGGAGTGGGGACTCCCCAGGAGACAGGCCTCACACAGTGAGCTCACCCCTCAGCTCCTTGGCTTCCCCACTGTG

CCGCTTTGGGCAAGTTGCT

HOXA1

ATGGACAACGCGCGGATGAATTCCTTCCTCGAGTACCCAATTTTGTCTAGTGGAGACAGTGGCACTTGCAGTGCCCG

AGCCTATCCATCAGACCACAGAATTACAACATTCCAAAGCTGTGCGGTGTCAGCCAACAGTTGCGGCGGAGACGACC

GCTTCCTGGTCGGAAGAGGGGTTCAAATTGGATCACCTCACCATCACCATCACCACCACCATCACCACCCCCAACCG

GCGACTTACCAAACCAGCGGCAATTTGGGCGTGAGCTATAGCCATTCCTCATGTGGACCTTCCTATGGGTCTCAGAA

TTTCTCCGCCCCTTATAGCCCATACGCCCTGAACCAAGAGGCCGATGTATCAGGAGGCTATCCCCAGTGCGCGCCAG

CGGTTTACTCAGGTAATCTTTCTAGCCCGATGGTCCAGCACCACCATCACCATCAAGGTTATGCCGGCGGTGCAGTC

GGATCCCCACAATACATACACCATAGTTACGGCCAAGAGCACCAATCCCTGGCCCTCGCTACATATAACAACTCACT

GTCTCCGCTTCATGCTTCCCACCAAGAAGCTTGTCGGAGTCCCGCCTCAGAAACTTCCTCTCCAGCTCAGACTTTTG

ATTGGATGAAGGTCAAGCGGAATCCGCCTAAAACGGGCAAAGTAGGTGAATATGGCTATTTGGGACAGCCTAATGCT

GTCCGCACCAATTTCACAACAAAACAGCTTACTGAACTCGAGAAGGAATTTCATTTTAATAAGTATTTGACTCGAGC

GAGACGAGTCGAAATCGCCGCTAGTCTTCAACTTAACGAGACCCAGGTTAAGATATGGTTCCAGAACAGAAGAATGA

AACAAAAAAAGCGGGAGAAGGAAGGACTCCTCCCTATATCACCAGCCACACCCCCAGGTAACGACGAGAAGGCGGAG

GAATCTTCAGAGAAGAGTTCCAGCTCCCCTTGTGTTCCTTCTCCTGGTAGCTCAACCAGCGATACCCTCACGACGAG

TCAC

HOXA10

ATGTGTCAAGGCAATTCCAAAGGTGAAAACGCAGCCAACTGGCTCACGGCAAAGAGTGGTCGGAAGAAGCGCTGCCC

CTACACGAAGCACCAGACACTGGAGCTGGAGAAGGAGTTTCTGTTCAATATGTACCTTACTCGAGAGCGGCGCCTAG

AGATTAGCCGCAGCGTCCACCTCACGGACAGACAAGTGAAAATCTGGTTTCAGAACCGCAGGATGAAACTGAAGAAA

ATGAATCGAGAAAACCGGATCCGGGAGCTCACAGCCAACTTTAATTTTTCC

HOXA11

ATGGATTTTGATGAGCGTGGTCCCTGCTCCTCTAACATGTATTTGCCAAGTTGTACTTACTACGTCTCGGGTCCAGA

TTTCTCCAGCCTCCCTTCTTTTCTGCCCCAGACCCCGTCTTCGCGCCCAATGACATACTCCTACTCCTCCAACCTGC

CCCAGGTCCAACCCGTGCGCGAAGTGACCTTCAGAGAGTACGCCATTGAGCCCGCCACTAAATGGCACCCCCGCGGC

AATCTGGCCCACTGCTACTCCGCGGAGGAGCTCGTGCACAGAGACTGCCTGCAGGCGCCCAGCGCGGCCGGCGTGCC

TGGCGACGTGCTGGCCAAGAGCTCGGCCAACGTCTACCACCACCCCACCCCCGCAGTCTCGTCCAATTTCTATAGCA

CCGTGGGCAGGAACGGCGTCCTGCCACAGGCTTTCGACCAGTTTTTCGAGACAGCCTACGGCACCCCGGAAAACCTC

GCCTCCTCCGACTACCCCGGGGACAAGAGCGCCGAGAAGGGGCCCCCGGCGGCCACGGCGACCTCCGCGGCGGCGGC

GGCGGCTGCAACGGGCGCGCCGGCAACTTCAAGTTCGGACAGCGGCGGCGGCGGCGGCTGCCGGGAGATGGCGGCGG

CAGCAGAGGAGAAAGAGCGGCGGCGGCGCCCCGAGAGCAGCAGCAGCCCCGAGTCGTCTTCCGGCCACACTGAGGAC

AAGGCCGGCGGCTCCAGTGGCCAACGCACCCGCAAAAAGCGCTGCCCCTATACCAAGTACCAGATCCGAGAGCTGGA

ACGGGAGTTCTTCTTCAGCGTCTACATTAACAAAGAGAAGCGCCTGCAACTGTCCCGCATGCTCAACCTCACTGATC

GTCAAGTCAAAATCTGGTTTCAGAACAGGAGAATGAAGGAAAAAAAAAATTAACAGAGACCGTTTACAGTACTACTCA

GCAAATCCACTCCTCTTG

HOXB6

ATGAGTTCCTATTTCGTGAACTCCACCTTCCCCGTCACTCTGGCCAGCGGGCAGGAGTCCTTCCTGGGCCAGCTACC

GCTCTATTCGTCGGGCTATGCGGACCCGCTGAGACATTACCCCGCGCCCTACGGGCCAGGGCCGGGCCAGGACAAGG

GCTTTGCCACTTCCTCCTATTACCCGCCGGCGGGCGGTGGCTACGGCCGAGCGGCGCCCTGCGACTACGGGCCGGCG

CCGGCCTTCTACCGCGAGAAAGAGTCGGCCTGCGCACTCTCCGGCGCCGACGAGCAGCCCCCGTTCCACCCCGAGCC

GCGGAAGTCGGACTGCGCGCAGGACAAGAGCGTGTTCGGCGAGACAGAAGAGCAGAAGTGCTCCACTCCGGTCTACC

CGTGGATGCAGCGGATGAATTCGTGCAACAGTTCCTCCTTTGGGCCCAGCGGCCGGCGAGGCCGCCAGACATACACA

CGTTACCAGACGCTGGAGCTGGAGAAGGAGTTTCACTACAATCGCTACCTGACGCGGCGGCGGCGCATCGAGATCGC

GCACGCCCTGTGCCTGACGGAGAGGCAGATCAAGATATGGTTCCAGAACCGACGCATGAAGTGGAAAAAGGAGAGCA

AACTGCTCAGCGCGTCTCAGCTCAGTGCCGAGGAGGAGGAAGAAAAACAGGCCGAG

KLF4

ATGGCTGTCAGCGACGCGCTGCTCCCATCTTTCTCCACGTTCGCGTCTGGCCCGGCGGGAAGGGAGAAGACACTGCG

TCAAGCAGGTGCCCCGAATAACCGCTGGCGGGAGGAGCTCTCCCACATGAAGCGACTTCCCCCAGTGCTTCCCGGCC

GCCCCTATGACCTGGCGGCGGCGACCGTGGCCACAGACCTGGAGAGCGGCGGAGCCGGTGCGGCTTGCGGCGGTAGC

AACCTGGCGCCCCTACCTCGGAGAGAGACCGAGGAGTTCAACGATCTCCTGGACCTGGACTTTATTCTCTCCAATTC

GCTGACCCATCCTCCGGAGTCAGTGGCCGCCACCGTGTCCTCGTCAGCGTCAGCCTCCTCTTCGTCGTCGCCGTCGA

GCAGCGGCCCTGCCAGCGCGCCCTCCACCTGCAGCTTCACCTATCCGATCCGGGCCGGGAACGACCCGGGCGTGGCG

CCGGGCGGCACGGGCGGAGGCCTCCTCTATGGCAGGGAGTCCGCTCCCCCTCCGACGGCTCCCTTCAACCTGGCGGA

CATCAACGACGTGAGCCCCTCGGGCGGCTTCGTGGCCGAGCTCCTGCGGCCAGAATTGGACCCGGTGTACATTCCGC

CGCAGCAGCCGCAGCCGCCAGGTGGCGGGCTGATGGGCAAGTTCGTGCTGAAGGCGTCGCTGAGCGCCCCTGGCAGC

GAGTACGGCAGCCCGTCGGTCATCAGCGTCAGCAAAGGCAGCCCTGACGGCAGCCACCCGGTGGTGGTGGCGCCCTA

CAACGGCGGGCCGCCGCGCACGTGCCCCAAGATCAAGCAGGAGGCGGTCTCTTCGTGCACCCACTTGGGCGCTGGAC

CCCCTCTCAGCAATGGCCACCGGCCGGCTGCACACGACTTCCCCCTGGGGCGGCAGCTCCCCAGCAGGACTACCCCG

ACCCTGGGTCTTGAGGAAGTGCTGAGCAGCAGGGACTGTCACCCTGCCCTGCCGCTTCCTCCCGGCTTCCATCCCCA

CCCGGGGCCCAATTACCCATCCTTCCTGCCCGATCAGATGCAGCCGCAAGTCCCGCCGCTCCATTACCAAGAGCTCA

TGCCACCCGGTTCCTGCATGCCAGAGGAGCCCAAGCCAAAGAGGGGAAGACGATCGTGGCCCCGGAAAAGGACCGCC

ACCCACACTTGTGATTACGCGGGCTGCGGCAAAACCTACACAAAGAGTTCCCATCTCAAGGCACACCTGCGAACCCA

CACAGGTGAGAAACCTTACCACTGTGACTGGGACGGCTGTGGATGGAAATTCGCCCGCTCAGATGAACTGACCAGGC

ACTACCGTAAACACACGGGGCACCGCCCGTTCCAGTGCCAAAAATGCGACCGAGCATTTTCCAGGTCGGACCACCTC
GCCTTACACATGAAGAGGCATTTT

LHX3

ATGGAGGCGCGCGGGGAGCTGGGCCCGGCCCGGGAGTCGGCGGGAGGCGACCTGCTGCTAGCACTGCTGGCGCGGAG
GGCGGACCTGCGCCGAGAGATCCCGCTGTGCGCTGGCTGTGACCAGCACATCCTGGACCGCTTCATCCTCAAGGCTC
TGGACCGCCACTGGCACAGCAAGTGTCTCAAGTGCAGCGACTGCCACACGCCACTGGCCGAGCGCTGCTTCAGCCGA
GGGGAGAGCGTTTACTGCAAGGACGACTTTTTCAAGCGCTTCGGGACCAAGTGCGCCGCGTGCCAGCTGGGCATCCC
GCCCACGCAGGTGGTGCGCCGCGCCCAGGACTTCGTGTACCACCTGCACTGCTTTGCCTGCGTCGTGTGCAAGCGGC
AGCTGGCCACGGGCGACGAGTTCTACCTCATGGAGGACAGCCGGCTCGTGTGCAAGGCGGACTACGAAACCGCCAAG
CAGCGAGAGGCCGAGGCCACGGCCAAGCGGCCGCGCACGACCATCACCGCCAAGCAGCTGGAGACGCTGAAGAGCGC
TTACAACACCTCGCCCAAGCCGGCGCGCCACGTGCGCGAGCAGCTCTCGTCCGAGACGGGCCTGGACATGCGCGTGG
TGCAGGTTTGGTTCCAGAACCGCCGGGCCAAGGAGAAGAGGCTGAAGAAGGACGCCGGCCGGCAGCGCTGGGGGCAG
TATTTCCGCAACATGAAGCGCTCCCGCGGCGGCTCCAAGTCGGACAAGGACAGCGTTCAGGAGGGGCAGGACAGCGA
CGCTGAGGTCTCCTTCCCCGATGAGCCTTCCTTGGCGGAAATGGGCCCGGCCAATGGCCTCTACGGGAGCTTGGGGG
AACCCACCCAGGCCTTGGGCCGGCCCTCGGGAGCCCTGGGCAACTTCTCCCTGGAGCATGGAGGCCTGGCAGGCCCA
GAGCAGTACCGAGAGCTGCGTCCCGGCAGCCCCTACGGTGTCCCCCCATCCCCCGCCGCCCCGCAGAGCCTCCCTGG
CCCCCAGCCCCTCCTCTCCAGCCTGGTGTACCCAGACACCAGCTTGGGCCTTGTGCCCTCGGGAGCCCCCGGCGGGC
CCCCACCCATGAGGGTGCTGGCAGGGAACGGACCCAGTTCTGACCTATCCACGGGGAGCAGCGGGGGTTACCCCGAC
TTCCCTGCCAGCCCCGCCTCCTGGCTGGATGAGGTAGACCACGCTCAGTTCTCAGGCCTCATGGGCCCAGCTTTCTT
GTAC

LMX1A

ATGGAAGGAATCATGAACCCCTACACGGCTCTGCCCACCCCACAGCAGCTCCTGGCCATCGAGCAGAGTGTCTACAG
CTCAGATCCCTTCCGACAGGGTCTCACCCCACCCCAGATGCCTGGAGACCACATGCACCCTTATGGTGCCGAGCCCC
TTTTCCATGACCTGGATAGCGACGACACCTCCCTCAGTAACCTGGGTGACTGTTTCCTAGCAACCTCAGAAGCTGGG
CCTCTGCAGTCCAGAGTGGGAAACCCCATTGACCATCTGTACTCCATGCAGAATTCTTACTTCACATCT

MEF2C

ATGGGGAGAAAAAAGATTCAGATTACGAGGATTATGGATGAACGTAACAGACAGGTGACATTTACAAAGAGGAAATT

TGGGTTGATGAAGAAGGCTTATGAGCTGAGCGTGCTGTGTGACTGTGAGATTGCGCTGATCATCTTCAACAGCACCA

ACAAGCTGTTCCAGTATGCCAGCACCGACATGGACAAAGTGCTTCTCAAGTACACGGAGTACAACGAGCCGCATGAG

AGCCGGACAAACTCAGACATCGTGGAGACGTTGAGAAAGAAGGGCCTTAATGGCTGTGACAGCCCAGACCCCGATGC

GGACGATTCCGTAGGTCACAGCCCTGAGTCTGAGGACAAGTACAGGAAAATTAACGAAGATATTGATCTAATGATCA

GCAGGCAAAGATTGTGTGCTGTTCCACCTCCCAACTTCGAGATGCCAGTCTCCATCCCAGTGTCCAGCCACAACAGT

TTGGTGTACAGCAACCCTGTCAGCTCACTGGGAAACCCCAACCTATTGCCACTGGCTCACCCTTCTCTGCAGAGGAA

TAGTATGTCTCCTGGTGTAACACATCGACCTCCAAGTGCAGGTAACACAGGTGGTCTGATGGGTGGAGACCTCACGT

CTGGTGCAGGCACCAGTGCAGGGAACGGGTATGGCAATCCCCGAAACTCACCAGGTCTGCTGGTCTCACCTGGTAAC

TTGAACAAGAATATGCAAGCAAAATCTCCTCCCCCAATGAATTTAGGAATGAATAACCGTAAACCAGATCTCCGAGT

TCTTATTCCACCAGGCAGCAAGAATACGATGCCATCAGTGTCTGAGGATGTCGACCTGCTTTTGAATCAAAGGATAA

ATAACTCCCAGTCGGCTCAGTCATTGGCTACCCCAGTGGTTTCCGTAGCAACTCCTACTTTACCAGGACAAGGAATG

GGAGGATATCCATCAGCCATTTCAACAACATATGGTACCGAGTACTCTCTGAGTAGTGCAGACCTGTCATCTCTGTC

TGGGTTTAACACCGCCAGCGCTCTTCACCTTGGTTCAGTAACTGGCTGGCAACAGCAACACCTACATAACATGCCAC

CATCTGCCCTCAGTCAGTTGGGAGCTTGCACTAGCACTCATTTATCTCAGAGTTCAAATCTCTCCCTGCCTTCTACT

CAAAGCCTCAACATCAAGTCAGAACCTGTTTCTCCTCCTAGAGACCGTACCACCACCCCTTCGAGATACCCACAACA

CACGCGCCACGAGGCGGGGAGATCTCCTGTTGACAGCTTGAGCAGCTGTAGCAGTTCGTACGACGGGAGCGACCGAG

AGGATCACCGGAACGAATTCCACTCCCCCATTGGACTCACCAGACCTTCGCCGGACGAAAGGGAAAGTCCCTCAGTC

AAGCGCATGCGACTTTCTGAAGGATGGGCAACA

MESP1

ATGGCCCAGCCCCTGTGCCCGCCGCTCTCCGAGTCCTGGATGCTCTCTGCGGCCTGGGGCCCAACTCGGCGGCCGCC

GCCCTCCGACAAGGACTGCGGCCGCTCCCTCGTCTCGTCCCCAGACTCATGGGGCAGCACCCCAGCCGACAGCCCCG

TGGCGAGCCCCGCGCGGCCAGGCACCCTCCGGGACCCCCGCGCCCCCTCCGTAGGTAGGCGCGGCGCGCGCAGCAGC

CGCCTGGGCAGCGGGCAGAGGCAGAGCGCCAGTGAGCGGGAGAAACTGCGCATGCGCACGCTGGCCCGCGCCCTGCA

CGAGCTGCGCCGCTTTCTACCGCCGTCCGTGGCGCCCGCGGGCCAGAGCCTGACCAAGATCGAGACGCTGCGCCTGG

CTATCCGCTATATCGGCCACCTGTCGGCCGTGCTAGGCCTCAGCGAGGAGAGTCTCCAGCGCCGGTGCCGGCAGCGC

GGTGACGCGGGGTCCCCTCGGGGCTGCCCGCTGTGCCCCGACGACTGCCCCGCGCAGATGCAGACACGGACGCAGGC

TGAGGGGCAGGGGCAGGGGCGCGGGCTGGGCCTGGTATCCGCCGTCCGCGCCGGGGCGTCCTGGGGATCCCCGCCTG

CCTGCCCCGGAGCCCGAGCTGCACCCGAGCCGCGCGACCCGCCTGCGCTGTTCGCCGAGGCGGCGTGCCCGGAAGGG

CAGGCGATGGAGCCAAGCCCACCGTCCCCGCTCCTTCCGGGCGACGTGCTGGCTCTGTTGGAGACCTGGATGCCCCT

CTCGCCTCTGGAGTGGCTGCCTGAGGAGCCCAAGTTG

MITF

ATGCTGGAAATGCTAGAATATAATCACTATCAGGTGCAGACCCACCTCGAAAACCCCACCAAGTACCACATACAGCA

AGCCCAACGGCAGCAGGTAAAGCAGTACCTTTCTACCACTTTAGCAAATAAACATGCCAACCAAGTCCTGAGCTTGC

CATGTCCAAACCAGCCTGGCGATCATGTCATGCCACCGGTGCCGGGGAGCAGCGCACCCAACAGCCCCATGGCTATG

CTTACGCTTAACTCCAACTGTGAAAAAGAGGGATTTTATAAGTTTGAAGAGCAAAACAGGGCAGAGAGCGAGTGCCC

AGGCATGAACACACATTCACGAGCGTCCTGTATGCAGATGGATGATGTAATCGATGACATCATTAGCCTAGAATCAA

GTTATAATGAGGAAATCTTGGGCTTGATGGATCCTGCTTTGCAAATGGCAAATACGTTGCCTGTCTCGGGAAACTTG

ATTGATCTTTATGGAAACCAAGGTCTGCCCCCACCAGGCCTCACCATCAGCAACTCCTGTCCAGCCAACCTTCCCAA

CATAAAAAGGGAGCTCACAGAGTCTGAAGCAAGAGCACTGGCCAAAGAGAGGCAGAAAAAGGACAATCACAACCTGA

TTGAACGAAGAAGAAGATTTAACATAAATGACCGCATTAAAGAACTAGGTACTTTGATTCCCAAGTCAAATGATCCA

GACATGCGCTGGAACAAGGGAACCATCTTAAAAGCATCCGTGGACTATATCCGAAAGTTGCAACGAGAACAGCAACG

CGCAAAAGAACTTGAAAACCGACAGAAGAAACTGGAGCACGCCAACCGGCATTTGTTGCTCAGAATACAGGAACTTG

AAATGCAGGCTCGAGCTCATGGACTTTCCCTTATTCCATCCACGGGTCTCTGCTCTCCAGATTTGGTGAATCGGATC

ATCAAGCAAGAACCCGTTCTTGAGAACTGCAGCCAAGACCTCCTTCAGCATCATGCAGACCTAACCTGTACAACAAC

TCTCGATCTCACGGATGGCACCATCACCTTCAACAACAACCTCGGAACTGGGACTGAGGCCAACCAAGCCTATAGTG

TCCCCACAAAAATGGGATCCAAACTGGAAGACATCCTGATGGACGACACCCTTTCTCCCGTCGGTGTCACTGATCCA

CTCCTTTCCTCAGTGTCCCCCGGAGCTTCCAAAACAAGCAGCCGGAGGAGCAGTATGAGCATGGAAGAGACGGAGCA

CACTTGT

MYC

ATGCCCCTCAACGTTAGCTTCACCAACAGGAACTATGACCTCGACTACGACTCGGTGCAGCCGTATTTCTACTGCGA

CGAGGAGGAGAACTTCTACCAGCAGCAGCAGCAGAGCGAGCTGCAGCCCCCGGCGCCCAGCGAGGATATCTGGAAGA

AATTCGAGCTGCTGCCCACCCCGCCCCTGTCCCCTAGCCGCCGCTCCGGGCTCTGCTCGCCCTCCTACGTTGCGGTC

ACACCCTTCTCCCTTCGGGGAGACAACGACGGCGGTGGCGGGAGCTTCTCCACGGCCGACCAGCTGGAGATGGTGAC

CGAGCTGCTGGGAGGAGACATGGTGAACCAGAGTTTCATCTGCGACCCGGACGACGAGACCTTCATCAAAAACATCA

TCATCCAGGACTGTATGTGGAGCGGCTTCTCGGCCGCCGCCAAGCTCGTCTCAGAGAAGCTGGCCTCCTACCAGGCT

GCGCGCAAAGACAGCGGCAGCCCGAACCCCGCCCGCGGCCACAGCGTCTGCTCCACCTCCAGCTTGTACCTGCAGGA

TCTGAGCGCCGCCGCCTCAGAGTGCATCGACCCCTCGGTGGTCTTCCCCTACCCTCTCAACGACAGCAGCTCGCCCA

AGTCCTGCGCCTCGCAAGACTCCAGCGCCTTCTCTCCGTCCTCGGATTCTCTGCTCTCCTCGACGGAGTCCTCCCCG

CAGGGCAGCCCCGAGCCCCTGGTGCTCCATGAGGAGACACCGCCCACCACCAGCAGCGACTCTGAGGAGGAACAAGA

AGATGAGGAAGAAATCGATGTTGTTTCTGTGGAAAAGAGGCAGGCTCCTGGCAAAAGGTCAGAGTCTGGATCACCTT

CTGCTGGAGGCCACAGCAAACCTCCTCACAGCCCACTGGTCCTCAAGAGGTGCCACGTCTCCACACATCAGCACAAC

TACGCAGCGCCTCCCTCCACTCGGAAGGACTATCCTGCTGCCAAGAGGGTCAAGTTGGACAGTGTCAGAGTCCTGAG

ACAGATCAGCAACAACCGAAAATGCACCAGCCCCAGGTCCTCGGACACCGAGGAGAATGTCAAGAGGCGAACACACA

ACGTCTTGGAGCGCCAGAGGAGGAACGAGCTAAAACGGAGCTTTTTTGCCCTGCGTGACCAGATCCCGGAGTTGGAA

AACAATGAAAAGGCCCCCAAGGTAGTTATCCTTAAAAAAGCCACAGCATACATCCTGTCCGTCCAAGCAGAGGAGCA

AAAGCTCATTTCTGAAGAGGACTTGTTGCGGAAACGACGAGAACAGTTGAAACACAAACTTGAACAGCTACGGAACT

CTTGTGCG

MYCL

ATGGACTACGACTCGTACCAGCACTATTTCTACGACTATGACTGCGGGGAGGATTTCTACCGCTCCACGGCGCCCAG

CGAGGACATCTGGAAGAAATTCGAGCTGGTGCCATCGCCCCCCACGTCGCCGCCCTGGGGCTTGGGTCCCGGCGCAG

GGGACCCGGCCCCCGGGATTGGTCCCCCGGAGCCGTGGCCCGGAGGGTGCACCGGAGACGAAGCGGAATCCCGGGGC

CACTCGAAAGGCTGGGGCAGGAACTACGCCTCCATCATACGCCGTGACTGCATGTGGAGCGGCTTCTCGGCCCGGGA

ACGGCTGGAGAGAGCTGTGAGCGACCGGCTCGCTCCTGGCGCGCCCCGGGGGAACCCGCCCAAGGCGTCCGCCGCCC

CGGACTGCACTCCCAGCCTCGAAGCCGGCAACCCGGCGCCCGCCGCCCCCTGTCCGCTGGGCGAACCCAAGACCCAG

GCCTGCTCCGGGTCCGAGAGCCCAAGCGACTCGGGTAAGGACCTCCCCGAGCCATCCAAGAGGGGGCCACCCCATGG

GTGGCCAAAGCTCTGCCCCTGCCTGAGGTCAGGCATTGGCTCTTCTCAAGCTCTTGGGCCATCTCCGCCTCTCTTTG

GC

MYCN

ATGCCGAGTTGTTCCACGTCTACGATGCCAGGAATGATATGCAAGAACCCCGACTTGGAGTTTGACTCTTTGCAACC

ATGCTTTTATCCGGATGAAGACGACTTTTATTTCGGCGGCCCGGACAGCACCCCTCCTGGAGAGGACATCTGGAAAA

AATTCGAACTTTTGCCTACACCCCCACTCAGTCCCTCTCGAGGATTTGCGGAACACAGCAGTGAACCGCCGTCTTGG

GTGACAGAGATGCTCCTCGAGAACGAATTGTGGGGAAGCCCTGCGGAGGAAGACGCTTTCGGGCTCGGTGGACTCGG

AGGTCTCACGCCGAACCCAGTCATACTGCAGGATTGCATGTGGTCTGGATTCTCAGCTCGGGAGAAGCTGGAACGGG

CAGTTTCTGAGAAACTCCAACATGGCCGGGGCCCTCCAACAGCGGGTTCTACCGCACAGTCCCCTGGTGCTGGAGCC

GCTAGTCCCGCGGGGAGAGGCCATGGGGGCGCGGCAGGAGCGGGTAGGGCCGGCGCTGCGTTGCCTGCTGAGCTTGC

GCACCCCGCCGCTGAATGTGTAGATCCCGCGGTAGTGTTTCCGTTCCCCGTTAATAAGCGAGAACCGGCACCGGTGC

CAGCCGCTCCTGCGTCTGCACCCGCGGCAGGTCCTGCTGTCGCCTCAGGAGCAGGTATTGCCGCTCCTGCAGGGGCA

CCAGGAGTAGCCCCTCCAAGGCCCGGCGGTAGGCAAACCTCCGGCGGCGACCACAAAGCACTCTCAACGAGCGGAGA

GGATACACTGTCCGATAGTGATGACGAGGACGACGAAGAGGAGGACGAGGAGGAGGAGATAGATGTTGTCACGGTCG

AGAAGCGAAGGAGTTCTTCAAATACAAAAGCGGTAACGACATTCACGATAACAGTAAGACCTAAGAACGCAGCCCTC

GGTCCAGGGCGGGCCCAGTCCAGTGAGCTTATACTTAAGCGCTGCCTGCCGATTCACCAGCAGCATAACTACGCGGC

CCCTAGTCCCTACGTTGAGAGCGAGGATGCCCCCCCACAAAAAAAAAATAAAGTCTGAAGCGTCCCCCCGCCCCCTGA

AATCCGTAATCCCCCCAAAGGCGAAGTCACTCAGTCCCAGGAATTCAGATTCCGAGGACTCCGAACGGCGGCGGAAT

CATAACATACTTGAGAGACAACGACGCAATGACCTGAGGTCTTCTTTTTTGACCCTCCGAGATCACGTCCCCGAGCT

GGTTAAGAATGAGAAAGCTGCGAAGGTAGTCATACTGAAAAAGGCCACCGAGTATGTCCATAGTTTGCAAGCTGAGG

AGCACCAGCTTCTCCTTGAAAAGGAGAAACTTCAGGCACGACAACAGCAATTGCTGAAAAAGATTGAGCATGCACGC

ACTTGT

MYOD1

ATGGAGCTACTGTCGCCACCGCTCCGCGACGTAGACCTGACGGCCCCCGACGGCTCTCTCTGCTCCTTTGCCACAAC

GGACGACTTCTATGACGACCCGTGTTTCGACTCCCCGGACCTGCGCTTCTTCGAAGACCTGGACCCGCGCCTGATGC

ACGTGGGCGCGCTCCTGAAACCCGAAGAGCACTCGCACTTCCCCGCGGCGGTGCACCCGGCCCCGGGCGCACGTGAG

GACGAGCATGTGCGCGCGCCCAGCGGGCACCACCAGGCGGGCCGCTGCCTACTGTGGGCCTGCAAGGCGTGCAAGCG

CAAGACCACCAACGCCGACCGCCGCAAGGCCGCCACCATGCGCGAGCGGCGCCGCCTGAGCAAAGTAAATGAGGCCT

TTGAGACACTCAAGCGCTGCACGTCGAGCAATCCAAACCAGCGGTTGCCCAAGGTGGAGATCCTGCGCAACGCCATC

CGCTATATCGAGGGCCTGCAGGCTCTGCTGCGCGACCAGGACGCCGCGCCCCCTGGCGCCGCAGCCGCCTTCTATGC

GCCGGGCCCGCTGCCCCCGGGCCGCGGCGGCGAGCACTACAGCGGCGACTCCGACGCGTCCAGCCCGCGCTCCAACT

GCTCCGACGGCATGATGGACTACAGCGGCCCCCCGAGCGGCGCCCGGCGGCGGAACTGCTACGAAGGCGCCTACTAC

AACGAGGCGCCCAGCGAACCCAGGCCCGGGAAGAGTGCGGCGGTGTCGAGCCTAGACTGCCTGTCCAGCATCGTGGA

GCGCATCTCCACCGAGAGCCCTGCGGCGCCCGCCCTCCTGCTGGCGGACGTGCCTTCTGAGTCGCCTCCGCGCAGGC

AAGAGGCTGCCGCCCCCAGCGAGGGAGAGAGCAGCGGCGACCCCACCCAGTCACCGGACGCCGCCCCGCAGTGCCCT

GCGGGTGCGAACCCCAACCCGATATACCAGGTGCTC

MYOG

ATGGAGCTGTATGAGACATCCCCCTACTTCTACCAGGAACCCCGCTTCTATGATGGGGAAAACTACCTGCCTGTCCA

CCTCCAGGGCTTCGAACCACCAGGCTACGAGCGGACGGAGCTCACCCTGAGCCCCGAGGCCCCAGGGCCCCTTGAGG

ACAAGGGGCTGGGGACCCCCGAGCACTGTCCAGGCCAGTGCCTGCCGTGGGCGTGTAAGGTGTGTAAGAGGAAGTCG

GTGTCCGTGGACCGGCGGCGGGCGGCCACACTGAGGGAGAAGCGCAGGCTCAAGAAGGTGAATGAGGCCTTCGAGGC

CCTGAAGAGAAGCACCCTGCTCAACCCCAACCAGCGGCTGCCCAAGGTGGAGATCCTGCGCAGTGCCATCCAGTACA

TCGAGCGCCTCCAGGCCCTGCTCAGCTCCCTCAACCAGGAGGAGCGTGACCTCCGCTACCGGGGCGGGGGCGGGCCC

CAGCCAGGGGTGCCCAGCGAATGCAGCTCTCACAGCGCCTCCTGCAGTCCAGAGTGGGGCAGTGCACTGGAGTTCAG

CGCCAACCCAGGGGATCATCTGCTCACGGCTGACCCTACAGATGCCCACAACCTGCACTCCCTCACCTCCATCGTGG

ACAGCATCACAGTGGAAGATGTGTCTGTGGCCTTCCCAGATGAAACCATGCCCAAC

NEUROD1

ATGACCAAATCGTACAGCGAGAGTGGGCTGATGGGCGAGCCTCAGCCCCAAGGTCCTCCAAGCTGGACAGACGAGTG

TCTCAGTTCTCAGGACGAGGAGCACGAGGCAGACAAGAAGGAGGACGACCTCGAAGCCATGAACGCAGAGGAGGACT

CACTGAGGAACGGGGGAGAGGAGGAGGACGAAGATGAGGACCTGGAAGAGGAGGAAGAAGAGGAAGAGGAGGATGAC

GATCAAAAGCCCAAGAGACGCGGCCCCAAAAAGAAGAAGATGACTAAGGCTCGCCTGGAGCGTTTTAAATTGAGACG

CATGAAGGCTAACGCCCGGGAGCGGAACCGCATGCACGGACTGAACGCGGCGCTAGACAACCTGCGCAAGGTGGTGC

CTTGCTATTCTAAGACGCAGAAGCTGTCCAAAATCGAGACTCTGCGCTTGGCCAAGAACTACATCTGGGCTCTGTCG

GAGATCCTGCGCTCAGGCAAAAGCCCAGACCTGGTCTCCTTCGTTCAGACGCTTTGCAAGGGCTTATCCCAACCCAC

CACCAACCTGGTTGCGGGCTGCCTGCAACTCAATCCTCGGACTTTTCTGCCTGAGCAGAACCAGGACATGCCCCCCC

ACCTGCCGACGGCCAGCGCTTCCTTCCCTGTACACCCCTACTCCTACCAGTCGCCTGGGCTGCCCAGTCCGCCTTAC

GGTACCATGGACAGCTCCCATGTCTTCCACGTTAAGCCTCCGCCGCACGCCTACAGCGCAGCGCTGGAGCCCTTCTT

TGAAAGCCCTCTGACTGATTGCACCAGCCCTTCCTTTGATGGACCCCTCAGCCCGCCGCTCAGCATCAATGGCAACT

TCTCTTTCAAACACGAACCGTCCGCCGAGTTTGAGAAAAATTATGCCTTTACCATGCACTATCCTGCAGCGACACTG

GCAGGGGCCCAAAGCCACGGATCAATCTTCTCAGGCACCGCTGCCCCTCGCTGCGAGATCCCCATAGACAATATTAT

GTCCTTCGATAGCCATTCACATCATGAGCGAGTCATGAGTGCCCAGCTCAATGCCATATTTCATGAT

NEUROG1

ATGCCAGCCCGCCTTGAGACCTGCATCTCCGACCTCGACTGCGCCAGCAGCAGCGGCAGTGACCTATCCGGCTTCCT

CACCGACGAGGAAGACTGTGCCAGACTCCAACAGGCAGCCTCCGCTTCGGGGCCGCCCGCGCCGGCCCGCAGGGGCG

CGCCCAATATCTCCCGGGCGTCTGAGGTTCCAGGGGCACAGGACGACGAGCAGGAGAGGCGGCGGCGCCGCGGCCGG

ACGCGGGTCCGCTCCGAGGCGCTGCTGCACTCGCTGCGCAGGAGCCGGCGCGTCAAGGCCAACGATCGCGAGCGCAA

CCGCATGCACAACTTGAACGCGGCCCTGGACGCACTGCGCAGCGTGCTGCCCTCGTTCCCCGACGACACCAAGCTCA

CCAAAATCGAGACGCTGCGCTTCGCCTACAACTACATCTGGGCTCTGGCCGAGACACTGCGCCTGGCGGATCAAGGG

CTGCCCGGAGGCGGTGCCCGGGAGCGCCTCCTGCCGCCGCAGTGCGTCCCCTGCCTGCCCGGTCCCCCAAGCCCCGC

CAGCGACGCGGAGTCCTGGGGCTCAGGTGCCGCCGCCGCCTCCCCGCTCTCTGACCCCAGTAGCCCAGCCGCCTCCG

AAGACTTCACCTACCGCCCCGGCGACCCTGTTTTCTCCTTCCCAAGCCTGCCCAAAGACTTGCTCCACACAACGCCC

TGTTTCATTCCTTACCAC

NEUROG3

ATGACACCACAACCATCTGGTGCTCCCACAGTCCAGGTGACGCGAGAGACTGAAAGATCATTCCCACGCGCGTCCGA

GGATGAGGTGACATGTCCAACTAGCGCACCCCCCTCTCCTACCCGGACCCGCGGGAATTGTGCTGAGGCCGAAGAGG

GAGGATGCAGAGGAGCACCAAGGAAACTTCGAGCCCGACGGGGTGGAAGAAGCCGCCCCAAGTCTGAGCTCGCCCTT

AGCAAGCAGCGCCGCAGTCGGAGGAAAAAGGCAAACGACCGGGAAAGGAATAGGATGCATAATCTTAATTCTGCTCT

GGACGCTCTGCGAGGCGTACTTCCTACTTTCCCGGATGACGCGAAATTGACCAAGATAGAGACTCTCCGGTTTGCAC

ATAATTACATCTGGGCTCTTACACAAACACTGAGAATTGCCGATCACAGTCTTTACGCTCTTGAGCCACCCGCCCCG

CACTGTGGCGAGCTGGGTAGCCCCGGCGGCTCTCCTGGAGACTGGGGGTCTTTGTATTCTCCTGTCAGCCAAGCGGG

ATCTTTGAGTCCGGCTGCCAGTCTCGAAGAAAGACCCGGACTCCTTGGAGCGACTTTTTCAGCATGTCTGTCCCCTG

GCTCATTGGCTTTCTCAGACTTTTTG

NRL

ATGGCCCTGCCTCCCAGCCCGCTGGCCATGGAATATGTCAATGACTTTGACTTGATGAAGTTTGAGGTAAAGCGGGA

ACCCTCTGAGGGCCGACCTGGCCCACCTACAGCCTCACTGGGATCCACACCTTACAGCTCAGTGCCTCCTTCACCCA

CCTTCAGTGAACCAGGCATGGTAGGGGCAACCGAGGGTACACGACCAGGTTTGGAGGAGCTGTACTGGCTTGCTACC

CTGCAGCAGCAGCTTGGGGCTGGGGAGGCATTGGGACTGAGTCCTGAAGAGGCCATGGAGCTACTGCAAGGTCAGGG

CCCAGTCCCTGTTGATGGACCCCATGGTTACTACCCAGGGAGCCCAGAGGAGACAGGAGCCCAGCACGTTCAGTTGG

CAGAGCGGTTTTCCGACGCGGCGCTTGTCTCGATGTCTGTGCGAGAACTAAACCGGCAGCTGCGGGGATGCGGGAGA

GACGAGGCTCTACGACTGAAGCAGAGGCGTCGAACGCTGAAGAACCGTGGCTATGCGCAAGCATGTCGTTCCAAGAG

GCTGCAACAGAGGCGAGGTCTTGAGGCCGAGCGCGCCCGTCTTGCAGCCCAGCTAGATGCGCTACGAGCTGAAGTAG

CACGTTTGGCAAGAGAGCGAGATCTCTACAAGGCTCGCTGTGACCGGCTAACCTCGAGTGGCCCCGGGTCCGGGGAT

CCCTCCCACCTTTTCCTCTGCCCAACTTTCTTGTACAAAGTTGTCCCC

ONECUT1

ATGAACGCGCAGCTGACCATGGAAGCGATCGGCGAGCTGCACGGGGTGAGCCATGAGCCGGTGCCCGCCCCTGCCGA

CCTGCTGGGCGGCAGCCCCCACGCGCGCAGCTCCGTGGCGCACCGCGGCAGCCACCTGCCCCCCGCGCACCCGCGCT

CCATGGGCATGGCGTCCCTGCTGGACGGCGGCAGCGGCGGCGGAGATTACCACCACCACCACCGGGCCCCTGAGCAC

AGCCTGGCCGGCCCCCTGCATCCCACCATGACCATGGCCTGCGAGACTCCCCCAGGTATGAGCATGCCCACCACCTA

CACCACCTTGACCCCTCTGCAGCCGCTGCCTCCCATCTCCACAGTCTCGGACAAGTTCCCCCACCATCACCACCACC

ACCATCACCACCACCACCCGCACCACCACCAGCGCCTGGCGGGCAACGTGAGCGGTAGCTTCACGCTCATGCGGGAT

GAGCGCGGGCTGGCCTCCATGAATAACCTCTATACCCCCTACCACAAGGACGTGGCCGGCATGGGCCAGAGCCTCTC

GCCCCTCTCCAGCTCCGGTCTGGGCAGCATCCACAACTCCCAGCAAGGGCTCCCCCACTATGCCCACCCGGGGGCCG

CCATGCCCACCGACAAGATGCTCACCCCCAACGGCTTCGAAGCCCACCACCCGGCCATGCTCGGCCGCCACGGGGAG

CAGCACCTCACGCCCACCTCGGCCGGCATGGTGCCCATCAACGGCCTTCCTCCGCACCATCCCCACGCCCACCTGAA

CGCCCAGGGCCACGGGCAACTCCTGGGCACAGCCCGGGAGCCCAACCCTTCGGTGACCGGCGCGCAGGTCAGCAATG

GAAGTAATTCAGGGCAGATGGAAGAGATCAATACCAAAGAGGTGGCGCAGCGTATCACCACCGAGCTCAAGCGCTAC

AGCATCCCACAGGCCATCTTCGCGCAGAGGGTGCTCTGCCGCTCCCAGGGGACCCTCTCGGACCTGCTGCGCAACCC

CAAACCCTGGAGCAAACTCAAATCCGGCCGGGAGACCTTCCGGAGGATGTGGAAGTGGCTGCAGGAGCCGGAGTTCC

AGCGCATGTCCGCGCTCCGCTTAGCAGCATGCAAAAGGAAAGAACAAGAACATGGGAAGGATAGAGGCAACACACCC

AAAAAGCCCAGGTTGGTCTTCACAGATGTCCAGCGTCGAACTCTACATGCAATATTCAAGGAAAATAAGCGTCCATC
CAAAGAATTGCAAATCACCATTTCCCAGCAGCTGGGGTTGGAGCTGAGCACTGTCAGCAACTTCTTCATGAACGCAA
GAAGGAGGAGTCTGGACAAGTGGCAGGACGAGGGCAGCTCCAATTCAGGCAACTCATCTTCTTCATCAAGCACTTGT
ACCAAAGCA

OTX2

ATGATGTCTTATCTTAAGCAACCGCCTTACGCAGTCAATGGGCTGAGTCTGACCACTTCGGGTATGGACTTGCTGCA
CCCCTCCGTGGGCTACCCGGGGCCCTGGGCTTCTTGTCCCGCAGCCACCCCCCGGAAACAGCGCCGGGAGAGGACGA
CGTTCACTCGGGCGCAGCTAGATGTGCTGGAAGCACTGTTTGCCAAGACCCGGTACCCAGACATCTTCATGCGAGAG
GAGGTGGCACTGAAAATCAACTTGCCCGAGTCGAGGGTGCAGGTATGGTTTAAGAATCGAAGAGCTAAGTGCCGCCA
ACAACAGCAACAACAGCAGAATGGAGGTCAAAACAAAGTGAGACCTGCCAAAAAGAAGACATCTCCAGCTCGGGAAG
TGAGTTCAGAGAGTGGAACAAGTGGCCAATTCACTCCCCCCTCTAGCACCTCAGTCCCGACCATTGCCAGCAGCAGT
GCTCCTGTGTCTATCTGGAGCCCAGCTTCCATCTCCCCACTGTCAGATCCCTTGTCCACCTCCTCTTCCTGCATGCA
GAGGTCCTATCCCATGACCTATACTCAGGCTTCAGGTTATAGTCAAGGATATGCTGGCTCAACTTCCTACTTTGGGG
GCATGGACTGTGGATCATATTTGACCCCTATGCATCACCAGCTTCCCGGACCAGGGGCCACACTCAGTCCCATGGGT
ACCAATGCAGTCACCAGCCATCTCAATCAGTCCCCAGCTTCTCTTTCCACCCAGGGATATGGAGCTTCAAGCTTGGG
TTTTAACTCAACCACTGATTGCTTGGATTATAAGGACCAAACTGCCTCCTGGAAGCTTAACTTCAATGCTGACTGCT
TGGATTATAAAGATCAGACATCCTCGTGGAAATTCCAGGTTTTG

PAX7

ATGGCGGCCCTTCCCGGCACGGTACCGAGAATGATGCGGCCGGCTCCGGGGCAGAACTACCCCCGCACGGGATTCCC
TTTGGAAGTGTCCACCCCGCTTGGCCAAGGCCGGGTCAATCAGCTGGGAGGGGTCTTCATCAATGGGCGACCCCTGC
CTAACCACATCCGCCACAAGATAGTGGAGATGGCCCACCATGGCATCCGGCCCTGTGTCATCTCCCGACAGCTGCGT
GTCTCCCACGGCTGCGTCTCCAAGATTCTTTGCCGCTACCAGGAGACCGGGTCCATCCGGCCTGGGGCCATCGGCGG
CAGCAAGCCCAGACAGGTGGCGACTCCGGATGTAGAGAAAAAGATTGAGGAGTACAAGAGGGAAAACCCAGGCATGT
TCAGCTGGGAGATCCGGGACAGGCTGCTGAAGGATGGGCACTGTGACCGAAGCACTGTGCCCTCAGTGAGTTCGATT
AGCCGCGTGCTCAGAATCAAGTTCGGGAAGAAAGAGGAGGAGGATGAAGCGGACAAGAAGGAGGACGACGGCGAAAA
GAAGGCCAAACACAGCATCGACGGCATCCTGGGCGACAAAGGGAACCGGCTGGACGAGGGCTCGGATGTGGAGTCGG

AACCTGACCTCCCACTGAAGCGCAAGCAGCGACGCAGTCGGACCACATTCACGGCCGAGCAGCTGGAGGAGCTGGAG

AAGGCCTTTGAGAGGACCCACTACCCAGACATATACACCCGCGAGGAGCTGGCGCAGAGGACCAAGCTGACAGAGGC

GCGTGTGCAGGTCTGGTTCAGTAACCGCCGCGCCCGTTGGCGTAAGCAGGCAGGAGCCAACCAGCTGGCGGCGTTCA

ACCACCTTCTGCCAGGAGGCTTCCCGCCCACCGGCATGCCCACGCTGCCCCCCTACCAGCTGCCGGACTCCACCTAC

CCCACCACCACCATCTCCCAAGATGGGGGCAGCACTGTGCACCGGCCTCAGCCCCTGCCACCGTCCACCATGCACCA

GGGCGGGCTGGCTGCAGCGGCTGCAGCCGCCGACACCAGCTCTGCCTACGGAGCCCGCCACAGCTTCTCCAGCTACT

CTGACAGCTTCATGAATCCGGCGGCGCCCTCCAACCACATGAACCCGGTCAGCAACGGCCTGTCTCCTCAGGTGATG

AGCATCTTGGGCAACCCCAGTGCGGTGCCCCCGCAGCCACAGGCTGACTTCTCCATCTCCCCGCTGCATGGCGGCCT

GGACTCGGCCACCTCCATCTCAGCCAGCTGCAGCCAGCGGGCCGACTCCATCAAGCCAGGAGACAGCCTGCCCACCT

CCCAGGCCTACTGCCCACCCACCTACAGCACCACCGGCTACAGCGTGGACCCCGTGGCCGGCTATCAGTACGGCCAG

TACGGCCAGAGTGAGTGCCTGGTGCCCTGGGCGTCCCCCGTCCCCATTCCTTCTCCCACCCCCAGGGCCTCCTGCTT

GTTTATGGAGAGCTACAAGGTGGTGTCAGGGTGGGGAATGTCCATTTCACAGATGGAAAAATTGAAGTCCAGCCAGA

TGGAACAGTTCACC

POU1F1

ATGAGTTGCCAAGCTTTTACTTCGGCTGATACCTTTATACCTCTGAATTCTGACGCCTCTGCAACTCTGCCTCTGAT

AATGCATCACAGTGCTGCCGAGTGTCTACCAGTCTCCAACCATGCCACCAATGTGATGTCTACAGCAACAGGACTTC

ATTATTCTGTTCCTTCCTGTCATTATGGAAACCAGCCATCAACCTATGGAGTGATGGCAGGTAGTTTAACCCCTTGT

CTTTATAAATTTCCTGACCACACCTTGAGTCATGGATTTCCTCCTATACACCAGCCTCTTCTGGCAGAGGACCCCAC

AGCTGCTGATTTCAAGCAGGAACTCAGGCGGAAAAGTAAATTGGTGGAAGAGCCAATAGACATGGATTCTCCAGAAA

TCAGAGAACTTGAAAAGTTTGCCAATGAATTTAAAGTGAGACGAATTAAATTAGGATACACCCAGACAAATGTTGGG

GAGGCCCTGGCAGCTGTGCATGGCTCTGAATTCAGTCAAACAACAATCTGCCGATTTGAAAATCTGCAGCTCAGCTT

TAAAAATGCATGCAAACTGAAAGCAATATTATCCAAATGGCTGGAGGAAGCTGAGCAAGTAGGAGCTTTGTACAATG

AAAAAGTGGGAGCAAATGAAAGGAAAAGAAAACGAAGAACAACTATAAGCATTGCTGCTAAAGATGCTCTGGAGAGA

CACTTTGGAGAACAGAATAAACCTTCTTCTCAAGAGATCATGAGGATGGCTGAAGAACTGAATCTGGAGAAAGAAGT

AGTAAGAGTTTGGTTTTGCAACCGGAGGCAGAGAGAAAAACGGGTGAAAACAAGTCTGAATCAGAGTTTATTTTCTA

TTTCTAAGGAACATCTTGAGTGCAGATCAGGCCTCATGGGCCCAGCTTTCTTGTAC

POU5F1

ATGGCGGGACACCTGGCTTCAGATTTTGCCTTCTCGCCCCCTCCAGGTGGTGGAGGTGATGGGCCAGGGGGGCCGGA

GCCGGGCTGGGTTGATCCTCGGACCTGGCTAAGCTTCCAAGGCCCTCCTGGAGGGCCAGGAATCGGGCCGGGGGTTG

GGCCAGGCTCTGAGGTGTGGGGGATTCCCCCATGCCCCCCGCCGTATGAGTTCTGTGGGGGGATGGCGTACTGTGGG

CCCCAGGTTGGAGTGGGGCTAGTGCCCCAAGGCGGCTTGGAGACCTCTCAGCCTGAGGGCGAAGCAGGAGTCGGGGT

GGAGAGCAACTCCGATGGGGCCTCCCCGGAGCCCTGCACCGTCACCCCTGGTGCCGTGAAGCTGGAGAAGGAGAAGC

TGGAGCAAAACCCGGAGGAGTCCCAGGACATCAAAGCTCTGCAGAAAGAACTCGAGCAATTTGCCAAGCTCCTGAAG

CAGAAGAGGATCACCCTGGGATATACACAGGCCGATGTGGGGCTCACCCTGGGGGTTCTATTTGGGAAGGTATTCAG

CCAAACGACCATCTGCCGCTTTGAGGCTCTGCAGCTTAGCTTCAAGAACATGTGTAAGCTGCGGCCCTTGCTGCAGA

AGTGGGTGGAGGAAGCTGACAACAATGAAAATCTTCAGGAGATATGCAAAGCAGAAACCCTCGTGCAGGCCCGAAAG

AGAAAGCGAACCAGTATCGAGAACCGAGTGAGAGGCAACCTGGAGAATTTGTTCCTGCAGTGCCCGAAACCCACACT

GCAGCAGATCAGCCACATCGCCCAGCAGCTTGGGCTCGAGAAGGATGTGGTCCGAGTGTGGTTCTGTAACCGGCGCC

AGAAGGGCAAGCGATCAAGCAGCGACTATGCACAACGAGAGGATTTTGAGGCTGCTGGGTCTCCTTTCTCAGGGGGA

CCAGTGTCCTTTCCTCTGGCCCCAGGGCCCCATTTTGGTACCCCAGGCTATGGGAGCCCTCACTTCACTGCACTGTA

CTCCTCGGTCCCTTTCCCTGAGGGGGAAGCCTTTCCCCCTGTCTCTGTCACCACTCTGGGCTCTCCCATGCATTCAA

AC

RUNX1

ATGGCTTCAGACAGCATATTTGAGTCATTTCCTTCGTACCCACAGTGCTTCATGAGAGAATGCATACTTGGAATGAA

TCCTTCTAGAGACGTCCACGATGCCAGCACGAGCCGCCGCTTCACGCCGCCTTCCACCGCGCTGAGCCCAGGCAAGA

TGAGCGAGGCGTTGCCGCTGGGCGCCCCGGACGCCGGCGCTGCCCTGGCCGGCAAGCTGAGGAGCGGCGACCGCAGC

ATGGTGGAGGTGCTGGCCGACCACCCGGGCGAGCTGGTGCGCACCGACAGCCCCAACTTCCTCTGCTCCGTGCTGCC

TACGCACTGGCGCTGCAACAAGACCCTGCCCATCGCTTTCAAGGTGGTGGCCCTAGGGGATGTTCCAGATGGCACTC

TGGTCACTGTGATGGCTGGCAATGATGAAAACTACTCGGCTGAGCTGAGAAATGCTACCGCAGCCATGAAGAACCAG

GTTGCAAGATTTAATGACCTCAGGTTTGTCGGTCGAAGTGGAAGAGGGAAAAGCTTCACTCTGACCATCACTGTCTT

CACAAACCCACCGCAAGTCGCCACCTACCACAGAGCCATCAAAATCACAGTGGATGGGCCCCGAGAACCTCGAAGAC

ATCGGCAGAAACTAGATGATCAGACCAAGCCCGGGAGCTTGTCCTTTTCCGAGCGGCTCAGTGAACTGGAGCAGCTG

CGGCGCACAGCCATGAGGGTCAGCCCACACCACCCAGCCCCCACGCCCAACCCTCGTGCCTCCCTGAACCACTCCAC

TGCCTTTAACCCTCAGCCTCAGAGTCAGATGCAGGATACAAGGCAGATCCAACCATCCCCACCGTGGTCCTACGATC

AGTCCTACCAATACCTGGGATCCATTGCCTCTCCTTCTGTGCACCCAGCAACGCCCATTTCACCTGGACGTGCCAGC

GGCATGACAACCCTCTCTGCAGAACTTTCCAGTCGACTCTCAACGGCACCCGACCTGACAGCGTTCAGCGACCCGCG

CCAGTTCCCCGCGCTGCCCTCCATCTCCGACCCCCGCATGCACTATCCAGGCGCCTTCACCTACTCCCCGACGCCGG

TCACCTCGGGCATCGGCATCGGCATGTCGGCCATGGGCTCGGCCACGCGCTACCACACCTACCTGCCGCCGCCCTAC

CCCGGCTCGTCGCAAGCGCAGGGAGGCCCGTTCCAAGCCAGCTCGCCCTCCTACCACCTGTACTACGGCGCCTCGGC

CGGCTCCTACCAGTTCTCCATGGTGGGCGGCGAGCGCTCGCCGCCGCGCATCCTGCCGCCCTGCACCAACGCCTCCA

CCGGCTCCGCGCTGCTCAACCCCAGCCTCCCGAACCAGAGCGACGTGGTGGAGGCCGAGGGCAGCCACAGCAACTCC

CCCACCAACATGGCGCCCTCCGCGCGCCTGGAGGAGGCCGTGTGGAGGCCCTAC

SIX1

ATGTCGATGCTGCCGTCGTTTGGCTTTACGCAGGAGCAAGTGGCGTGCGTGTGCGAGGTTCTGCAGCAAGGCGGAAA

CCTGGAGCGCCTGGGCAGGTTCCTGTGGTCACTGCCCGCCTGCGACCACCTGCACAAGAACGAGAGCGTACTCAAGG

CCAAGGCGGTGGTCGCCTTCCACCGCGGCAACTTCCGTGAGCTCTACAAGATCCTGGAGAGCCACCAGTTCTCGCCT

CACAACCACCCCAAACTGCAGCAACTGTGGCTGAAGGCGCATTACGTGGAGGCCGAGAAGCTGTGCGGCCGACCCCT

GGGCGCCGTGGGCAAATATCGGGTGCGCCGAAAATTTCCACTGCCGCGCACCATCTGGGACGGCGAGGAGACCAGCT

ACTGCTTCAAGGAGAAGTCGAGGGGTGTCCTGCGGGAGTGGTACGCGCACAATCCCTACCCATCGCCGCGTGAGAAG

CGGGAGCTGGCCGAGGCCACCGGCCTCACCACCACCCAGGTCAGCAACTGGTTTAAGAACCGGAGGCAAAGAGACCG

GGCCGCGGAGGCCAAGGAAAGGGAGAACACCGAAAACAATAACTCCTCCTCCAACAAGCAGAACCAACTCTCTCCTC

TGGAAGGGGGCAAGCCGCTCATGTCCAGCTCAGAAGAGGAATTCTCACCTCCCCAAAGTCCAGACCAGAACTCGGTC

CTTCTGCTGCAGGGCAATATGGGCCACGCCAGGAGCTCAAACTATTCTCTCCCGGGCTTAACAGCCTCGCAGCCCAG

TCACGGCCTGCAGACCCACCAGCATCAGCTCCAAGACTCTCTGCTCGGCCCCCTCACCTCCAGTCTGGTGGACTTGG

GGTCC

SIX2

ATGTCCATGCTGCCCACCTTCGGCTTCACGCAGGAGCAAGTGGCGTGCGTGTGCGAGGTGCTGCAGCAGGGCGGCAA

CATCGAGCGGCTGGGCCGCTTCCTGTGGTCGCTGCCCGCCTGCGAGCACCTTCACAAGAATGAAAGCGTGCTCAAGG

CCAAGGCCGTGGTGGCCTTCCACCGCGGCAACTTCCGCGAGCTCTACAAGATCCTGGAGAGCCACCAGTTCTCGCCG

CACAACCACGCCAAGCTGCAGCAGCTGTGGCTCAAGGCACACTACATCGAGGCGGAGAAGCTGCGCGGCCGACCCCT

GGGCGCCGTGGGCAAATACCGCGTGCGCCGCAAATTCCCGCTGCCGCGCTCCATCTGGGACGGCGAGGAGACCAGCT

ACTGCTTCAAGGAAAAGAGTCGCAGCGTGCTGCGCGAGTGGTACGCGCACAACCCCTACCCTTCACCCCGCGAGAAG

CGTGAGCTGACGGAGGCCACGGGCCTCACCACCACACAGGTCAGCAACTGGTTCAAGAACCGGCGGCAGCGCGACCG

GGCGGCCGAGGCCAAGGAAAGGGAGAACAACGAGAACTCCAATTCTAACAGCCACAACCCGCTGAATGGCAGCGGCA

AGTCGGTGTTAGGCAGCTCGGAGGATGAGAAGACTCCATCGGGGACGCCAGACCACTCATCATCCAGCCCCGCACTG

CTCCTCAGCCCGCCGCCCCCTGGGCTGCCGTCCCTGCACAGCCTGGGCCACCCTCCGGGCCCCAGCGCAGTGCCAGT

GCCGGTGCCAGGCGGAGGTGGAGCGGACCCACTGCAACACCACCATGGCCTGCAGGACTCCATCCTCAACCCCATGT

CAGCCAACCTCGTGGACCTGGGCTCC

SNAI2

ATGCCGCGCTCCTTCCTGGTCAAGAAGCATTTCAACGCCTCCAAAAAGCCAAACTACAGCGAACTGGACACACATAC

AGTGATTATTTCCCCGTATCTCTATGAGAGTTACTCCATGCCTGTCATACCACAACCAGAGATCCTCAGCTCAGGAG

CATACAGCCCCATCACTGTGTGGACTACCGCTGCTCCATTCCACGCCCAGCTACCCAATGGCCTCTCTCCTCTTTCC

GGATACTCCTCATCTTTGGGGCGAGTGAGTCCCCCTCCTCCATCTGACACCTCCTCCAAGGACCACAGTGGCTCAGA

AAGCCCCATTAGTGATGAAGAGGAAAGACTACAGTCCAAGCTTTCAGACCCCCATGCCATTGAAGCTGAAAAGTTTC

AGTGCAATTTATGCAATAAGACCTATTCAACTTTTTCTGGGCTGGCCAAACATAAGCAGCTGCACTGCGATGCCCAG

TCTAGAAAATCTTTCAGCTGTAAATACTGTGACAAGGAATATGTGAGCCTGGGCGCCCTGAAGATGCATATTCGGAC

CCACACATTACCTTGTGTTTGCAAGATCTGCGGCAAGGCGTTTTCCAGACCCTGGTTGCTTCAAGGACACATTAGAA

CTCACACGGGGGAGAAGCCTTTTTCTTGCCCTCACTGCAACAGAGCATTTGCAGACAGGTCAAATCTGAGGGCTCAT

CTGCAGACCCATTCTGATGTAAAGAAATACCAGTGCAAAAACTGCTCCAAAACCTTCTCCAGAATGTCTCTCCTGCA

CAAACATGAGGAATCTGGCTGCTGTGTAGCACAC

SOX10

ATGGCGGAGGAGCAGGACCTATCGGAGGTGGAGCTGAGCCCCGTGGGCTCGGAGGAGCCCCGCTGCCTGTCCCCGGG

GAGCGCGCCCTCGCTAGGGCCCGACGGCGGCGGCGGCGGATCGGGCCTGCGAGCCAGCCCGGGGCCAGGCGAGCTGG

GCAAGGTCAAGAAGGAGCAGCAGGACGGCGAGGCGGACGATGACAAGTTCCCCGTGTGCATCCGCGAGGCCGTCAGC

CAGGTGCTCAGCGGCTACGACTGGACGCTGGTGCCCATGCCCGTGCGCGTCAACGGCGCCAGCAAAAGCAAGCCGCA

CGTCAAGCGGCCCATGAACGCCTTCATGGTGTGGGCTCAGGCAGCGCGCAGGAAGCTCGCGGACCAGTACCCGCACC

TGCACAACGCTGAGCTCAGCAAGACGCTGGGCAAGCTCTGGAGGCTGCTGAACGAAAGTGACAAGCGCCCCTTCATC

GAGGAGGCTGAGCGGCTCCGTATGCAGCACAAGAAAGACCACCCGGACTACAAGTACCAGCCCAGGCGGCGGAAGAA

CGGGAAGGCCGCCCAGGGCGAGGCGGAGTGCCCCGGTGGGGAGGCCGAGCAAGGTGGGACCGCCGCCATCCAGGCCC

ACTACAAGAGCGCCCACTTGGACCACCGGCACCCAGGAGAGGGCTCCCCCATGTCAGATGGGAACCCCGAGCACCCC

TCAGGCCAGAGCCATGGCCCACCCACCCCTCCAACCACCCCGAAGACAGAGCTGCAGTCGGGCAAGGCAGACCCGAA

GCGGGACGGGCGCTCCATGGGGGAGGGCGGGAAGCCTCACATCGACTTCGGCAACGTGGACATTGGTGAGATCAGCC

ACGAGGTAATGTCCAACATGGAGACCTTTGATGTGGCTGAGTTGGACCAGTACCTGCCGCCCAATGGGCACCCAGGC

CATGTGAGCAGCTACTCAGCAGCCGGCTATGGGCTGGGCAGTGCCCTGGCCGTGGCCAGTGGACACTCCGCCTGGAT

CTCCAAGCCACCAGGCGTGGCTCTGCCCACGGTCTCACCACCTGGTGTGGATGCCAAAGCCCAGGTGAAGACAGAGA

CCGCGGGGCCCCAGGGGCCCCCACACTACACCGACCAGCCATCCACCTCACAGATCGCCTACACCTCCCTCAGCCTG

CCCCACTATGGCTCAGCCTTCCCCTCCATCTCCCGCCCCCAGTTTGACTACTCTGACCATCAGCCCTCAGGACCCTA

TTATGGCCACTCGGGCCAGGCCTCTGGCCTCTACTCGGCCTTCTCCTATATGGGGCCCTCGCAGCGGCCCCTCTACA

CGGCCATCTCTGACCCCAGCCCCTCAGGGCCCCAGTCCCACAGCCCCACACACTGGGAGCAGCCAGTATATACGACA

CTGTCCCGGCCC

SOX2

ATGTACAACATGATGGAGACGGAGCTGAAGCCGCCGGGCCCGCAGCAAACTTCGGGGGGCGGCGGCGGCAACTCCAC

CGCGGCGGCGGCCGGCGGCAACCAGAAAAACAGCCCGGACCGCGTCAAGCGGCCCATGAATGCCTTCATGGTGTGGT

CCCGCGGGCAGCGGCGCAAGATGGCCCAGGAGAACCCCAAGATGCACAACTCGGAGATCAGCAAGCGCCTGGGCGCC

GAGTGGAAACTTTTGTCGGAGACGGAGAAGCGGCCGTTCATCGACGAGGCTAAGCGGCTGCGAGCGCTGCACATGAA

GGAGCACCCGGATTATAAATACCGGCCCCGGCGGAAAACCAAGACGCTCATGAAGAAGGATAAGTACACGCTGCCCG

GCGGGCTGCTGGCCCCCGGCGGCAATAGCATGGCGAGCGGGGTCGGGGTGGGCGCCGGCCTGGGCGCGGGCGTGAAC

CAGCGCATGGACAGTTACGCGCACATGAACGGCTGGAGCAACGGCAGCTACAGCATGATGCAGGACCAGCTGGGCTA

CCCGCAGCACCCGGGCCTCAATGCGCACGGCGCAGCGCAGATGCAGCCCATGCACCGCTACGACGTGAGCGCCCTGC

AGTACAACTCCATGACCAGCTCGCAGACCTACATGAACGGCTCGCCCACCTACAGCATGTCCTACTCGCAGCAGGGC

ACCCCTGGCATGGCTCTTGGCTCCATGGGTTCGGTGGTCAAGTCCGAGGCCAGCTCCAGCCCCCCTGTGGTTACCTC

TTCCTCCCACTCCAGGGCGCCCTGCCAGGCCGGGGACCTCCGGGACATGATCAGCATGTATCTCCCCGGCGCCGAGG

TGCCGGAACCCGCCGCCCCCAGCAGACTTCACATGTCCCAGCACTACCAGAGCGGCCCGGTGCCCGGCACGGCCATT
AACGGCACACTGCCCCTCTCACACATG

SOX3

ATGCGACCTGTTCGAGAGAACTCATCAGGTGCGAGAAGCCCGCGGGTTCCTGCTGATTTGGCGCGGAGCATTTTGAT
AAGCCTACCCTTCCCGCCGGACTCGCTGGCCCACAGGCCCCCAAGCTCCGCTCCGACGGAGTCCCAGGGCCTTTTCA
CCGTGGCCGCTCCAGCCCCGGGAGCGCCTTCTCCTCCCGCCACGCTGGCGCACCTTCTTCCCGCCCCGGCAATGTAC
AGCCTTCTGGAGACTGAACTCAAGAACCCCGTAGGGACACCCACACAAGCGGCGGGCACCGGCGGCCCCGCAGCCCC
GGGAGGCGCAGGCAAGAGTAGTGCGAACGCAGCCGGCGGCGCGAACTCGGGCGGCGGCAGCAGCGGTGGTGCGAGCG
GAGGTGGCGGGGGTACAGACCAGGACCGTGTGAAACGGCCCATGAACGCCTTCATGGTATGGTCCCGCGGGCAGCGG
CGCAAAATGGCCCTGGAGAACCCCAAGATGCACAATTCTGAGATCAGCAAGCGCTTGGGCGCCGACTGGAAACTGCT
GACCGACGCCGAGAAGCGACCATTCATCGACGAGGCCAAGCGACTTCGCGCCGTGCACATGAAGGAGTATCCGGACT
ACAAGTACCGACCGCGCCGCAAGACCAAGACGCTGCTCAAGAAAGATAAGTACTCCCTGCCCAGCGGCCTCCTGCCT
CCCGGTGCCGCGGCCGCCGCCGCCGCTGCCGCGGCCGCAGCCGCTGCCGCCAGCAGTCCGGTGGGCGTGGGCCAGCG
CCTGGACACGTACACGCACGTGAACGGCTGGGCCAACGGCGCGTACTCGCTGGTGCAGGAGCAGCTGGGCTACGCGC
AGCCCCCGAGCATGAGCAGCCCGCCGCCGCCGCCCGCGCTGCCGCCGATGCACCGCTACGACATGGCCGGCCTGCAG
TACAGCCCAATGATGCCGCCCGGCGCTCAGAGCTACATGAACGTCGCTGCCGCGGCCGCCGCCGCCTCGGGCTACGG
GGGCATGGCGCCCTCAGCCACAGCAGCCGCGGCCGCCGCCTACGGGCAGCAGCCCGCCACCGCCGCGGCCGCAGCTG
CGGCCGCAGCCGCCATGAGCCTGGGCCCCATGGGCTCGGTAGTGAAGTCTGAGCCCAGCTCGCCGCCGCCCGCCATC
GCATCGCACTCTCAGCGCGCGTGCCTCGGCGACCTGCGCGACATGATCAGCATGTACCTGCCACCCGGCGGGGACGC
GGCCGACGCCGCCTCTCCGCTGCCCGGCGGTCGCCTGCACGGCGTGCACCAGCACTACCAGGGCGCCGGGACTGCAG
TCAACGGAACGGTGCCGCTGACCCACATC

SPI1

ATGTTACAGGCGTGCAAAATGGAAGGGTTTCCCCTCGTCCCCCCTCAGCCATCAGAAGACCTGGTGCCCTATGACAC
GGATCTATACCAACGCCAAACGCACGAGTATTACCCCTATCTCAGCAGTGATGGGGAGAGCCATAGCGACCATTACT
GGGACTTCCACCCCCACCACGTGCACAGCGAGTTCGAGAGCTTCGCCGAGAACAACTTCACGGAGCTCCAGAGCGTG
CAGCCCCCGCAGCTGCAGCAGCTCTACCGCCACATGGAGCTGGAGCAGATGCACGTCCTCGATACCCCCATGGTGCC

ACCCCATCCCAGTCTTGGCCACCAGGTCTCCTACCTGCCCCGGATGTGCCTCCAGTACCCATCCCTGTCCCCAGCCC

AGCCCAGCTCAGATGAGGAGGAGGGCGAGCGGCAGAGCCCCCCACTGGAGGTGTCTGACGGCGAGGCGGATGGCCTG

GAGCCCGGGCCTGGGCTCCTGCCTGGGGAGACAGGCAGCAAGAAGAAGATCCGCCTGTACCAGTTCCTGTTGGACCT

GCTCCGCAGCGGCGACATGAAGGACAGCATCTGGTGGGTGGACAAGGACAAGGGCACCTTCCAGTTCTCGTCCAAGC

ACAAGGAGGCGCTGGCGCACCGCTGGGGCATCCAGAAGGGCAACCGCAAGAAGATGACCTACCAGAAGATGGCGCGC

GCGCTGCGCAACTACGGCAAGACGGGCGAGGTCAAGAAGGTGAAGAAGAAGCTCACCTACCAGTTCAGCGGCGAAGT

GCTGGGCCGCGGGGGCCTGGCCGAGCGGCGCCACCCGCCCCAC

SPIB

ATGCTCGCCCTGGAGGCTGCACAGCTCGACGGGCCACACTTCAGCTGTCTGTACCCAGATGGCGTCTTCTATGACCT

GGACAGCTGCAAGCATTCCAGCTACCCTGATTCAGAGGGGGCTCCTGACTCCCTGTGGGACTGGACTGTGGCCCCAC

CTGTCCCAGCCACCCCCTATGAAGCCTTCGACCCGGCAGCAGCCGCTTTTAGCCACCCCCAGGCTGCCCAGCTCTGC

TACGAACCCCCCACCTACAGCCCTGCAGGGAACCTCGAACTGGCCCCCAGCCTGGAGGCCCCGGGGCCTGGCCTCCC

CGCATACCCCACGGAGAACTTCGCTAGCCAGACCCTGGTTCCCCCGGCATATGCCCCGTACCCCAGCCCTGTGCTAT

CAGAGGAGGAAGACTTACCGTTGGACAGCCCTGCCCTGGAGGTCTCGGACAGCGAGTCGGATGAGGCCCTCGTGGCT

GGCCCCGAGGGGAAGGGATCCGAGGCAGGGACTCGCAAGAAGCTGCGCCTGTACCAGTTCCTGCTGGGGCTACTGAC

GCGCGGGGACATGCGTGAGTGCGTGTGGTGGGTGGAGCCAGGCGCCGGCGTCTTCCAGTTCTCCTCCAAGCACAAGG

AACTCCTGGCGCGCCGCTGGGGCCAGCAGAAGGGGAACCGCAAGCGCATGACCTACCAGAAGCTGGCGCGCGCCCTC

CGAAACTACGCCAAGACCGGCGAGATCCGCAAGGTCAAGCGCAAGCTCACCTACCAGTTCGACAGCGCGCTGCTGCC

TGCAGTCCGCCGGGCCTTG

SPIC

ATGACGTGTGTTGAACAAGACAAGCTGGGTCAAGCATTTGAAGATGCTTTTGAGGTTCTGAGGCAACATTCAACTGG

AGATCTTCAGTACTCGCCAGATTACAGAAATTACCTGGCTTTAATCAACCATCGTCCTCATGTCAAAGGAAATTCCA

GCTGCTATGGAGTGTTGCCTACAGAGGAGCCTGTCTATAATTGGAGAACGGTAATTAACAGTGCTGCGGACTTCTAT

TTTGAAGGAAATATTCATCAATCTCTGCAGAACATAACTGAAAACCAGCTGGTACAACCCACTCTTCTCCAGCAAAA

GGGGGGAAAAGGCAGGAAGAAGCTCCGACTGTTTGAATACCTTCACGAATCCCTGTATAATCCGGAGATGGCATCTT

GTATTCAGTGGGTAGATAAAACCAAAGGCATCTTTCAGTTTGTATCAAAAAACAAAGAAAAACTTGCCGAGCTTTGG

GGGAAAAGAAAAGGCAACAGGAAGACCATGACTTACCAGAAAATGGCCAGGGCACTCAGAAATTACGGAAGAAGTGG

GGAAATTACCAAAATCCGGAGGAAGCTGACTTACCAGTTCAGTGAGGCCATTCTCCAAAGACTCTCTCCATCCTATT

TCCTGGGGAAAGAGATCTTCTATTCACAGTGTGTTCAACCTGATCAAGAATATCTCAGTTTAAATAACTGGAATGCA

AATTATAATTATACATATGCCAATTACCATGAGCTAAATCACCATGATTGC

SRY

ATGCAATCATATGCTTCTGCTATGTTAAGCGTATTCAACAGCGATGATTACAGTCCAGCTGTGCAAGAGAATATTCC

CGCTCTCCGGAGAAGCTCTTCCTTCCTTTGCACTGAAAGCTGTAACTCTAAGTATCAGTGTGAAACGGGAGAAAACA

GTAAAGGCAACGTCCAGGATAGAGTGAAGCGACCCATGAACGCATTCATCGTGTGGTCTCGCGATCAGAGGCGCAAG

ATGGCTCTAGAGAATCCCAGAATGCGAAACTCAGAGATCAGCAAGCAGCTGGGATACCAGTGGAAAATGCTTACTGA

AGCCGAAAAATGGCCATTCTTCCAGGAGGCACAGAAATTACAGGCCATGCACAGAGAGAAATACCCGAATTATAAGT

ATCGACCTCGTCGGAAGGCGAAGATGCTGCCGAAGAATTGCAGTTTGCTTCCCGCAGATCCCGCTTCGGTACTCTGC

AGCGAAGTGCAACTGGACAACAGGTTGTACAGGGATGACTGTACGAAAGCCACACACTCAAGAATGGAGCACCAGCT

AGGCCACTTACCGCCCATCAACGCAGCCAGCTCACCGCAGCAACGGGACCGCTACAGCCACTGGACAAAGCTG

TBX5

ATGGCCGACGCAGACGAGGGCTTTGGCCTGGCGCACACGCCTCTGGAGCCTGACGCAAAAGACCTGCCCTGCGATTC

GAAACCCGAGAGCGCGCTCGGGGCCCCCAGCAAGTCCCCGTCGTCCCCGCAGGCCGCCTTCACCCAGCAGGGCATGG

AGGGAATCAAAGTGTTTCTCCATGAAAGAGAACTGTGGCTAAAATTCCACGAAGTGGGCACGGAAATGATCATAACC

AAGGCTGGAAGGCGGATGTTTCCCAGTTACAAAGTGAAGGTGACGGGCCTTAATCCCAAAACGAAGTACATTCTTCT

CATGGACATTGTACCTGCCGACGATCACAGATACAAATTCGCAGATAATAAATGGTCTGTGACGGGCAAAGCTGAGC

CCGCCATGCCTGGCCGCCTGTACGTGCACCCAGACTCCCCCGCCACCGGGGCGCATTGGATGAGGCAGCTCGTCTCC

TTCCAGAAACTCAAGCTCACCAACAACCACCTGGACCCATTTGGGCATATTATTCTAAATTCCATGCACAAATACCA

GCCTAGATTACACATCGTGAAAGCGGATGAAAATAATGGATTTGGCTCAAAAAATACAGCGTTCTGCACTCACGTCT

TTCCTGAGACTGCGTTTATAGCAGTGACTTCCTACCAGAACCACAAGATCACGCAATTAAAGATTGAGAATAATCCC

TTTGCCAAAGGATTTCGGGGCAGTGATGACATGGAGCTGCACAGAATGTCAAGAATGCAAAGTAAAGAATATCCCGT

GGTCCCCAGGAGCACCGTGAGGCAAAAAGTGGCCTCCAACCACAGTCCTTTCAGCAGCGAGTCTCGAGCTCTCTCCA

CCTCATCCAATTTGGGGTCCCAATACCAGTGTGAGAATGGTGTTTCCGGCCCCTCCCAGGACCTCCTGCCTCCACCC

AACCCATACCCACTGCCCCAGGAGCATAGCCAAATTTACCATTGTACCAAGAGGAAAGAGGAAGAATGTTCCACCAC

AGACCATCCCTATAAGAAGCCCTACATGGAGACATCACCCAGTGAAGAAGATTCCTTCTACCGCTCTAGCTATCCAC

AGCAGCAGGGCCTGGGTGCCTCCTACAGGACAGAGTCGGCACAGCGGCAAGCTTGCATGTATGCCAGCTCTGCGCCC

CCCAGCGAGCCTGTGCCCAGCCTAGAGGACATCAGCTGCAACACGTGGCCAAGCATGCCTTCCTACAGCAGCTGCAC

CGTCACCACCGTGCAGCCCATGGACAGGCTACCCTACCAGCACTTCTCCGCTCACTTCACCTCGGGGCCCCTGGTCC

CTCGGCTGGCTGGCATGGCCAACCATGGCTCCCCACAGCTGGGAGAGGGAATGTTCCAGCACCAGACCTCCGTGGCC

CACCAGCCTGTGGTCAGGCAGTGTGGGCCTCAGACTGGCCTGCAGTCCCCTGGCACCCTTCAGCCCCCTGAGTTCCT

CTACTCTCATGGCGTGCCAAGGACTCTATCCCCTCATCAGTACCACTCTGTGCACGGAGTTGGCATGGTGCCAGAGT

GGAGCGACAATAGCTTG

TFAP2C

ATGTTGTGGAAAATAACCGATAATGTCAAGTACGAAGAGGACTGCGAGGATCGCCACGACGGGAGCAGCAATGGGAA

TCCGCGGGTCCCCCACCTCTCCTCCGCCGGGCAGCACCTCTACAGCCCCGCGCCACCCCTCTCCCACACTGGAGTCG

CCGAATATCAGCCGCCACCCTACTTTCCCCCTCCCTACCAGCAGCTGGCCTACTCCCAGTCGGCCGACCCCTACTCG

CATCTGGGGGAAGCGTACGCCGCCGCCATCAACCCCCTGCACCAGCCGGCGCCCACAGGCAGCCAGCAGCAGGCCTG

GCCCGGCCGCCAGAGCCAGGAGGGAGCGGGGCTGCCCTCGCACCACGGGCGCCCGGCCGGCCTACTGCCCCACCTCT

CCGGGCTGGAGGCGGGCGCGGTGAGCGCCCGCAGGGATGCCTACCGCCGCTCCGACCTGCTGCTGCCCCACGCACAC

GCCCTGGATGCCGCGGGCCTGGCCGAGAACCTGGGGCTCCACGACATGCCTCACCAGATGGACGAGGTGCAGAATGT

CGACGACCAGCACCTGTTGCTGCACGATCAGACAGTCATTCGCAAAGGTCCCATTTCCATGACCAAGAACCCTCTGA

ACCTCCCCTGTCAGAAGGAGCTGGTGGGGGCCGTAATGAACCCCACTGAGGTCTTCTGCTCAGTCCCTGGAAGATTG

TCGCTCCTCAGCTCTACGTCTAAATACAAAGTGACAGTGGCTGAAGTACAGAGGCGACTGTCCCCACCTGAATGCTT

AAATGCCTCGTTACTGGGAGGTGTTCTCAGAAGAGCCAAATCGAAAAATGGAGGCCGGTCCTTGCGGGAGAAGTTGG

ACAAGATTGGGTTGAATCTTCCGGCCGGGAGGCGGAAAGCCGCTCATGTGACTCTCCTGACATCCTTAGTAGAAGGT

GAAGCTGTTCATTTGGCTAGGGACTTTGCCTATGTCTGTGAAGCCGAATTTCCTAGTAAACCAGTGGCAGAATATTT

AACCAGACCTCATCTTGGAGGACGAAATGAGATGGCAGCTAGGAAGAACATGCTATTGGCGGCCCAGCAACTGTGTA

AAGAATTCACAGAACTTCTCAGCCAAGACCGGACACCCCATGGGACCAGCAGGCTCGCCCCAGTCTTGGAGACGAAC

ATACAGAACTGCTTGTCTCATTTCAGCCTGATTACCCACGGGTTTGGCAGCCAGGCCATCTGTGCCGCGGTGTCTGC

CCTGCAGAACTACATCAAAGAAGCCCTGATTGTCATAGACAAATCCTACATGAACCCTGGAGACCAGAGTCCAGCTG
ATTCTAACAAAACCCTGGAGAAAATGGAGAAACACAGGAAA

KLF1

ATGGCGACTGCGGAGACAGCACTTCCATCAATCTCAACACTCACTGCACTGGGGCCATTTCCAGATACCCAGGACGA
TTTCCTTAAGTGGTGGCGGTCCGAAGAGGCTCAAGACATGGGACCTGGTCCGCCGGATCCCACCGAACCTCCTCTGC
ATGTCAAAAGTGAAGATCAGCCTGGCGAGGAAGAGGATGACGAAAGGGGTGCCGACGCCACTTGGGACTTGGATCTT
CTCCTTACCAATTTCTCTGGTCCGGAACCTGGCGGGGCACCACAGACGTGCGCTCTCGCTCCCTCAGAAGCGAGCGG
GGCTCAGTACCCACCCCCTCCCGAAACTCTGGGAGCCTATGCTGGGGGTCCTGGACTGGTGGCTGGGTTGCTTGGTA
GTGAGGACCATTCTGGCTGGGTACGCCCCGCTTTGAGGGCCCGCGCTCCGGACGCCTTTGTGGGACCGGCGCTCGCT
CCTGCACCGGCTCCGGAACCAAAAGCCCTCGCGCTGCAGCCCGTGTACCCCGGACCCGGAGCCGGATCCTCAGGGGG
ATACTTCCCACGGACCGGACTCAGCGTTCCAGCGGCTTCCGGGGCGCCATACGGATTGTTGAGCGGCTACCCGGCTA
TGTATCCCGCTCCCCAGTACCAAGGACACTTCCAATTGTTCCGGGGTCTTCAAGGGCCTGCGCCCGGGCCTGCTACC
AGTCCCAGTTTCCTCAGTTGTCTGGGACCGGGAACTGTTGGCACTGGACTTGGCGGGACTGCAGAGGACCCAGGCGT
TATAGCAGAGACAGCGCCAAGTAAAAGGGGCCGACGAAGCTGGGCCAGGAAACGCCAAGCTGCGCACACTTGTGCCC
ATCCAGGTTGCGGTAAATCCTACACGAAGAGCAGTCATCTTAAAGCACATCTTCGCACACACACGGGCGAGAAGCCC
TACGCCTGTACTTGGGAAGGTTGCGGCTGGAGATTCGCTAGATCTGACGAGCTCACCCGGCATTATCGAAAACACAC
TGGCCAGCGACCGTTCCGGTGCCAACTCTGCCCAAGGGCGTTCAGTCGCTCAGATCATCTGGCTTTGCATATGAAGC
GACACCTT

KLF2

ATGGCCCTTAGTGAACCCATTCTTCCCAGCTTTTCCACGTTCGCGTCTCCTTGCCGAGAGAGAGGCCTTCAGGAAAG
GTGGCCGAGGGCTGAACCCGAGTCTGGAGGTACGGATGATGATCTTAACAGTGTGCTCGATTTCATACTCTCAATGG
GACTGGACGGGCTGGGAGCGGAGGCAGCTCCTGAACCACCACCACCCCCTCCGCCCCCAGCGTTTTACTACCCGGAG
CCAGGTGCGCCGCCGCCATATTCAGCCCCGGCGGGTGGCTTGGTGTCCGAGCTCCTCCGGCCTGAATTGGATGCCCC
GCTCGGCCCGGCGCTGCATGGTAGATTTCTGCTCGCGCCTCCGGGTCGACTCGTTAAGGCTGAACCTCCTGAGGCTG
ATGGTGGAGGTGGCTACGGATGTGCCCCCGGGCTTACCCGAGGACCGAGAGGTCTTAAGCGGGAAGGGGCACCTGGC
CCGGCTGCAAGCTGTATGCGGGGGCCCGGTGGGAGGCCTCCCCCGCCCCCTGATACACCCCCCCTTAGTCCAGATGG

ACCAGCTCGACTTCCCGCACCTGGCCCCAGAGCGAGTTTCCCCCCTCCATTTGGAGGACCGGGGTTTGGCGCCCCAG

GTCCTGGACTTCACTACGCCCCTCCTGCCCCCCCAGCTTTTGGTCTTTTCGACGATGCTGCTGCTGCCGCAGCAGCC

TTGGGCCTTGCGCCGCCCGCAGCCAGGGGACTGCTCACGCCACCGGCAAGCCCCCTGGAGCTCCTTGAAGCCAAGCC

GAAGCGAGGACGCAGATCATGGCCGCGCAAGCGGACAGCTACGCATACCTGCTCATATGCGGGCTGCGGAAAAACCT

ACACAAAGAGTTCACACCTTAAAGCGCACCTTCGCACACACACAGGCGAGAAACCATATCATTGTAACTGGGACGGA

TGTGGATGGAAATTTGCTCGGTCTGATGAGCTTACGAGACATTATCGAAAGCATACCGGACATCGGCCCTTTCAATG

CCATCTTTGTGACAGAGCTTTTTCCCGGTCTGACCACCTCGCTCTGCACATGAAGAGGCACATG

KLF3

ATGCTCATGTTTGACCCAGTTCCTGTCAAGCAAGAGGCCATGGACCCTGTCTCAGTGTCATACCCATCTAATTACAT

GGAATCCATGAAGCCTAACAAGTATGGGGTCATCTACTCCACACCATTGCCTGAGAAGTTCTTTCAGACCCCAGAAG

GTCTGTCGCACGGAATACAGATGGAGCCAGTGGACCTCACGGTGAACAAGCGGAGTTCACCCCCTTCGGCTGGGAAT

TCGCCCTCCTCTCTGAAGTTCCCGTCCTCACACCGGAGAGCCTCGCCTGGGTTGAGCATGCCTTCTTCCAGCCCACC

GATAAAAAAATACTCACCCCCTTCTCCAGGCGTGCAGCCCTTCGGCGTGCCGCTGTCCATGCCACCAGTGATGGCAG

CTGCCCTCTCGCGGCATGGAATACGGAGCCCGGGGATCCTGCCCGTCATCCAGCCGGTGGTGGTGCAGCCCGTCCCC

TTTATGTACACAAGTCACCTCCAGCAGCCTCTCATGGTCTCCTTATCGGAGGAGATGGAAAATTCCAGTAGTAGCAT

GCAAGTACCTGTAATTGAATCATATGAGAAGCCTATATCACAGAAAAAAATTAAAATAGAACCTGGGATCGAACCAC

AGAGGACAGATTATTATCCTGAAGAAATGTCACCCCCCTTAATGAACTCAGTGTCCCCCCCGCAAGCATTGTTGCAA

GAGAATCACCCTTCGGTCATCGTGCAGCCTGGGAAGAGACCTTTACCTGTGGAATCCCCGGATACTCAAAGGAAGCG

GAGGATACACAGATGTGATTATGATGGATGCAACAAAGTGTACACTAAAAGCTCCCACTTGAAAGCACACAGAAGAA

CACACACAGGAGAAAAACCCTACAAATGTACATGGGAAGGGTGCACATGGAAGTTTGCTCGGTCTGATGAACTAACA

AGACATTTCCGAAAACATACTGGAATCAAACCTTTCCAGTGCCCGGACTGTGACCGCAGCTTCTCCCGTTCTGACCA

TCTTGCCCTCCATAGGAAACGCCACATGCTAGTC

KLF5

ATGGCTACAAGGGTGCTGAGCATGAGCGCCCGCCTGGGACCCGTGCCCCAGCCGCCGGCGCCGCAGGACGAGCCGGT

GTTCGCGCAGCTCAAGCCGGTGCTGGGCGCCGCGAATCCGGCCCGCGACGCGGCGCTCTTCCCCGGCGAGGAGCTGA

AGCACGCGCACCACCGCCCGCAGGCGCAGCCCGCGCCCGCGCAGGCCCCGCAGCCGGCCCAGCCGCCCGCCACCGGC

CCGCGGCTGCCTCCAGAGGACCTGGTCCAGACAAGATGTGAAATGGAGAAGTATCTGACACCTCAGCTTCCTCCAGT

TCCTATAATTCCAGAGCATAAAAAGTATAGACGAGACAGTGCCTCAGTCGTAGACCAGTTCTTCACTGACACTGAAG

GGTTACCTTACAGTATCAACATGAACGTCTTCCTCCCTGACATCACTCACCTGAGAACTGGCCTCTACAAATCCCAG

AGACCGTGCGTAACACACATCAAGACAGAACCTGTTGCCATTTTCAGCCACCAGAGTGAAACGACTGCCCCTCCTCC

GGCCCCGACCCAGGCCCTCCCTGAGTTCACCAGTATATTCAGCTCACACCAGACCGCAGCTCCAGAGGTGAACAATA

TTTTCATCAAACAAGAACTTCCTACACCAGATCTTCATCTTTCTGTCCCTACCCAGCAGGGCCACCTGTACCAGCTA

CTGAATACACCGGATCTAGATATGCCCAGTTCTACAAATCAGACAGCAGCAATGGACACTCTTAATGTTTCTATGTC

AGCTGCCATGGCAGGCCTTAACACACACACCTCTGCTGTTCCGCAGACTGCAGTGAAACAATTCCAGGGCATGCCCC

CTTGCACATACACAATGCCAAGTCAGTTTCTTCCACAACAGGCCACTTACTTTCCCCCGTCACCACCAAGCTCAGAG

CCTGGAAGTCCAGATAGACAAGCAGAGATGCTCCAGAATTTAACCCCACCTCCATCCTATGCTGCTACAATTGCTTC

TAAACTGGCAATTCACAATCCAAATTTACCCACCACCCTGCCAGTTAACTCACAAAACATCCAACCTGTCAGATACA

ATAGAAGGAGTAACCCCGATTTGGAGAAACGACGCATCCACTACTGCGATTACCCTGGTTGCACAAAAGTTTATACC

AAGTCTTCTCATTTAAAAGCTCACCTGAGGACTCACACTGGTGAAAAGCCATACAAGTGTACCTGGGAAGGCTGCGA

CTGGAGGTTCGCGCGATCGGATGAGCTGACCCGCCACTACCGGAAGCACACAGGCGCCAAGCCCTTCCAGTGCGGGG

TGTGCAACCGCAGCTTCTCGCGCTCTGACCACCTGGCCCTGCATATGAAGAGGCACCAGAAC

KLF6

ATGGACGTGCTCCCCATGTGCAGCATCTTCCAGGAGCTCCAGATCGTGCACGAGACCGGCTACTTCTCGGCGCTGCC

GTCTCTGGAGGAGTACTGGCAACAGACCTGCCTAGAGCTGGAACGTTACCTCCAGAGCGAGCCCTGCTATGTTTCAG

CCTCAGAAATCAAATTTGACAGCCAGGAAGATCTGTGGACCAAAATCATTCTGGCTCGGGAGAAAAAGGAGGAATCC

GAACTGAAGATATCTTCCAGTCCTCCAGAGGACACTCTCATCAGCCCGAGCTTTTGTTACAACTTAGAGACCAACAG

CCTGAACTCAGATGTCAGCAGCGAATCCTCTGACAGCTCCGAGGAACTTTCTCCCACGGCCAAGTTTACCTCCGACC

CCATTGGCGAAGTTTTGGTCAGCTCGGGAAAATTGAGCTCCTCTGTCACCTCCACGCCTCCATCTTCTCCGGAACTG

AGCAGGGAACCTTCTCAACTGTGGGGTTGCGTGCCCGGGGAGCTGCCCTCGCCAGGGAAGGTGCGCAGCGGGACTTC

GGGGAAGCCAGGTGACAAGGGAAATGGCGATGCCTCCCCCGACGGCAGGAGGAGGGTGCACCGGTGCCACTTTAACG

GCTGCAGGAAAGTTTACACCAAAAGCTCCCACTTGAAAGCACACCAGCGGACGCACACAGGAGAAAAGCCTTACAGA

TGCTCATGGGAAGGGTGTGAGTGGCGTTTTGCAAGAAGTGATGAGTTAACCAGGCACTTCCGAAAGCACACCGGGGC

CAAGCCTTTTAAATGCTCCCACTGTGACAGGTGTTTTTCCAGGTCTGACCACCTGGCCCTGCACATGAAGAGGCACC
TC

KLF7

ATGGACGTGTTGGCTAGTTATAGTATATTCCAGGAGCTACAACTTGTCCACGACACCGGCTACTTCTCAGCTTTACC
ATCCCTGGAGGAGACCTGGCAGCAGACATGCCTTGAATTGGAACGCTACCTACAGACGGAGCCCCGGAGGATCTCAG
AGACCTTTGGTGAGGACTTGGACTGTTTCCTCCACGCTTCCCCTCCCCCGTGCATTGAGGAAAGCTTCCGTCGCTTA
GACCCCCTGCTGCTCCCCGTGGAAGCGGCCATCTGTGAGAAGAGCTCGGCAGTGGACATCTTGCTCTCTCGGGACAA
GTTGCTATCTGAGACCTGCCTCAGCCTCCAGCCGGCCAGCTCTTCTCTAGACAGCTACACAGCCGTCAACCAGGCCC
AGCTCAACGCAGTGACCTCATTAACGCCCCCATCGTCCCCTGAGCTCAGCCGCCATCTGGTCAAAACCTCACAAACT
CTCTCTGCCGTGGATGGCACGGTGACGTTGAAACTGGTGGCCAAGAAGGCTGCTCTCAGCTCCGTAAAGGTGGGAGG
GGTCGCAACAGCTGCAGCAGCCGTGACGGCTGCGGGGGCCGTTAAGAGTGGACAGAGCGACAGTGACCAAGGAGGGC
TAGGGGCTGAAGCATGTCCCGAAAACAAGAAGAGGGTTCACCGCTGTCAGTTTAACGGGTGCCGGAAAGTTTATACA
AAAAGCTCCCACTTAAAGGCCCACCAGAGGACTCACACAGGTGAGAAGCCTTATAAGTGCTCATGGGAGGGATGTGA
GTGGCGTTTTGCACGAAGCGATGAGCTCACGAGGCACTACAGGAAACACACAGGTGCAAAGCCCTTCAAATGCAACC
ACTGCGACAGGTGTTTTTCCAGGTCTGACCATCTTGCCCTCCACATGAAGAGACATATC

KLF8

ATGGTCGATATGGATAAACTCATAAACAACTTGGAGGTCCAACTTAATTCAGAAGGTGGCTCAATGCAGGTATTCAA
GCAGGTCACTGCTTCTGTTCGGAACAGAGATCCCCCTGAGATAGAATACAGAAGTAATATGACTTCTCCAACACTCC
TGGATGCCAACCCCATGGAGAACCCAGCACTGTTTAATGACATCAAGATTGAGCCCCCAGAAGAACTTTTGGCTAGT
GATTTCAGCCTGCCCCAAGTGGAACCAGTTGACCTCTCCTTTCACAAGCCCAAGGCTCCTCTCCAGCCTGCTAGCAT
GCTACAAGCTCCAATACGTCCCCCCAAGCCACAGTCTTCTCCCCAGACCCTTGTGGTGTCCACGTCAACATCTGACA
TGAGCACTTCAGCAAACATTCCTACTGTTCTGACCCCAGGCTCTGTCCTGACCTCCTCTCAGAGCACTGGTAGCCAG
CAGATCTTACATGTCATTCACACTATCCCCTCAGTCAGTCTGCCAAATAAGATGGGTGGCCTGAAGACCATCCCAGT
GGTAGTGCAGTCTCTGCCCATGGTGTATACTACTTTGCCTGCAGATGGGGGCCCTGCAGCCATTACAGTCCCACTCA
TTGGAGGAGATGGTAAAAATGCTGGATCAGTGAAAGTTGACCCCACCTCCATGTCTCCACTGGAAATTCCAAGTGAC
AGTGAGGAGAGTACAATTGAGAGTGGATCCTCAGCCTTGCAGAGTCTGCAGGGACTACAGCAAGAACCAGCAGCAAT

GGCCCAAATGCAGGGAGAAGAGTCGCTTGACTTGAAGAGAAGACGGATTCACCAATGTGACTTTGCAGGATGCAGCA

AAGTGTACACCAAAAGCTCTCACCTGAAAGCTCACCGCAGAATCCATACAGGAGAGAAGCCTTATAAATGCACCTGG

GATGGCTGCTCCTGGAAATTTGCTCGCTCAGATGAGCTCACTCGCCATTTCCGCAAGCACACAGGCATCAAGCCTTT

TCGGTGCACAGACTGCAACCGCAGCTTTTCTCGTTCTGACCACCTGTCCCTGCATCGCCGTCGCCATGACACCATG

KLF9

ATGTCCGCGGCCGCCTACATGGACTTCGTGGCTGCCCAGTGTCTGGTTTCCATTTCGAACCGCGCTGCGGTGCCGGA

GCATGGGGTCGCTCCGGACGCCGAGCGGCTGCGACTACCTGAGCGCGAGGTGACCAAGGAGCACGGTGACCCGGGGG

ACACCTGGAAGGATTACTGCACACTGGTCACCATCGCCAAGAGCTTGTTGGACCTGAACAAGTACCGACCCATCCAG

ACCCCCTCCGTGTGCAGCGACAGTCTGGAAAGTCCAGATGAGGATATGGGATCCGACAGCGACGTGACCACCGAATC

TGGGTCGAGTCCTTCCCACAGCCCGGAGGAGAGACAGGATCCTGGCAGCGCGCCCAGCCCGCTCTCCCTCCTCCATC

CTGGAGTGGCTGCGAAGGGGAAACACGCCTCCGAAAAGAGGCACAAGTGCCCCTACAGTGGCTGTGGGAAAGTCTAT

GGAAAATCCTCCCATCTCAAAGCCCATTACAGAGTGCATACAGGTGAACGGCCCTTTCCCTGCACGTGGCCAGACTG

CCTTAAAAAGTTCTCCCGCTCAGACGAGCTGACCCGCCACTACCGGACCCACACTGGGGAAAAGCAGTTCCGCTGTC

CGCTGTGTGAGAAGCGCTTCATGAGGAGTGACCACCTCACAAAGCACGCCCGGCGGCACACCGAGTTCCACCCCAGC

ATGATCAAGCGATCGAAAAAGGCGCTGGCCAACGCTTTG

KLF10

ATGCTCAACTTCGGTGCCTCTCTCCAGCAGACTGCGGAGGAAAGAATGGAAATGATTTCTGAAAGGCCAAAAGAGAG

TATGTATTCCTGGAACAAAACTGCAGAGAAAAGTGATTTTGAAGCTGTAGAAGCACTTATGTCAATGAGCTGCAGTT

GGAAGTCTGATTTTAAGAAATACGTTGAAAACAGACCTGTTACACCAGTATCTGATTTGTCAGAGGAAGAGAATCTG

CTTCCGGGAACACCTGATTTTCATACAATCCCAGCATTTTGTTTGACTCCACCTTACAGTCCTTCTGACTTTGAACC

CTCTCAAGTGTCAAATCTGATGGCACCAGCGCCATCTACTGTACACTTCAAGTCACTCTCAGATACTGCCAAACCTC

ACATTGCCGCACCTTTCAAAGAGGAAGAAAAGAGCCCAGTATCTGCCCCCAAACTCCCCAAAGCTCAGGCAACAAGT

GTGATTCGTCATACAGCTGATGCCCAGCTATGTAACCACCAGACCTGCCCAATGAAAGCAGCCAGCATCCTCAACTA

TCAGAACAATTCTTTTAGAAGAAGAACCCACCTAAATGTTGAGGCTGCAAGAAAGAACATACCATGTGCCGCTGTGT

CACCAAACAGATCCAAATGTGAGAGAAACACAGTGGCAGATGTTGATGAGAAAGCAAGTGCTGCACTTTATGACTTT

TCTGTGCCTTCCTCAGAGACGGTCATCTGCAGGTCTCAGCCAGCCCCTGTGTCCCCACAACAGAAGTCAGTGTTGGT

CTCTCCACCTGCAGTATCTGCAGGGGGAGTGCCACCTATGCCGGTCATCTGCCAGATGGTTCCCCTTCCTGCCAACA

ACCCTGTTGTGACAACAGTCGTTCCCAGCACTCCTCCCAGCCAGCCACCAGCCGTTTGCCCCCCTGTTGTGTTCATG

GGCACACAAGTCCCCAAAGGCGCTGTCATGTTTGTGGTACCCCAGCCCGTTGTGCAGAGTTCAAAGCCTCCGGTGGT

GAGCCCGAATGGCACCAGACTCTCTCCCATTGCCCCTGCTCCTGGGTTTTCCCCTTCAGCAGCAAAAGTCACTCCTC

AGATTGATTCATCAAGGATAAGGAGTCACATCTGTAGCCACCCAGGATGTGGCAAGACATACTTTAAAAGTTCCCAT

CTGAAGGCCCACACGAGGACGCACACAGGAGAAAAGCCTTTCAGCTGTAGCTGGAAAGGTTGTGAAAGGAGGTTTGC

CCGTTCTGATGAACTGTCCAGACACAGGCGAACCCACACGGGTGAGAAGAAATTTGCGTGCCCCATGTGTGACCGGC

GGTTCATGAGGAGTGACCATTTGACCAAGCATGCCCGGCGCCATCTATCAGCCAAGAAGCTACCAAACTGGCAGATG

GAAGTGAGCAAGCTAAATGACATTGCTCTACCTCCAACCCCTGCTCCCACACAG

KLF11

ATGCATACTCCTGATTTCGCTGGACCTGACGACGCCCGAGCCGTGGACATTATGGACATTTGTGAATCTATACTCGA

AAGAAAGAGACATGATTCAGAGCGAAGTACATGCTCTATCCTCGAGCAAACAGACATGGAGGCGGTAGAAGCTCTGG

TGTGCATGTCCAGTTGGGGTCAGAGATCCCAGAAGGGGGACTTGCTTAGAATCCGACCGCTTACTCCAGTTTCCGAT

AGCGGCGACGTAACAACTACTGTTCATATGGACGCAGCCACGCCTGAGCTGCCCAAAGACTTTCACAGCCTCTCAAC

TCTTTGCATCACTCCACCACAGTCCCCCGATCTTGTCGAACCATCAACCCGGACCCCTGTTAGCCCGCAAGTTACAG

ATTCAAAGGCGTGTACCGCGACCGATGTTCTGCAGAGTTCAGCGGTTGTAGCGCGGGCATTGAGCGGAGGGGCTGAA

CGAGGTCTGTTGGGTCTTGAACCCGTACCGAGTTCTCCTTGTAGAGCCAAGGGTACTAGTGTTATTCGGCATACCGG

CGAGAGTCCGGCAGCTTGTTTCCCCACCATACAAACCCCAGACTGTCGCCTTAGTGATTCCCGGGAAGGGGAGGAAC

AGCTGTTGGGCCACTTCGAGACACTTCAAGATACACACTTGACAGATAGCTTGCTGTCCACCAACCTGGTGTCATGT

CAACCTTGTTTGCACAAGTCCGGGGGTCTCCTTCTGACTGACAAAGGTCAACAAGCGGGATGGCCTGGCGCTGTCCA

AACATGCAGTCCTAAAAACTACGAAAATGATTTGCCTAGGAAAACCACGCCGCTTATCAGTGTGAGTGTTCCCGCTC

CACCTGTCCTGTGCCAGATGATCCCTGTAACCGGGCAATCATCTATGTTGCCTGCGTTCTTGAAGCCCCCCCCACAA

CTGTCCGTTGGTACTGTTCGCCCGATCCTTGCGCAAGCAGCGCCCGCCCCGCAACCCGTGTTCGTGGGGCCCGCTGT

CCCGCAGGGTGCAGTCATGTTGGTTCTTCCCCAGGGGGCCCTCCCGCCACCAGCTCCGTGTGCAGCGAATGTCATGG

CTGCCGGAAACACGAAATTGTTGCCCCTTGCACCCGCTCCAGTTTTCATAACGAGCTCACAGAATTGTGTGCCACAA

GTCGACTTCTCACGAAGACGGAACTATGTGTGCTCTTTCCCAGGTTGCAGAAAAACATATTTCAAATCCTCTCATCT

GAAAGCACATCTTCGGACCCATACAGGAGAGAAGCCTTTTAATTGTAGCTGGGATGGCTGTGATAAAAAATTCGCAA

GAAGTGATGAGCTCAGTCGACATCGCAGGACGCATACCGGGGAAAAAAAATTCGTTTGTCCAGTTTGTGACAGAAGA

TTTATGAGGTCCGACCATCTCACCAAGCACGCGCGACGCCACATGACTACAAAGAAAATTCCTGGCTGGCAAGCCGA

GGTGGGAAAACTCAACCGAATCGCTTCCGCTGAATCCCCCGGCAGCCCGCTGGTAAGTATGCCTGCCAGTGCC

KLF12

ATGAACATTCACATGAAGCGCAAGACGATAAAGAACATCAATACATTCGAGAACCGAATGTTGATGTTGGATGGCAT

GCCCGCTGTACGGGTAAAAACCGAGCTCCTGGAGTCTGAACAAGGATCCCCAAACGTCCACAACTACCCGGATATGG

AGGCAGTGCCGCTCTTGCTCAACAATGTGAAGGGAGAGCCGCCTGAGGACTCTCTCTCCGTAGATCATTTCCAGACA

CAGACTGAGCCCGTAGATCTTTCAATTAACAAAGCCAGAACATCTCCTACTGCGGTAAGTTCTTCTCCCGTAAGTAT

GACAGCAAGTGCATCTAGTCCAAGTTCTACGAGCACTAGCAGTTCTTCATCTAGTAGACTTGCTAGTTCACCAACGG

TGATCACAAGTGTTTCTAGCGCCAGCAGCAGCTCAACGGTACTGACTCCCGGTCCACTCGTGGCAAGCGCTAGTGGC

GTGGGTGGCCAACAATTTCTCCATATTATTCACCCCGTGCCTCCGTCTAGTCCGATGAATCTCCAGAGCAACAAGCT

TAGTCACGTACATAGGATCCCCGTCGTCGTCCAGTCAGTTCCCGTCGTCTACACAGCTGTGCGATCCCCTGGGAATG

TCAATAATACTATAGTTGTTCCTTTGCTTGAGGATGGTAGGGGCCATGGGAAAGCACAGATGGACCCCCGCGGCTTG

TCACCGAGACAGTCTAAATCCGATAGTGACGACGATGATTTGCCTAACGTAACACTGGACTCTGTGAACGAGACCGG

GAGTACCGCTCTGTCAATCGCTAGGGCCGTACAGGAGGTCCACCCAAGCCCTGTGTCACGAGTCCGAGGTAACAGGA

TGAATAATCAGAAATTTCCCTGTAGCATCAGCCCATTTTCTATAGAGTCCACTCGGAGACAGCGACGAAGTGAATCA

CCCGACTCCAGAAAAAGGAGGATACATCGCTGTGACTTTGAGGGCTGTAACAAGGTCTACACAAAAAGTTCACACCT

CAAGGCGCATCGACGGACGCATACTGGGGAAAAACCGTACAAATGCACCTGGGAGGGATGCACGTGGAAATTTGCAC

GCTCTGACGAGTTGACACGCCACTATCGAAAGCATACGGGCGTAAAGCCGTTTAAATGCGCTGATTGCGACAGGAGT

TTTAGCCGCTCTGATCACCTTGCTCTTCACCGGAGGCGACACATGCTTGTT

KLF13

ATGGCTGCGGCTGCATATGTGGATCATTTTGCGGCTGAGTGCCTGGTGTCAATGTCTAGTAGAGCGGTGGTACACGG

TCCCAGAGAAGGCCCAGAATCACGCCCAGAGGGCGCCGCCGTCGCTGCAACACCGACGCTGCCTCGGGTCGAGGAGC

GCCGCGACGGGAAGGACAGTGCGTCACTTTTCGTAGTAGCGAGAATATTGGCAGATCTGAATCAACAGGCTCCAGCA

CCTGCGCCCGCTGAACGCCGGGAGGGCGCCGCTGCCAGAAAGGCCAGAACACCATGCCGCTTGCCGCCACCTGCGCC

AGAACCCACAAGTCCAGGTGCCGAAGGTGCGGCGGCTGCCCCTCCTTCACCGGCCTGGTCTGAACCAGAACCAGAGG

CAGGTCTTGAACCTGAGCGCGAACCCGGCCCTGCAGGCTCTGGGGAACCTGGCCTGAGGCAGCGGGTGAGGCGCGGC

CGGAGCAGGGCCGACCTGGAATCACCGCAAAGGAAACATAAATGCCATTATGCTGGTTGCGAAAAGGTTTATGGAAA

GTCATCCCACCTGAAAGCACACCTCCGCACTCACACGGGTGAGCGACCTTTTGCGTGTTCCTGGCAAGACTGCAATA

AAAAGTTTGCTAGATCTGATGAACTTGCACGGCATTATCGAACTCATACCGGTGAAAAGAAGTTCTCATGCCCTATA

TGTGAGAAACGGTTCATGCGCTCTGACCACTTGACGAAACATGCAAGACGACATGCTAATTTTCATCCGGGGATGTT

GCAGAGACGGGGAGGGGGAAGTAGGACTGGAAGTCTCTCCGACTATTCCCGATCCGACGCTTCCTCACCAACGATTA

GCCCCGCAAGCAGTCCC

KLF14

ATGTCAGCCGCAGTCGCATGCCTTGATTACTTCGCGGCCGAGTGTCTTGTTTCCATGTCAGCGGGGGCTGTCGTTCA

CAGAAGACCACCAGACCCGGAGGGAGCGGGAGGGGCAGCTGGATCTGAAGTCGGCGCGGCTCCACCTGAATCAGCGC

TTCCCGGCCCTGGTCCTCCAGGTCCCGCTAGCGTGCCCCAACTCCCACAAGTGCCTGCTCCGAGTCCTGGAGCGGGC

GGAGCAGCCCCGCATCTCCTTGCAGCATCAGTGTGGGCCGATCTTCGCGGAAGCTCCGGGGAGGGCTCCTGGGAAAA

CAGCGGAGAGGCCCCGCGAGCTTCAAGCGGCTTTTCCGATCCAATCCCTTGCAGTGTTCAAACCCCATGCTCCGAGC

TCGCGCCCGCGTCCGGAGCTGCGGCAGTGTGCGCACCTGAAAGCTCATCCGATGCGCCGGCCGTTCCATCTGCGCCA

GCTGCTCCCGGTGCACCCGCAGCATCTGGCGGCTTTAGTGGTGGAGCTCTTGGGGCGGGTCCCGCCCCTGCGGCGGA

TCAAGCTCCTCGCAGGCGCAGTGTTACGCCCGCAGCAAAACGGCATCAATGCCCCTTTCCTGGTTGTACAAAAGCAT

ACTATAAGTCATCCCATCTCAAGAGTCACCAGAGGACGCATACAGGTGAGAGACCTTTTAGCTGTGACTGGCTCGAT

TGCGACAAGAAATTTACGCGGAGCGACGAACTTGCGCGGCACTACCGCACTCACACTGGAGAAAAGAGGTTCTCTTG

TCCCCTGTGTCCCAAGCAGTTCTCACGCAGTGATCACTTGACAAAACATGCTAGGAGACATCCAACATACCATCCCG

ACATGATAGAGTATCGAGGTAGGCGACGCACACCTAGAATTGATCCTCCGCTGACTAGTGAAGTCGAGTCAAGTGCC

AGTGGAAGCGGACCGGGTCCCGCGCCCTCATTTACAACCTGTCTT

KLF15

ATGGTGGACCACTTACTTCCAGTGGACGAGAACTTCTCGTCGCCAAAATGCCCAGTTGGGTATCTGGGTGATAGGCT

GGTTGGCCGGCGGGCATATCACATGCTGCCCTCACCCGTCTCTGAAGATGACAGCGATGCCTCCAGCCCCTGCTCCT

GTTCCAGTCCCGACTCTCAAGCCCTCTGCTCCTGCTATGGTGGAGGCCTGGGCACCGAGAGCCAGGACAGCATCTTG

GACTTCCTATTGTCCCAGGCCACGCTGGGCAGTGGCGGGGGCAGCGGCAGTAGCATTGGGGCCAGCAGTGGCCCCGT

GGCCTGGGGGCCCTGGCGAAGGGCAGCGGCCCCTGTGAAGGGGGAGCATTTCTGCTTGCCCGAGTTTCCTTTGGGTG

ATCCTGATGACGTCCCACGGCCCTTCCAGCCTACCCTGGAGGAGATTGAAGAGTTTCTGGAGGAGAACATGGAGCCT

GGAGTCAAGGAGGTCCCTGAGGGCAACAGCAAGGACTTGGATGCCTGCAGCCAGCTCTCAGCTGGGCCACACAAGAG

CCACCTCCATCCTGGGTCCAGCGGGAGAGAGCGCTGTTCCCCTCCACCAGGTGGTGCCAGTGCAGGAGGTGCCCAGG

GCCCAGGTGGGGGCCCCACGCCTGATGGCCCCATCCCAGTGTTGCTGCAGATCCAGCCCGTGCCTGTGAAGCAGGAA

TCGGGCACAGGGCCTGCCTCCCCTGGGCAAGCCCCAGAGAATGTCAAGGTTGCCCAGCTCCTGGTCAACATCCAGGG

GCAGACCTTCGCACTCGTGCCCCAGGTGGTACCCTCCTCCAACTTGAACCTGCCCTCCAAGTTTGTGCGCATTGCCC

CTGTGCCCATTGCCGCCAAGCCTGTTGGATCGGGACCCCTGGGGCCTGGCCCTGCCGGTCTCCTCATGGGCCAGAAG

TTCCCCAAGAACCCAGCCGCAGAACTCATCAAAATGCACAAATGTACTTTCCCTGGCTGCAGCAAGATGTACACCAA

AAGCAGCCACCTCAAGGCCCACCTGCGCCGGCACACGGGTGAGAAGCCCTTCGCCTGCACCTGGCCAGGCTGCGGCT

GGAGGTTCTCGCGCTCTGACGAGCTGTCGCGGCACAGGCGCTCGCACTCAGGTGTGAAGCCGTACCAGTGTCCTGTG

TGCGAGAAGAAGTTCGCGCGGAGCGACCACCTCTCCAAGCACATCAAGGTGCACCGCTTCCCGCGGAGCAGCCGCTC

CGTGCGCTCCGTGAAC

KLF16

ATGTCAGCCGCGGTCGCGTGCGTGGATTATTTTGCAGCAGATGTGCTGATGGCAATTTCATCCGGTGCAGTAGTTCA

TCGCGGAAGACCAGGTCCTGAGGGTGCGGGGCCTGCGGCCGGGTTGGATGTTCGCGCCGCGCGCAGGGAAGCCGCTT

CTCCCGGAACACCTGGCCCTCCTCCTCCTCCGCCGGCGGCATCAGGCCCGGGTCCTGGTGCAGCTGCGGCTCCTCAC

CTGTTGGCAGCCTCCATACTGGCTGACCTGCGAGGGGGGCCAGGCGCTGCACCTGGTGGCGCGAGTCCAGCAAGTTC

CAGCTCCGCGGCGTCCTCCCCGAGTAGTGGGCGAGCTCCGGGCGCGGCACCTTCTGCTGCCGCTAAATCACACCGAT

GCCCTTTCCCAGACTGCGCGAAGGCGTATTATAAGTCCAGTCATTTGAAATCACACTTGAGGACACATACCGGCGAG

AGACCTTTTGCGTGCGACTGGCAGGGTTGTGATAAGAAATTTGCGAGAAGCGACGAACTGGCCCGCCATCACCGCAC

CCACACAGGGGAAAAAAGATTCTCATGCCCACTCTGTTCTAAGCGCTTCACGCGAAGCGACCATCTTGCAAAGCACG

CTAGGAGACACCCTGGGTTCCACCCCGACCTCTTGCGACGACCTGGCGCCCGGTCTACTAGCCCGTCTGACTCATTG

CCGTGCTCTCTCGCAGGGTCCCCTGCTCCGAGCCCCGCACCGTCCCCAGCTCCTGCCGGGCTT

KLF17

ATGTACGGCCGACCGCAGGCTGAGATGGAACAGGAGGCTGGGGAGCTGAGCCGGTGGCAGGCGGCGCACCAGGCTGC

CCAGGATAACGAGAACTCAGCGCCCATCTTGAACATGTCTTCATCTTCTGGAAGCTCTGGAGTGCACACCTCTTGGA

ACCAAGGCCTACCAAGCATTCAGCACTTTCCTCACAGCGCAGAGATGCTGGGGTCCCCTTTGGTGTCTGTTGAGGCG

CCGGGGCAGAATGTGAATGAAGGGGGGCCACAGTTCAGTATGCCACTGCCTGAGCGTGGTATGAGCTACTGCCCCCA

AGCGACTCTCACTCCTTCCCGGATGATTTACTGTCAGAGAATGTCTCCCCCTCAGCAAGAGATGACGATTTTCAGTG

GGCCCCAACTAATGCCCGTAGGAGAGCCCAATATTCCAAGGGTAGCCAGGCCCTTCGGTGGGAATCTAAGGATGCCC

CCCAATGGGCTGCCAGTCTCGGCTTCCACTGGAATCCCAATAATGTCCCACACTGGGAACCCTCCAGTGCCTTACCC

TGGCCTCTCGACAGTACCTTCTGACGAAACATTGTTGGGCCCGACTGTGCCTTCCACTGAGGCCCAGGCAGTGCTCC

CCTCCATGGCTCAGATGTTGCCCCCGCAAGATGCCCATGACCTTGGGATGCCCCCAGCTGAGTCCCAGTCATTGCTG

GTTTTAGGATCTCAGGACTCTCTTGTCAGTCAGCCAGACTCTCAAGAAGGCCCATTTCTACCAGAGCAGCCCGGACC

TGCTCCACAGACAGTAGAGAAGAACTCCAGGCCTCAGGAAGGGACTGGTAGAAGGGGCTCCTCAGAGGCAAGGCCTT

ACTGCTGCAACTACGAGAACTGCGGAAAAGCTTATACCAAACGCTCCCACCTCGTGAGCCACCAGCGCAAGCACACA

GGTGAGAGGCCATATTCTTGCAACTGGGAAAGTTGTTCATGGTCTTTCTTCCGTTCTGATGAGCTTAGACGACATAT

GCGGGTACACACCAGATATCGACCATATAAATGTGATCAGTGCAGCCGGGAGTTCATGAGGTCTGACCATCTCAAGC

AACACCAGAAGACTCATCGGCCGGGACCCTCAGACCCACAGGCCAACAACAACAATGGAGAGCAGGACAGTCCTCCT

GCTGCTGGTCCT

MYC ΔMBI

ATGCCCCTCAACGTTAGCTTCACCAACAGGAACTATGACCTCGACTACGACTCGGTGCAGCCGTATTTCTACTGCGA

CGAGGAGGAGAACTTCTACCAGCAGCAGCAGCAGAGCGAGCTGCAGCCCCCGGCGGGATCAGGTAGCGGTAGCCGCC

GCTCCGGGCTCTGCTCGCCCTCCTACGTTGCGGTCACACCCTTCTCCCTTCGGGGAGACAACGACGGCGGTGGCGGG

AGCTTCTCCACGGCCGACCAGCTGGAGATGGTGACCGAGCTGCTGGGAGGAGACATGGTGAACCAGAGTTTCATCTG

CGACCCGGACGACGAGACCTTCATCAAAAACATCATCATCCAGGACTGTATGTGGAGCGGCTTCTCGGCCGCCGCCA

AGCTCGTCTCAGAGAAGCTGGCCTCCTACCAGGCTGCGCGCAAAGACAGCGGCAGCCCGAACCCCGCCCGCGGCCAC

AGCGTCTGCTCCACCTCCAGCTTGTACCTGCAGGATCTGAGCGCCGCCGCCTCAGAGTGCATCGACCCCTCGGTGGT

CTTCCCCTACCCTCTCAACGACAGCAGCTCGCCCAAGTCCTGCGCCTCGCAAGACTCCAGCGCCTTCTCTCCGTCCT

CGGATTCTCTGCTCTCCTCGACGGAGTCCTCCCCGCAGGGCAGCCCCGAGCCCCTGGTGCTCCATGAGGAGACACCG

CCCACCACCAGCAGCGACTCTGAGGAGGAACAAGAAGATGAGGAAGAAATCGATGTTGTTTCTGTGGAAAAGAGGCA

GGCTCCTGGCAAAAGGTCAGAGTCTGGATCACCTTCTGCTGGAGGCCACAGCAAACCTCCTCACAGCCCACTGGTCC

TCAAGAGGTGCCACGTCTCCACACATCAGCACAACTACGCAGCGCCTCCCTCCACTCGGAAGGACTATCCTGCTGCC

AAGAGGGTCAAGTTGGACAGTGTCAGAGTCCTGAGACAGATCAGCAACAACCGAAAATGCACCAGCCCCAGGTCCTC

GGACACCGAGGAGAATGTCAAGAGGCGAACACACAACGTCTTGGAGCGCCAGAGGAGGAACGAGCTAAAACGGAGCT

TTTTTGCCCTGCGTGACCAGATCCCGGAGTTGGAAAACAATGAAAAGGCCCCCAAGGTAGTTATCCTTAAAAAAGCC

ACAGCATACATCCTGTCCGTCCAAGCAGAGGAGCAAAAGCTCATTTCTGAAGAGGACTTGTTGCGGAAACGACGAGA

ACAGTTGAAACACAAACTTGAACAGCTACGGAACTCTTGTGCG

MYC ΔMBII

ATGCCCCTCAACGTTAGCTTCACCAACAGGAACTATGACCTCGACTACGACTCGGTGCAGCCGTATTTCTACTGCGA

CGAGGAGGAGAACTTCTACCAGCAGCAGCAGCAGAGCGAGCTGCAGCCCCCGGCGCCCAGCGAGGATATCTGGAAGA

AATTCGAGCTGCTGCCCACCCCGCCCCTGTCCCCTAGCCGCCGCTCCGGGCTCTGCTCGCCCTCCTACGTTGCGGTC

ACACCCTTCTCCCTTCGGGGAGACAACGACGGCGGTGGCGGGAGCTTCTCCACGGCCGACCAGCTGGAGATGGTGAC

CGAGCTGCTGGGAGGAGACATGGTGAACCAGAGTTTCATCTGCGACCCGGACGACGAGACCTTCATCAAAAACATCG

GATCAGGTAGCGGTCTCGTCTCAGAGAAGCTGGCCTCCTACCAGGCTGCGCGCAAAGACAGCGGCAGCCCGAACCCC

GCCCGCGGCCACAGCGTCTGCTCCACCTCCAGCTTGTACCTGCAGGATCTGAGCGCCGCCGCCTCAGAGTGCATCGA

CCCCTCGGTGGTCTTCCCCTACCCTCTCAACGACAGCAGCTCGCCCAAGTCCTGCGCCTCGCAAGACTCCAGCGCCT

TCTCTCCGTCCTCGGATTCTCTGCTCTCCTCGACGGAGTCCTCCCCGCAGGGCAGCCCCGAGCCCCTGGTGCTCCAT

GAGGAGACACCGCCCACCACCAGCAGCGACTCTGAGGAGGAACAAGAAGATGAGGAAGAAATCGATGTTGTTTCTGT

GGAAAAGAGGCAGGCTCCTGGCAAAAGGTCAGAGTCTGGATCACCTTCTGCTGGAGGCCACAGCAAACCTCCTCACA

GCCCACTGGTCCTCAAGAGGTGCCACGTCTCCACACATCAGCACAACTACGCAGCGCCTCCCTCCACTCGGAAGGAC

TATCCTGCTGCCAAGAGGGTCAAGTTGGACAGTGTCAGAGTCCTGAGACAGATCAGCAACAACCGAAAATGCACCAG

CCCCAGGTCCTCGGACACCGAGGAGAATGTCAAGAGGCGAACACACAACGTCTTGGAGCGCCAGAGGAGGAACGAGC

TAAAACGGAGCTTTTTTGCCCTGCGTGACCAGATCCCGGAGTTGGAAAACAATGAAAAGGCCCCCAAGGTAGTTATC

CTTAAAAAAGCCACAGCATACATCCTGTCCGTCCAAGCAGAGGAGCAAAAGCTCATTTCTGAAGAGGACTTGTTGCG

GAAACGACGAGAACAGTTGAAACACAAACTTGAACAGCTACGGAACTCTTGTGCG

MYC ΔNLS

ATGCCCCTCAACGTTAGCTTCACCAACAGGAACTATGACCTCGACTACGACTCGGTGCAGCCGTATTTCTACTGCGA

CGAGGAGGAGAACTTCTACCAGCAGCAGCAGCAGAGCGAGCTGCAGCCCCCGGCGCCCAGCGAGGATATCTGGAAGA

AATTCGAGCTGCTGCCCACCCCGCCCCTGTCCCCTAGCCGCCGCTCCGGGCTCTGCTCGCCCTCCTACGTTGCGGTC

ACACCCTTCTCCCTTCGGGGAGACAACGACGGCGGTGGCGGGAGCTTCTCCACGGCCGACCAGCTGGAGATGGTGAC

CGAGCTGCTGGGAGGAGACATGGTGAACCAGAGTTTCATCTGCGACCCGGACGACGAGACCTTCATCAAAAACATCA

TCATCCAGGACTGTATGTGGAGCGGCTTCTCGGCCGCCGCCAAGCTCGTCTCAGAGAAGCTGGCCTCCTACCAGGCT

GCGCGCAAAGACAGCGGCAGCCCGAACCCCGCCCGCGGCCACAGCGTCTGCTCCACCTCCAGCTTGTACCTGCAGGA

TCTGAGCGCCGCCGCCTCAGAGTGCATCGACCCCTCGGTGGTCTTCCCCTACCCTCTCAACGACAGCAGCTCGCCCA

AGTCCTGCGCCTCGCAAGACTCCAGCGCCTTCTCTCCGTCCTCGGATTCTCTGCTCTCCTCGACGGAGTCCTCCCCG

CAGGGCAGCCCCGAGCCCCTGGTGCTCCATGAGGAGACACCGCCCACCACCAGCAGCGACTCTGAGGAGGAACAAGA

AGATGAGGAAGAAATCGATGTTGTTTCTGTGGAAAAGAGGCAGGCTCCTGGCAAAAGGTCAGAGTCTGGATCACCTT

CTGCTGGAGGCCACAGCAAACCTCCTCACAGCCCACTGGTCCTCAAGAGGTGCCACGTCTCCACACATCAGCACAAC

TACGCAGCGCCTCCCTCCACTCGGAAGGACTATGGATCAGGTAGCGGTAGTGTCAGAGTCCTGAGACAGATCAGCAA

CAACCGAAAATGCACCAGCCCCAGGTCCTCGGACACCGAGGAGAATGTCAAGAGGCGAACACACAACGTCTTGGAGC

GCCAGAGGAGGAACGAGCTAAAACGGAGCTTTTTTGCCCTGCGTGACCAGATCCCGGAGTTGGAAAACAATGAAAAG

GCCCCCAAGGTAGTTATCCTTAAAAAAGCCACAGCATACATCCTGTCCGTCCAAGCAGAGGAGCAAAAGCTCATTTC

TGAAGAGGACTTGTTGCGGAAACGACGAGAACAGTTGAAACACAAACTTGAACAGCTACGGAACTCTTGTGCG

MYC Δb

ATGCCCCTCAACGTTAGCTTCACCAACAGGAACTATGACCTCGACTACGACTCGGTGCAGCCGTATTTCTACTGCGA

CGAGGAGGAGAACTTCTACCAGCAGCAGCAGCAGAGCGAGCTGCAGCCCCCGGCGCCCAGCGAGGATATCTGGAAGA

AATTCGAGCTGCTGCCCACCCCGCCCCTGTCCCCTAGCCGCCGCTCCGGGCTCTGCTCGCCCTCCTACGTTGCGGTC

ACACCCTTCTCCCTTCGGGGAGACAACGACGGCGGTGGCGGGAGCTTCTCCACGGCCGACCAGCTGGAGATGGTGAC

CGAGCTGCTGGGAGGAGACATGGTGAACCAGAGTTTCATCTGCGACCCGGACGACGAGACCTTCATCAAAAACATCA

TCATCCAGGACTGTATGTGGAGCGGCTTCTCGGCCGCCGCCAAGCTCGTCTCAGAGAAGCTGGCCTCCTACCAGGCT

GCGCGCAAAGACAGCGGCAGCCCGAACCCCGCCCGCGGCCACAGCGTCTGCTCCACCTCCAGCTTGTACCTGCAGGA

TCTGAGCGCCGCCGCCTCAGAGTGCATCGACCCCTCGGTGGTCTTCCCCTACCCTCTCAACGACAGCAGCTCGCCCA

AGTCCTGCGCCTCGCAAGACTCCAGCGCCTTCTCTCCGTCCTCGGATTCTCTGCTCTCCTCGACGGAGTCCTCCCCG

CAGGGCAGCCCCGAGCCCCTGGTGCTCCATGAGGAGACACCGCCCACCACCAGCAGCGACTCTGAGGAGGAACAAGA

AGATGAGGAAGAAATCGATGTTGTTTCTGTGGAAAAGAGGCAGGCTCCTGGCAAAAGGTCAGAGTCTGGATCACCTT

CTGCTGGAGGCCACAGCAAACCTCCTCACAGCCCACTGGTCCTCAAGAGGTGCCACGTCTCCACACATCAGCACAAC

TACGCAGCGCCTCCCTCCACTCGGAAGGACTATCCTGCTGCCAAGAGGGTCAAGTTGGACAGTGTCAGAGTCCTGAG

ACAGATCAGCAACAACCGAAAATGCACCAGCCCCAGGTCCTCGGACACCGAGGAGAATGTCGGATCAGGTAGCGGTG

AGCTAAAACGGAGCTTTTTTGCCCTGCGTGACCAGATCCCGGAGTTGGAAAACAATGAAAAGGCCCCCAAGGTAGTT

ATCCTTAAAAAAGCCACAGCATACATCCTGTCCGTCCAAGCAGAGGAGCAAAAGCTCATTTCTGAAGAGGACTTGTT

GCGGAAACGACGAGAACAGTTGAAACACAAACTTGAACAGCTACGGAACTCTTGTGCG

MYC ΔHLH

ATGCCCCTCAACGTTAGCTTCACCAACAGGAACTATGACCTCGACTACGACTCGGTGCAGCCGTATTTCTACTGCGA

CGAGGAGGAGAACTTCTACCAGCAGCAGCAGCAGAGCGAGCTGCAGCCCCCGGCGCCCAGCGAGGATATCTGGAAGA

AATTCGAGCTGCTGCCCACCCCGCCCCTGTCCCCTAGCCGCCGCTCCGGGCTCTGCTCGCCCTCCTACGTTGCGGTC

ACACCCTTCTCCCTTCGGGGAGACAACGACGGCGGTGGCGGGAGCTTCTCCACGGCCGACCAGCTGGAGATGGTGAC

CGAGCTGCTGGGAGGAGACATGGTGAACCAGAGTTTCATCTGCGACCCGGACGACGAGACCTTCATCAAAAACATCA

TCATCCAGGACTGTATGTGGAGCGGCTTCTCGGCCGCCGCCAAGCTCGTCTCAGAGAAGCTGGCCTCCTACCAGGCT

GCGCGCAAAGACAGCGGCAGCCCGAACCCCGCCCGCGGCCACAGCGTCTGCTCCACCTCCAGCTTGTACCTGCAGGA

TCTGAGCGCCGCCGCCTCAGAGTGCATCGACCCCTCGGTGGTCTTCCCCTACCCTCTCAACGACAGCAGCTCGCCCA

AGTCCTGCGCCTCGCAAGACTCCAGCGCCTTCTCTCCGTCCTCGGATTCTCTGCTCTCCTCGACGGAGTCCTCCCCG

CAGGGCAGCCCCGAGCCCCTGGTGCTCCATGAGGAGACACCGCCCACCACCAGCAGCGACTCTGAGGAGGAACAAGA

AGATGAGGAAGAAATCGATGTTGTTTCTGTGGAAAAGAGGCAGGCTCCTGGCAAAAGGTCAGAGTCTGGATCACCTT

CTGCTGGAGGCCACAGCAAACCTCCTCACAGCCCACTGGTCCTCAAGAGGTGCCACGTCTCCACACATCAGCACAAC

TACGCAGCGCCTCCCTCCACTCGGAAGGACTATCCTGCTGCCAAGAGGGTCAAGTTGGACAGTGTCAGAGTCCTGAG

ACAGATCAGCAACAACCGAAAATGCACCAGCCCCAGGTCCTCGGACACCGAGGAGAATGTCAAGAGGCGAACACACA

ACGTCTTGGAGCGCCAGAGGAGGAACGGATCAGGTAGCGGTCAAAAGCTCATTTCTGAAGAGGACTTGTTGCGGAAA

CGACGAGAACAGTTGAAACACAAACTTGAACAGCTACGGAACTCTTGTGCG

MYC ΔLZ

ATGCCCCTCAACGTTAGCTTCACCAACAGGAACTATGACCTCGACTACGACTCGGTGCAGCCGTATTTCTACTGCGA

CGAGGAGGAGAACTTCTACCAGCAGCAGCAGCAGAGCGAGCTGCAGCCCCCGGCGCCCAGCGAGGATATCTGGAAGA

AATTCGAGCTGCTGCCCACCCCGCCCCTGTCCCCTAGCCGCCGCTCCGGGCTCTGCTCGCCCTCCTACGTTGCGGTC

ACACCCTTCTCCCTTCGGGGAGACAACGACGGCGGTGGCGGGAGCTTCTCCACGGCCGACCAGCTGGAGATGGTGAC

CGAGCTGCTGGGAGGAGACATGGTGAACCAGAGTTTCATCTGCGACCCGGACGACGAGACCTTCATCAAAAACATCA

TCATCCAGGACTGTATGTGGAGCGGCTTCTCGGCCGCCGCCAAGCTCGTCTCAGAGAAGCTGGCCTCCTACCAGGCT

GCGCGCAAAGACAGCGGCAGCCCGAACCCCGCCCGCGGCCACAGCGTCTGCTCCACCTCCAGCTTGTACCTGCAGGA

TCTGAGCGCCGCCGCCTCAGAGTGCATCGACCCCTCGGTGGTCTTCCCCTACCCTCTCAACGACAGCAGCTCGCCCA

AGTCCTGCGCCTCGCAAGACTCCAGCGCCTTCTCTCCGTCCTCGGATTCTCTGCTCTCCTCGACGGAGTCCTCCCCG

CAGGGCAGCCCCGAGCCCCTGGTGCTCCATGAGGAGACACCGCCCACCACCAGCAGCGACTCTGAGGAGGAACAAGA

AGATGAGGAAGAAATCGATGTTGTTTCTGTGGAAAAGAGGCAGGCTCCTGGCAAAAGGTCAGAGTCTGGATCACCTT

CTGCTGGAGGCCACAGCAAACCTCCTCACAGCCCACTGGTCCTCAAGAGGTGCCACGTCTCCACACATCAGCACAAC

TACGCAGCGCCTCCCTCCACTCGGAAGGACTATCCTGCTGCCAAGAGGGTCAAGTTGGACAGTGTCAGAGTCCTGAG

ACAGATCAGCAACAACCGAAAATGCACCAGCCCCAGGTCCTCGGACACCGAGGAGAATGTCAAGAGGCGAACACACA

ACGTCTTGGAGCGCCAGAGGAGGAACGAGCTAAAACGGAGCTTTTTTGCCCTGCGTGACCAGATCCCGGAGTTGGAA

AACAATGAAAAGGCCCCCAAGGTAGTTATCCTTAAAAAAGCCACAGCATACATCCTGTCCGTCCAAGCAGAGGAG

MYC ΔNTD

ATGGGATCAGGTAGCGGTCTCGTCTCAGAGAAGCTGGCCTCCTACCAGGCTGCGCGCAAAGACAGCGGCAGCCCGAA

CCCCGCCCGCGGCCACAGCGTCTGCTCCACCTCCAGCTTGTACCTGCAGGATCTGAGCGCCGCCGCCTCAGAGTGCA

TCGACCCCTCGGTGGTCTTCCCCTACCCTCTCAACGACAGCAGCTCGCCCAAGTCCTGCGCCTCGCAAGACTCCAGC

GCCTTCTCTCCGTCCTCGGATTCTCTGCTCTCCTCGACGGAGTCCTCCCCGCAGGGCAGCCCCGAGCCCCTGGTGCT

CCATGAGGAGACACCGCCCACCACCAGCAGCGACTCTGAGGAGGAACAAGAAGATGAGGAAGAAATCGATGTTGTTT

CTGTGGAAAAGAGGCAGGCTCCTGGCAAAAGGTCAGAGTCTGGATCACCTTCTGCTGGAGGCCACAGCAAACCTCCT

CACAGCCCACTGGTCCTCAAGAGGTGCCACGTCTCCACACATCAGCACAACTACGCAGCGCCTCCCTCCACTCGGAA

GGACTATCCTGCTGCCAAGAGGGTCAAGTTGGACAGTGTCAGAGTCCTGAGACAGATCAGCAACAACCGAAAATGCA

CCAGCCCCAGGTCCTCGGACACCGAGGAGAATGTCAAGAGGCGAACACACAACGTCTTGGAGCGCCAGAGGAGGAAC

GAGCTAAAACGGAGCTTTTTTGCCCTGCGTGACCAGATCCCGGAGTTGGAAAACAATGAAAAGGCCCCCAAGGTAGT

TATCCTTAAAAAAGCCACAGCATACATCCTGTCCGTCCAAGCAGAGGAGCAAAAGCTCATTTCTGAAGAGGACTTGT

TGCGGAAACGACGAGAACAGTTGAAACACAAACTTGAACAGCTACGGAACTCTTGTGCG

MYC ΔCTD

ATGCCCCTCAACGTTAGCTTCACCAACAGGAACTATGACCTCGACTACGACTCGGTGCAGCCGTATTTCTACTGCGA

CGAGGAGGAGAACTTCTACCAGCAGCAGCAGCAGAGCGAGCTGCAGCCCCCGGCGCCCAGCGAGGATATCTGGAAGA

AATTCGAGCTGCTGCCCACCCCGCCCCTGTCCCCTAGCCGCCGCTCCGGGCTCTGCTCGCCCTCCTACGTTGCGGTC

ACACCCTTCTCCCTTCGGGGAGACAACGACGGCGGTGGCGGGAGCTTCTCCACGGCCGACCAGCTGGAGATGGTGAC

CGAGCTGCTGGGAGGAGACATGGTGAACCAGAGTTTCATCTGCGACCCGGACGACGAGACCTTCATCAAAAACATCA

TCATCCAGGACTGTATGTGGAGCGGCTTCTCGGCCGCCGCCAAGCTCGTCTCAGAGAAGCTGGCCTCCTACCAGGCT

GCGCGCAAAGACAGCGGCAGCCCGAACCCCGCCCGCGGCCACAGCGTCTGCTCCACCTCCAGCTTGTACCTGCAGGA

TCTGAGCGCCGCCGCCTCAGAGTGCATCGACCCCTCGGTGGTCTTCCCCTACCCTCTCAACGACAGCAGCTCGCCCA

AGTCCTGCGCCTCGCAAGACTCCAGCGCCTTCTCTCCGTCCTCGGATTCTCTGCTCTCCTCGACGGAGTCCTCCCCG

CAGGGCAGCCCCGAGCCCCTGGTGCTCCATGAGGAGACACCGCCCACCACCAGCAGCGACTCTGAGGAGGAACAAGA

AGATGAGGAAGAAATCGATGTTGTTTCTGTGGAAAAGAGGCAGGCTCCTGGCAAAAGGTCAGAGTCTGGATCACCTT

CTGCTGGAGGCCACAGCAAACCTCCTCACAGCCCACTGGTCCTCAAGAGGTGCCACGTCTCCACACATCAGCACAAC

TACGCAGCGCCTCCCTCCACTCGGAAGGACTATCCTGCTGCCAAGAGGGTCAAGTTGGACAGTGTCAGAGTCCTGAG

ACAGATCAGCAACAACCGAAAATGCACCAGCCCCAGGTCCTCGGACACCGAGGAGAATGTC

MYC Glu39Ala

ATGCCCCTCAACGTTAGCTTCACCAACAGGAACTATGACCTCGACTACGACTCGGTGCAGCCGTATTTCTACTGCGA

CGAGGAGGAGAACTTCTACCAGCAGCAGCAGCAGAGCGcGCTGCAGCCCCCGGCGCCCAGCGAGGATATCTGGAAGA

AATTCGAGCTGCTGCCCACCCCGCCCCTGTCCCCTAGCCGCCGCTCCGGGCTCTGCTCGCCCTCCTACGTTGCGGTC

ACACCCTTCTCCCTTCGGGGAGACAACGACGGCGGTGGCGGGAGCTTCTCCACGGCCGACCAGCTGGAGATGGTGAC

CGAGCTGCTGGGAGGAGACATGGTGAACCAGAGTTTCATCTGCGACCCGGACGACGAGACCTTCATCAAAAACATCA

TCATCCAGGACTGTATGTGGAGCGGCTTCTCGGCCGCCGCCAAGCTCGTCTCAGAGAAGCTGGCCTCCTACCAGGCT

GCGCGCAAAGACAGCGGCAGCCCGAACCCCGCCCGCGGCCACAGCGTCTGCTCCACCTCCAGCTTGTACCTGCAGGA

TCTGAGCGCCGCCGCCTCAGAGTGCATCGACCCCTCGGTGGTCTTCCCCTACCCTCTCAACGACAGCAGCTCGCCCA

AGTCCTGCGCCTCGCAAGACTCCAGCGCCTTCTCTCCGTCCTCGGATTCTCTGCTCTCCTCGACGGAGTCCTCCCCG

CAGGGCAGCCCCGAGCCCCTGGTGCTCCATGAGGAGACACCGCCCACCACCAGCAGCGACTCTGAGGAGGAACAAGA

AGATGAGGAAGAAATCGATGTTGTTTCTGTGGAAAAGAGGCAGGCTCCTGGCAAAAGGTCAGAGTCTGGATCACCTT

CTGCTGGAGGCCACAGCAAACCTCCTCACAGCCCACTGGTCCTCAAGAGGTGCCACGTCTCCACACATCAGCACAAC

TACGCAGCGCCTCCCTCCACTCGGAAGGACTATCCTGCTGCCAAGAGGGTCAAGTTGGACAGTGTCAGAGTCCTGAG

ACAGATCAGCAACAACCGAAAATGCACCAGCCCCAGGTCCTCGGACACCGAGGAGAATGTCAAGAGGCGAACACACA

ACGTCTTGGAGCGCCAGAGGAGGAACGAGCTAAAACGGAGCTTTTTTGCCCTGCGTGACCAGATCCCGGAGTTGGAA

AACAATGAAAAGGCCCCCAAGGTAGTTATCCTTAAAAAAGCCACAGCATACATCCTGTCCGTCCAAGCAGAGGAGCA

AAAGCTCATTTCTGAAGAGGACTTGTTGCGGAAACGACGAGAACAGTTGAAACACAAACTTGAACAGCTACGGAACT

CTTGTGCG

MYC Thr58Ala

ATGCCCCTCAACGTTAGCTTCACCAACAGGAACTATGACCTCGACTACGACTCGGTGCAGCCGTATTTCTACTGCGA

CGAGGAGGAGAACTTCTACCAGCAGCAGCAGCAGAGCGAGCTGCAGCCCCCGGCGCCCAGCGAGGATATCTGGAAGA

AATTCGAGCTGCTGCCCGCCCCGCCCCTGTCCCCTAGCCGCCGCTCCGGGCTCTGCTCGCCCTCCTACGTTGCGGTC

ACACCCTTCTCCCTTCGGGGAGACAACGACGGCGGTGGCGGGAGCTTCTCCACGGCCGACCAGCTGGAGATGGTGAC

CGAGCTGCTGGGAGGAGACATGGTGAACCAGAGTTTCATCTGCGACCCGGACGACGAGACCTTCATCAAAAACATCA

TCATCCAGGACTGTATGTGGAGCGGCTTCTCGGCCGCCGCCAAGCTCGTCTCAGAGAAGCTGGCCTCCTACCAGGCT

GCGCGCAAAGACAGCGGCAGCCCGAACCCCGCCCGCGGCCACAGCGTCTGCTCCACCTCCAGCTTGTACCTGCAGGA

TCTGAGCGCCGCCGCCTCAGAGTGCATCGACCCCTCGGTGGTCTTCCCCTACCCTCTCAACGACAGCAGCTCGCCCA

AGTCCTGCGCCTCGCAAGACTCCAGCGCCTTCTCTCCGTCCTCGGATTCTCTGCTCTCCTCGACGGAGTCCTCCCCG

CAGGGCAGCCCCGAGCCCCTGGTGCTCCATGAGGAGACACCGCCCACCACCAGCAGCGACTCTGAGGAGGAACAAGA

AGATGAGGAAGAAATCGATGTTGTTTCTGTGGAAAAGAGGCAGGCTCCTGGCAAAAGGTCAGAGTCTGGATCACCTT

CTGCTGGAGGCCACAGCAAACCTCCTCACAGCCCACTGGTCCTCAAGAGGTGCCACGTCTCCACACATCAGCACAAC

TACGCAGCGCCTCCCTCCACTCGGAAGGACTATCCTGCTGCCAAGAGGGTCAAGTTGGACAGTGTCAGAGTCCTGAG

ACAGATCAGCAACAACCGAAAATGCACCAGCCCCAGGTCCTCGGACACCGAGGAGAATGTCAAGAGGCGAACACACA

ACGTCTTGGAGCGCCAGAGGAGGAACGAGCTAAAACGGAGCTTTTTTGCCCTGCGTGACCAGATCCCGGAGTTGGAA

AACAATGAAAAGGCCCCCAAGGTAGTTATCCTTAAAAAAGCCACAGCATACATCCTGTCCGTCCAAGCAGAGGAGCA

AAAGCTCATTTCTGAAGAGGACTTGTTGCGGAAACGACGAGAACAGTTGAAACACAAACTTGAACAGCTACGGAACT

CTTGTGCG

MYC Ser62Ala

ATGCCCCTCAACGTTAGCTTCACCAACAGGAACTATGACCTCGACTACGACTCGGTGCAGCCGTATTTCTACTGCGA

CGAGGAGGAGAACTTCTACCAGCAGCAGCAGCAGAGCGAGCTGCAGCCCCCGGCGCCCAGCGAGGATATCTGGAAGA

AATTCGAGCTGCTGCCCACCCCGCCCCTGGCCCCTAGCCGCCGCTCCGGGCTCTGCTCGCCCTCCTACGTTGCGGTC

ACACCCTTCTCCCTTCGGGGAGACAACGACGGCGGTGGCGGGAGCTTCTCCACGGCCGACCAGCTGGAGATGGTGAC

CGAGCTGCTGGGAGGAGACATGGTGAACCAGAGTTTCATCTGCGACCCGGACGACGAGACCTTCATCAAAAACATCA

TCATCCAGGACTGTATGTGGAGCGGCTTCTCGGCCGCCGCCAAGCTCGTCTCAGAGAAGCTGGCCTCCTACCAGGCT

GCGCGCAAAGACAGCGGCAGCCCGAACCCCGCCCGCGGCCACAGCGTCTGCTCCACCTCCAGCTTGTACCTGCAGGA

TCTGAGCGCCGCCGCCTCAGAGTGCATCGACCCCTCGGTGGTCTTCCCCTACCCTCTCAACGACAGCAGCTCGCCCA

AGTCCTGCGCCTCGCAAGACTCCAGCGCCTTCTCTCCGTCCTCGGATTCTCTGCTCTCCTCGACGGAGTCCTCCCCG

CAGGGCAGCCCCGAGCCCCTGGTGCTCCATGAGGAGACACCGCCCACCACCAGCAGCGACTCTGAGGAGGAACAAGA

AGATGAGGAAGAAATCGATGTTGTTTCTGTGGAAAAGAGGCAGGCTCCTGGCAAAAGGTCAGAGTCTGGATCACCTT

CTGCTGGAGGCCACAGCAAACCTCCTCACAGCCCACTGGTCCTCAAGAGGTGCCACGTCTCCACACATCAGCACAAC

TACGCAGCGCCTCCCTCCACTCGGAAGGACTATCCTGCTGCCAAGAGGGTCAAGTTGGACAGTGTCAGAGTCCTGAG

ACAGATCAGCAACAACCGAAAATGCACCAGCCCCAGGTCCTCGGACACCGAGGAGAATGTCAAGAGGCGAACACACA

ACGTCTTGGAGCGCCAGAGGAGGAACGAGCTAAAACGGAGCTTTTTTGCCCTGCGTGACCAGATCCCGGAGTTGGAA

AACAATGAAAAGGCCCCCAAGGTAGTTATCCTTAAAAAAGCCACAGCATACATCCTGTCCGTCCAAGCAGAGGAGCA

AAAGCTCATTTCTGAAGAGGACTTGTTGCGGAAACGACGAGAACAGTTGAAACACAAACTTGAACAGCTACGGAACT

CTTGTGCG

Barcode Sequences

Gene	Barcode Sequence
mCherry	TTCTCATTTTGGAATATAAA
ASCL1	TTCGTGATAAAACTATCTGC
ASCL3	GTATATTAACTTGTACGCAG

ASCL4	GCGAATTATTACAGGCCTCT
ASCL5	TATCATGTTCTATTCAAACA
ATF7	CTTGTTCCACATTTTACTTT
CDX2	ATTGATTCGTTTTGAGAACT
CRX	GTTTTTCTCGAACTGACACA
ERG	CTCTGGTGAGTTCCACGCGA
ESRRG	CCTACAGACTTCTAAACGAT
ETV2	CACTGCCATTAGTCTTTTGT
FLI1	TTCGACGAGTTCGTAGCATG
FOXA1	GTTTTTTGAAGCGCCCCTTA
FOXA2	CTTTTTTAGCCGTCTTTAAC
FOXA3	AACGCGCGTTTTCAGCGCTA
FOXP1	GGGATGATTACTCTAGATCC
GATA1	TGTTACTTGCACCTCATTTC
GATA2	GCCACCCTTTTTCCTAAATC
GATA4	CAGTTCTTAGTATGATCTGA
GATA6	TCACTATTTGGAGTAATATA
GLI1	GTTTAGAAATTTAATACGAG
HAND2	ATGTTTTTTCTTAACAATTG
HNF1A	GCGTTTGGCGCTAATGTGAT
HNF1B	AAAGTTATAGATTAATGTAA
HNF4A	ACACTCTATATATATT
HOXA1	TCTCAACCCTTTTAATACGT
HOXA10	ACCTTCTGTCGAGTCTCACG
HOXA11	TTAGGGATAGCCGACGCACT
HOXB6	CTTTTAAAATTTGCGTAAAA
KLF4	AAGACTCTCATAATTACCTG

LHX3	CATGTCGGACGACAGCATCG
LMX1A	TATAAGAATAAAGTTATTTT
MEF2C	AAGTCTGCATCGCATTTATG
MESP1	AGGGCGTAAATTCTGCCTTT
MITF	TCCATAAAGCCACACTTTTC
MYC	CGACATACATGTAGCTTGCG
MYCL	AACCGTTAGTTAATCCCAA
MYCN	AACCATGGTTATTCTAGGCT
MYOD1	TTTCAAGGATAGTGTAACTG
MYOG	TTGTCTAAGTATAGTTAAGG
NEUROD1	AGGCTTATTTTATTGATATT
NEUROG1	AAACTTTTCCGGTCAGGTTC
NEUROG3	TGGCCTATTATAATCGGTGA
NRL	TATTATAGTCTTTCTTAAAT
ONECUT1	CCATTCATTTATTTTTAATA
OTX2	AGTTTTCCTCATAATTCATCA
PAX7	AACACTAAAGTCACACGTTA
POU1F1	TATTCTCCATCATTTTTGGA
POU5F1	CAGACAGTGCGGCCTGCTCG
RUNX1	CTACCGTGTCCAACTCTTAG
SIX1	ACTGCGACCCATGTCCACTA
SIX2	ACTGTCACCTATCGCAGCAC
SNAI2	TCATAGTATTTAATTGAGAA
SOX10	AATCAATCCTCATATAGCAA
SOX2	CCTGTTGACACAAATGTGCT
SOX3	TGAGTCACCCTTTATAAAAA
SPI1	GTTTACTATTTCACTCCCGC

SPIB	GATCTTTAGTAACAGATCAA
SPIC	GTTGGCCACGGAGCTTAAGG
SRY	AGGGAGAACGGTGAGGGGAG
TBX5	AGTTTGCTCTAGGAATTTTC
TFAP2C	GGGGGGGGCGGGGGAGGCTG
KLF1	TAGAATGTTAACGCAACCTT
KLF2	AGACATTTTGGGTAAAGTCA
KLF3	TCGCAGAGTCCACACAAGTC
KLF5	CAACGTCTACCCGCTACGTT
KLF6	ATCCATCAGATACGATCAAT
KLF7	TTCCGATGCGTACGGTCGGA
KLF8	TTAACACAGTACCGAATGAC
KLF9	GGTTAAAAATAGTTTATATT
KLF10	TTTTGGGGAGGTGTTGGGGT
KLF11	TGTGATGGGAGGGTGGGGCG
KLF12	ATTGAGGGGGCGTGCCGGGT
KLF13	AAGTACATTTCATTTTACCA
KLF14	CTCTTATTTCTCCCGTATCT
KLF15	TTGTTGACCCGAATTATTCG
KLF16	AAGTACTGTTTTAAACAAGT
KLF17	GTAACGATAGCATCTTCTGA
MYC ΔMBI	CTCCGCGAATCGGTTTCGTT
c-MYC ΔMBII	ACATTCACTACTCGTTTATT
MYC ΔNLS	ATTAATTCTTTACATTTGAT
MYC Δb	GCTATACTATTGCGGGAGTG
MYC ΔHLH	TTCAACCGTTCTATTGCTGC
MYC ΔLZ	TGTGTTTTTTGACCACAAAG

MYC ΔNTD	TCGAAACTCAAAACCTTGGA
MYC ΔCTD	GTCGAGGATATCTTCTAATT
MYC Glu39Ala	AATTATTATTTGACTTAAAA
MYC Thr58Ala	CTGCGAACGTTTACCGAATT
MYC Ser62Ala	TGTTAGTTGGGGAGGGTCTT

DNA sequences from Chapter 3

EF1α + Kozak Sequence + mCherry + P2A + Hygromycin Resistance + Barcode

cgtgaggctccggtgcccgtcagtgggcagagcgcacatcgcccacagtccccgagaagttgggggggaggggtcggc

aattgaaccggtgcctagagaaggtggcgcggggtaaactgggaaagtgatgtcgtgtactggctccgccttttttcc

cgagggtgggggagaaccgtatataagtgcagtagtcgccgtgaacgttcttttcgcaacgggtttgccgccagaa

cacaggtaagtgccgtgtgtggttcccgcgggcctggcctctttacgggttatggcccttgcgtgccttgaattact

tccacctggctgcagtacgtgattcttgatcccgagcttcgggttggaagtgggtgggagagttcgaggccttgcgc

ttaaggagccccttcgcctcgtgcttgagttgaggcctggcctgggcgctggggccgccgcgtgcgaatctggtggc

accttcgcgcctgtctcgctgctttcgataagtctctagccatttaaaatttttgatgacctgctgcgacgcttttt

ttctggcaagatagtcttgtaaatgcgggccaagatctgcacactggtatttcggtttttggggccgcgggcggcga

cggggcccgtgcgtcccagcgcacatgttcggcgaggcggggcctgcgagcgcggccaccgagaatcggacgggggt

agtctcaagctggccggcctgctctggtgcctggcctcgcgccgccgtgtatcgccccgccctgggcggcaaggctg

gcccggtcggcaccagttgcgtgagcggaaagatggccgcttcccggccctgctgcagggagctcaaaatggaggac

gcggcgctcgggagagcgggcgggtgagtcacccacacaaaggaaaagggcctttccgtcctcagccgtcgcttcat

gtgactccacggagtaccgggcgccgtccaggcacctcgattagttctcgagcttttggagtacgtcgtctttaggt

tgggggggaggggttttatgcgatggagtttccccacactgagtgggtggagactgaagttaggccagcttggcactt

gatgtaattctccttggaatttgccctttttgagtttggatcttggttcattctcaagcctcagacagtggttcaaa

gtttttttcttccatttcaggtgtcgtgaggaattagcttggtactaatacgactcactatagggagacccaagctg

gctagttaagcttgatatcgaattcctgcagcccgggggatccaccggtcgccaccatggtgagcaagggcgaggag

gataacatggccatcatcaaggagttcatgcgcttcaaggtgcacatggagggctccgtgaacggccacgagttcga

gatcgagggcgagggcgagggccgcccctacgagggcacccagaccgccaagctgaaggtgaccaagggtggccccc

tgcccttcgcctgggacatcctgtcccctcagttcatgtacggctccaaggcctacgtgaagcaccccgccgacatc

cccgactacttgaagctgtccttccccgagggcttcaagtgggagcgcgtgatgaacttcgaggacggcggcgtggt

gaccgtgacccaggactcctccctgcaggacggcgagttcatctacaaggtgaagctgcgcggcaccaacttcccct

ccgacggccccgtaatgcagaagaagaccatgggctgggaggcctcctccgagcggatgtaccccgaggacggcgcc

ctgaagggcgagatcaagcagaggctgaagctgaaggacggcggccactacgacgctgaggtcaagaccacctacaa

ggccaagaagcccgtgcagctgcccggcgcctacaacgtcaacatcaagttggacatcacctcccacaacgaggact

acaccatcgtggaacagtacgaacgcgccgagggccgccactccaccggcggcatggacgagctgtacaagGGATCC

GGCGCAACAAACTTCTCTCTGCTGAAACAAGCCGGAGATGTCGAAGAGAATCCTGGACCGATGAAAAAGCCTGAACT

CACCGCGACGTCTGTCGAGAAGTTTCTGATCGAAAAGTTCGACAGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCG

AAGAATCTCGTGCTTTCAGCTTCGATGTAGGAGGGCGTGGATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTTC

TACAAAGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCGCTCCCGATTCCGGAAGTGCTTGACATTGGGGAATT

CAGCGAGAGCCTGACCTATTGCATCTCCCGCCGTGCACAGGGTGTCACGTTGCAAGACCTGCCTGAAACCGAACTGC

CCGCTGTTCTGCAGCCGGTCGCGGAGGCCATGGATGCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGGC

CCATTCGGACCGCAAGGAATCGGTCAATACACTACATGGCGTGATTTCATATGCGCGATTGCTGATCCCCATGTGTA

TCACTGGCAAACTGTGATGGACGACACCGTCAGTGCGTCCGTCGCGCAGGCTCTCGATGAGCTGATGCTTTGGGCCG

AGGACTGCCCCGAAGTCCGGCACCTCGTGCACGCGGATTTCGGCTCCAACAATGTCCTGACGGACAATGGCCGCATA

ACAGCGGTCATTGACTGGAGCGAGGCGATGTTCGGGGATTCCCAATACGAGGTCGCCAACATCTTCTTCTGGAGGCC

GTGGTTGGCTTGTATGGAGCAGCAGACGCGCTACTTCGAGCGGAGGCATCCGGAGCTTGCAGGATCGCCGCGGCTCC

GGGCGTATATGCTCCGCATTGGTCTTGACCAACTCTATCAGAGCTTGGTTGACGGCAATTTCGATGATGCAGCTTGG

GCGCAGGGTCGATGCGACGCAATCGTCCGATCCGGAGCCGGGACTGTCGGGCGTACACAAATCGCCCGCAGAAGCGC

GGCCGTCTGGACCGATGGCTGTGTAGAAGTACTCGCCGATAGTGGAAACCGACGCCCCAGCACTCGTCCGAGGGCAA

AGGAATAAGTTAGAACTATTTCCTGGCTGTTACGCGNNNNNNNNNNNNNNNNNNNNCTACTGACGGGCTTGTTAAAC

AGTG

KRAS (G12V)

atgactgaatataaacttgtggtagttggagctgttggcgtaggcaagagtgccttgacgatacagctaattcagaa

tcattttgtggacgaatatgatccaacaatagaggattcctacaggaagcaagtagtaattgatggagaaacctgtc

tcttggatattctcgacacagcaggtcaagaggagtacagtgcaatgagggaccagtacatgaggactggggagggc

tttctttgtgtatttgccataaataatactaaatcatttgaagatattcaccattatagagaacaaattaaaagagt

taaggactctgaagatgtacctatggtcctagtaggaaataaatgtgatttgccttccagaacagtagacacaaaac

aggctcaggacttagcaagaagttatggaattccttttattgaaacatcagcaaagacaagacagggtgttgatgat

gccttctatacattagttcgagaaattcgaaaacataaagaaaagatgagcaaagatggtaaaaagaagaaaaagaa

gtcaaagacaaagtgtgtaattatg

HRAS (G12V)

atgacggaatataagctggtggtggtgggcgccgtcggtgtgggcaagagtgcgctgaccatccagctgatccagaa

ccattttgtggacgaatacgaccccactatagaggattcctaccggaagcaggtggtcattgatggggagacgtgcc

tgttggacatcctggataccgccggccaggaggagtacagcgccatgcgggaccagtacatgcgcaccggggagggc

ttcctgtgtgtgtttgccatcaacaacaccaagtcttttgaggacatccaccagtacagggagcagatcaaacgggt

gaaggactcggatgacgtgcccatggtgctggtggggaacaagtgtgacctggctgcacgcactgtggaatctcggc

aggctcaggacctcgcccgaagctacggcatcccctacatcgagacctcggccaagacccggcagggagtggaggat

gccttctacacgttggtgcgtgagatccggcagcacaagctgcggaagctgaaccctcctgatgagagtggccccgg

ctgcatgagctgcaagtgtgtgctctcc

MEK1 (S218D, S222D)

atgcccaagaagaagccgacgcccatccagctgaacccggccccgacggctctgcagttaacgggaccagctctgc

ggagaccaacttggaggccttgcagaagaagctggaggagctagagcttgatgagcagcagcgaaagcgccttgagg

cctttcttacccagaagcagaaggtgggagaactgaaggatgacgactttgagaagatcagtgagctgggggctggc

aatggcggtgtggtgttcaaggtctcccacaagccttctggcctggtcatggccagaaagctaattcatctggagat

caaacccgcaatccggaaccagatcataagggagctgcaggttctgcatgagtgcaactctccgtacatcgtgggct

tctatggtgcgttctacagcgatggcgagatcagtatctgcatggagcacatggatggaggttctctggatcaagtc

ctgaagaaagctggaagaattcctgaacaaattttaggaaaagttagcattgctgtaataaaaggcctgacatatct

gagggagaagcacaagatcatgcacagagatgtcaagccctccaacatcctagtcaactcccgtggggagatcaagc

tctgtgactttggggtcagcgggcagctcatcgacgacatggccaacgacttcgtgggcacaaggtcctacatgtcg

ccagaaagactccaggggactcattactctgtgcagtcagacatctggagcatgggactgtctctggtagagatggc

ggttgggaggtatcccatccctcctccagatgccaaggagctggagctgatgtttgggtgccaggtggaaggagatg

cggctgagaccccacccaggccaaggacccccgggaggcccttagctcatacggaatggacagccgacctcccatg

gcaatttttgagttgttggattacatagtcaacgagcctcctccaaaactgcccagtggagtgttcagtctggaatt

tcaagattttgtgaataaatgcttaataaaaaaccccgcagagagagcagatttgaagcaactcatggttcatgctt

ttatcaagagatctgatgctgaggaagtggattttgcaggttggctctgctccaccatcggccttaaccagcccagc

acaccaacccatgctgctggcgtc

myr-FLAG-PIK3CA

atggggtcttcaaaatctaaaccaaaggacccccagccagcgccggcgcaggatccgaggttaccttgactacaaaga

cgatgacgacaagcaattgagcgacgagatccgaacaatgcctccaagaccatcatcaggtgaactgtggggcatcc

acttgatgcccccaagaatcctagtagaatgtttactaccaaatgggatgatagtgactttagaatgcctccgtgag

gctacgttaataacgataaagcatgaactatttaaagaagcaagaaaataccctctccatcaacttcttcaagatga

atcttcttacattttcgtaagtgttacccaagaagcagaaagggaagaatttttgatgaaacaagacgactttgtg

accttcggctttttcaaccctttttaaaagtaattgaaccagtaggcaaccgtgaagaaaagatcctcaatcgagaa

attggttttgctatcggcatgccagtgtgtgaattcgatatggttaaagatccagaagtacaggacttccgaagaaa

tattctcaatgtttgtaaagaagctgtggatcttagggatcttaattcacctcatagtagagcaatgtatgtttatc

ctccaaatgtagaatcttcaccagaactgccaaagcacatatataataaattggataaagggcaaataatagtggtg

atttgggtaatagtttctccaaataatgacaaacagaagtatactctgaaaatcaaccatgactgtgtgccagaaca

agtaattgctgaagcaatcaggaaaaaaactcgaagtatgttgctatcatctgaacaactaaaactctgtgttttag

aatatcagggcaagtatattttaaaagtgtgtggatgtgatgaatacttcctagaaaaatatcctctgagtcagtat

aagtatataagaagctgtataatgcttgggaggatgcccaatttgatgctgatggctaaagaaagcctctattctca

actgccaatggactgttttacaatgccatcatattccagacgcatctccacagctacgccatatatgaatggagaaa

catctacaaaatccctttgggttataaatagtgcactcagaataaaaattctttgtgcaacctatgtgaatgtaaat

attcgagacattgacaagatttatgttcgaacaggtatctaccatggaggagaacccttatgtgataatgtgaacac

tcaaagagtaccttgttccaatcccaggtggaatgaatggctgaattacgatatatacattcctgatcttcctcgtg

ctgctcgactttgcctttccatttgttctgttaaaggccgaaagggtgctaaagaggaacactgtccattggcctgg

ggaaatataaacttgtttgattacacagatactctagtatctggaaaaatggctttgaatctttggccagtacctca

tggactagaagatttgctgaaccctattggtgttactggatcaaatccaaataaagaaactccatgtttagagttgg

agtttgactggttcagcagtgtggtaaagtttccagatatgtcagtgattgaagagcatgccaattggtctgtatcc

cgtgaagcaggatttagttattcccatgcaggactgagtaacagactagctagagacaatgaattaagagaaaatga

taaagaacagctccgagcaatttgtacacgagatcctctatctgaaatcactgagcaagagaaagattttctgtgga

gccacagacactattgtgtaactatccccgaaattctacccaaattgcttctgtctgttaaatggaactctagagat

gaagtagctcagatgtactgcttggtaaaagattggcctccaatcaagcctgaacaggctatggagcttctggactg

caattacccagatcctatggttcgaggttttgctgttcggtgcttagaaaaatatttaacagatgacaaactttctc

agtacctaattcagctagtacaggtactaaaatatgaacagtatttggataacctgcttgtgagatttttactcaaa

aaagcgttaactaatcaaaggatcggtcacttttttcttttggcatttaaaatctgagatgcacaataaaacagttag

tcagaggtttggcctgcttttggagtcctattgccgtgcatgtgggatgtatctgaagcaccttaataggcaagttg

aggctatggaaaagctcattaacttgactgacattctcaaacaagagaagaaggatgaaacacaaaaggtacagatg

aagtttttagttgagcaaatgcggcgaccagatttcatggatgctctccagggctttctgtctcctctaaaccctgc

tcatcagctgggaaatctcaggcttgaagagtgtcgaattatgtcttctgcaaaaaggccactgtggttgaattggg

agaacccagacatcatgtcagaattactctttcagaacaatgagatcatctttaaaaatggggatgatttacggcaa

gatatgctaacccttcagattattcgcattatggaaaatatctggcaaaatcaaggtcttgatcttcgaatgttacc

ttatggatgtctgtcaatcggtgactgtgtgggacttatcgaggtggtgagaaattctcacactataatgcagattc

agtgtaaaggaggcctgaaaggtgcactgcagtttaacagccacacactccatcagtggctcaaagacaagaacaag

ggggaaatatatgatgcggccatcgatttgtttacacgatcatgtgctggatattgtgttgccaccttcattttggg

aattggagatcgtcacaatagtaatatcatggttaaagatgatggacaactgtttcatatagattttggacactttt

tggatcacaagaagaaaaaatttggttataaacgagagcgcgtgccgtttgttttgacacaagatttcttaatagtg

attagtaaaggagcccaagaatgcacaaagacaagagaatttgagaggtttcaggagatgtgttacaaggcttatct

agctattcggcagcatgccaatctcttcataaatcttttctcaatgatgcttggctctggaatgccagaactgcaat

cttttgatgatattgcatacattcgaaagaccctagctttagataaaactgagcaagaggctttggagtatttcatg

aaacaaatgaatgatgcacaccatggtggctggacaacaaaaatggattggatcttccacacaattaagcagcatgc

tttgaac

myr-FLAG-AKT1

atggggtcttcaaaatctaaaccaaaggaccccagccagcgccggcgcaggatccgaggttaccttgactacaaaga

cgatgacgacaagcaattgacaagtttgtacaaaaaagttggcatgagcgacgtggctattgtgaaggagggttggc

tgcacaaacgaggggagtacatcaagacctggcggccacgctacttcctcctcaagaatgatggcaccttcattggc

tacaaggagcggccgcaggatgtggaccaacgtgaggctcccctcaacaacttctctgtggcgcagtgccagctgat

gaagacggagcggccccggcccaacaccttcatcatccgctgcctgcagtggaccactgtcatcgaacgcaccttcc

atgtggagactcctgaggagcgggaggagtggacaaccgccatccagactgtggctgacggcctcaagaagcaggag

gaggaggagatggacttccggtcgggctcacccagtgacaactcaggggctgaagagatggaggtgtccctggccaa

gcccaagcaccgcgtgaccatgaacgagtttgagtacctgaagctgctgggcaagggcactttcggcaaggtgatcc

tggtgaaggagaaggccacaggccgctactacgccatgaagatcctcaagaaggaagtcatcgtggccaaggacgag

gtggcccacacactcaccgagaaccgcgtcctgcagaactccaggcacccctcctcacagccctgaagtactcttt

ccagacccacgaccgcctctgctttgtcatggagtacgccaacgggggcgagctgttcttccacctgtcccgggagc

gtgtgttctccgaggaccgggcccgcttctatggcgctgagattgtgtcagccctggactacctgcactcggagaag

aacgtggtgtaccgggacctcaagctggagaacctcatgctggacaaggacgggcacattaagatcacagacttcgg

gctgtgcaaggaggggatcaaggacggtgccaccatgaagacctttgcggcacacctgagtacctggcccccgagg

tgctggaggacaatgactacggccgtgcagtggactggtgggggctgggcgtggtcatgtacgagatgatgtgcggt

cgcctgcccttctacaaccaggaccatgagaagctttttgagctcatcctcatggaggagatccgcttcccgcgcac

gcttggtcccgaggccaagtccttgctttcagggctgctcaagaaggaccccaagcagaggcttggcgggggctccg

aggacgccaaggagatcatgcagcatcgcttctttgccggtatcgtgtggcagcacgtgtacgagaagaagctcagc

ccacccttcaagccccaggtcacgtcggagactgacaccaggtattttgatgaggagttcacggcccagatgatcac

catcacaccacctgaccaagatgacagcatggagtgtgtggacagcgagcgcaggccccacttcccccagttctcct

actcggccagcggcacggcc

FLAG-Rheb (Q64L)

atggactacaaagacgatgacaaggcggccgcgccgcagtccaagtcccggaagatcgcgatcctgggctaccggtc

tgtggggaaatcctcattgacgattcaatttgttgaaggccaatttgtggactcctacgatccaaccatagaaaaca

cttttacaaagttgatcacagtaaatggacaagaatatcatcttcaacttgtagacacagccgggctagatgaatat

tctatctttcctcagacatactccatagatattaatggctatattcttgtgtattctgttacatcaatcaaaagttt

tgaagtgattaaagttatccatggcaaattgttggatatggtggggaaagtacaaatacctattatgttggttggga

ataagaaagacctgcatatggaaagggtgatcagttatgaagaagggaaagctttggcagaatcttggaatgcagct

tttttggaatcttctgctaaagaaaatcagactgctgtggatgtttttcgaaggataattttggaggcagaaaaaat

ggacggggcagcttcacaaggcaagtcttcatgctcggtgatg

IKKα (S176E, S180E)

atggagcggcccccggggctgcggccgggcgcgggcgggccctgggagatgcgggagcggctgggcaccggcggctt

cgggaacgtctgtctgtaccagcatcgggaacttgatctcaaaatagcaattaagtcttgtcgcctagagctaagta

ccaaaaacagagaacgatggtgccatgaaatccagattatgaagaagttgaaccatgccaatgttgtaaaggcctgt

gatgttcctgaagaattgaatattttgattcatgatgtgcctcttctagcaatggaatactgttctggaggagatct

ccgaaagctgctcaacaaaccagaaaattgttgtggacttaaagaaagccagatactttctttactaagtgatatag

ggtctgggattcgatatttgcatgaaaacaaaattatacatcgagatctaaaacctgaaaacatagttcttcaggat

gttggtggaaagataatacataaaataattgatctgggatatgccaaagatgttgatcaaggagaactgtgtacaga

atttgtgggaacactgcagtatctggccccagagctctttgagaataagccttacacagccactgttgattattgga

gctttgggaccatggtatttgaatgtattgctggatataggccttttttgcatcatctgcagccatttacctggcat

gagaagattaagaagaaggatccaaagtgtatatttgcatgtgaagagatgtcaggagaagttcggtttagtagcca

tttacctcaaccaaatagcctttgtagtttagtagtagaacccatggaaaactggctacagttgatgttgaattggg

accctcagcagagaggaggacctgttgaccttactttgaagcagccaagatgttttgtattaatggatcacattttg

aatttgaagatagtacacatcctaaatatgacttctgcaaagataatttcttttctgttaccacctgatgaaagtct

tcattcactacagtctcgtattgagcgtgaaactggaataaatactggttctcaagaacttctttcagagacaggaa

tttctctggatcctcggaaaccagcctctcaatgtgttctagatggagttagaggctgtgatagctatatggtttat

ttgtttgataaaagtaaaactgtatatgaagggccatttgcttccagaagtttatctgattgtgtaaattatattgt

acaggacagcaaaatacagcttccaattatacagctgcgtaaagtgtgggctgaagcagtgcactatgtgtctggac

taaaagaagactatagcaggctctttcagggacaaagggcagcaatgttaagtcttcttagatataatgctaactta

acaaaaatgaagaacactttgatctcagcatcacaacaactgaaagctaaattggagtttttcacaaaagcattca

gcttgacttggagagatacagcgagcagatgacgtatgggatatcttcagaaaaaatgctaaaagcatggaaagaaa

tggaagaaaaggccatccactatgctgaggttggtgtcattggatacctggaggatcagattatgtctttgcatgct

gaaatcatggagctacagaagagcccctatggaagacgtcagggagacttgatggaatctctggaacagcgtgccat

tgatctatataagcagttaaaacacagaccttcagatcactcctacagtgacagcacagagatggtgaaaatcattg

tgcacactgtgcagagtcaggaccgtgtgctcaaggagctgtttggtcatttgagcaagttgttgggctgtaagcag

aagattattgatctactccctaaggtggaagtggccctcagtaatatcaaagaagctgacaatactgtcatgttcat

gcagggaaaaaggcagaaagaaatatggcatctccttaaaattgcctgtacacagagttctgcccggtcccttgtag

gatccagtctagaaggtgcagtaacccctcagacatcagcatggctgcccccgacttcagcagaacatgatcattct

ctgtcatgtgtggtaactcctcaagatggggagacttcagcacaaatgatagaagaaaatttgaactgccttggcca

tttaagcactattattcatgaggcaaatgaggaacagggcaatagtatgatgaatcttgattggagttggttaacag

aa

FLAG-IKKβ (S177E, S181E)

atggactacaaagacgatgacgacaagcttgcggccgcaatgagctggtcaccttccctgacaacgcagacatgtgg

ggcctgggaaatgaaagagcgccttgggacaggggatttggaaatgtcatccgatggcacaatcaggaaacaggtg

agcagattgccatcaagcagtgccggcaggagctcagcccccggaaccgagagcggtggtgcctggagatccagatc

atgagaaggctgacccaccccaatgtggtggctgcccgagatgtccctgaggggatgcagaacttggcgcccaatga

cctgcccctgctggccatggagtactgccaaggaggagatctccggaagtacctgaaccagtttgagaactgctgtg

gtctgcgggaaggtgccatcctcaccttgctgagtgacattgcctctgcgcttagataccttcatgaaaacagaatc

atccatcgggatctaaagccagaaaacatcgtcctgcagcaaggagaacagaggttaatacacaaaattattgacct

aggatatgccaaggagctggatcagggcgaactttgcacagaattcgtggggaccctgcagtacctggccccagagc

tactggagcagcagaagtacacagtgaccgtcgactactggagcttcggcaccctggcctttgagtgcatcacgggc

ttccggcccttcctccccaactggcagcccgtgcagtggcattcaaaagtgcggcagaagagtgaggtggacattgt

tgttagcgaagacttgaatggaacggtgaagttttcaagctctttaccctaccccaataatcttaacagtgtcctgg

ctgagcgactggagaagtggctgcaactgatgctgatgtggcacccccgacagaggggcacggatcccacgtatggg

cccaatggctgcttcaaggccctggatgacatcttaaacttaaagctggttcatatcttgaacatggtcacgggcac

catccacacctaccctgtgacagaggatgagagtctgcagagcttgaaggccagaatccaacaggacacgggcatcc

cagaggaggaccaggagctgctgcaggaagcgggcctggcgttgatccccgataagcctgccactcagtgtatttca

gacggcaagttaaatgagggccacacattggacatggatcttgtttttctctttgacaacagtaaaatcacctatga

gactcagatctccccacggccccaacctgaaagtgtcagctgtatccttcaagagcccaagaggaatctcgccttct

tccagctgaggaaggtgtggggccaggtctggcacagcatccagaccctgaaggaagattgcaaccggctgcagcag

ggacagcgagccgccatgatgaatctcctccgaaacaacagctgcctctccaaaatgaagaattccatggcttccat

gtctcagcagctcaaggccaagttggatttcttcaaaaccagcatccagattgacctggagaagtacagcgagcaaa

ccgagtttgggatcacatcagataaactgctgctggcctggagggaaatggagcaggctgtggagctctgtgggcgg

gagaacgaagtgaaactcctggtagaacggatgatggctctgcagaccgacattgtggacttacagaggagccccat

gggccggaagcaggggggaacgctggacgacctagaggagcaagcaagggagctgtacaggagactaagggaaaaac

ctcgagaccagcgaactgagggtgacagtcaggaaatggtacggctgctgcttcaggcaattcagagcttcgagaag

aaagtgcgagtgatctatacgcagctcagtaaaactgtggtttgcaagcagaaggcgctggaactgttgcccaaggt

ggaagaggtggtgagcttaatgaatgaggatgagaagactgttgtccggctgcaggagaagcggcagaaggagctct

ggaatctcctgaagattgcttgtagcaaggtccgtggtcctgtcagtggaagcccggatagcatgaatgcctctcga

cttagccagcctgggcagctgatgtctcagccctccacggcctccaacagcttacctgagccagccaagaagagtga

agaactggtggctgaagcacataacctctgcaccctgctagaaaatgccatacaggacactgtgagggaacaagacc

agagtttcacggccctagactggagctggttacagacggaagaagaagagcacagctgcctggagcaggcctca

STAT3 (A662C, N664C, V667L)

atggcccaatggaatcagctacagcagcttgacacacggtacctggagcagctccatcagctctacagtgacagctt

cccaatggagctgcggcagtttctggccccttggattgagagtcaagattgggcatatgcggccagcaaagaatcac

atgccactttggtgtttcataatctcctgggagagattgaccagcagtatagccgcttcctgcaagagtcgaatgtt

ctctatcagcacaatctacgaagaatcaagcagtttcttcagagcaggtatcttgagaagccaatggagattgcccg

gattgtggcccggtgcctgtgggaagaatcacgccttctacagactgcagccactgcggcccagcaaggggccagg

ccaaccaccccacagcagccgtggtgacggagaagcagcagatgctggagcagcaccttcaggatgtccggaagaga

gtgcaggatctagaacagaaaatgaaagtggtagagaatctccaggatgactttgatttcaactataaaaccctcaa

gagtcaaggagacatgcaagatctgaatggaaacaaccagtcagtgaccaggcagaagatgcagcagctggaacaga

tgctcactgcgctggaccagatgcggagaagcatcgtgagtgagctggcggggcttttgtcagcgatggagtacgtg

cagaaaactctcacggacgaggagctggctgactggaagaggcggcaacagattgcctgcattggaggcccgcccaa

catctgcctagatcggctagaaaactggataacgtcattagcagaatctcaacttcagacccgtcaacaaattaaga

aactggaggagttgcagcaaaaagtttcctacaaaggggaccccattgtacagcaccggccgatgctggaggagaga

atcgtggagctgtttagaaacttaatgaaaagtgcctttgtggtggagcggcagccctgcatgcccatgcatcctga

ccggcccctcgtcatcaagaccggcgtccagttcactactaaagtcaggttgctggtcaaattccctgagttgaatt

atcagcttaaaattaaagtgtgcattgacaaagactctggggacgttgcagctctcagaggatcccggaaatttaac

attctgggcacaaacacaaaagtgatgaacatggaagaatccaacaacggcagcctctctgcagaattcaaacactt

gaccctgagggagcagagatgtgggaatgggggccgagccaattgtgatgcttccctgattgtgactgaggagctgc

acctgatcacctttgagaccgaggtgtatcaccaaggcctcaagattgacctagagacccactccttgccagttgtg

gtgatctccaacatctgtcagatgccaaatgcctgggcgtccatcctgtggtacaacatgctgaccaacaatcccaa

gaatgtaaacttttttaccaagcccccaattggaacctgggatcaagtggccgaggtcctgagctggcagttctcct

ccaccaccaagcgaggactgagcatcgagcagctgactacactggcagagaaactcttgggacctggtgtgaattat

tcagggtgtcagatcacatgggctaaattttgcaaagaaaacatggctggcaagggcttctccttctgggtctggct

ggacaatatcattgaccttgtgaaaaagtacatcctggccctttggaacgaagggtacatcatgggctttatcagta

aggagcgggagcgggccatcttgagcactaagcctccaggcaccttcctgctaagattcagtgaaagcagcaaagaa

ggaggcgtcactttcacttgggtggagaaggacatcagcggtaagacccagatccagtccgtggaaccatacacaaa

gcagcagctgaacaacatgtcatttgctgaaatcatcatgggctataagatcatggattgtacctgtatcctgttgt

ctccactggtctatctctatcctgacattcccaaggaggaggcattcggaaagtattgtcggccagagagccaggag

catcctgaagctgacccaggtagcgctgccccatacctgaagaccaagtttatctgtgtgacaccaacgacctgcag

caataccattgacctgccgatgtccccccgcactttagattcattgatgcagtttggaaataatggtgaaggtgctg

aaccctcagcaggagggcagtttgagtccctcacctttgacatggagttgacctcggagtgcgctacctccccatg

<u>β-catenin (S33A, S37A, T41A, S45A)</u>

atggctactcaagctgacctgatggagttggacatggccatggagccggacagaaaagctgctgtcagccactggca

gcagcagtcttacttggatgctggaatccatgctggtgccaccgccacagctcctgccctgagtggcaagggcaacc

ctgaggaagaagatgttgacacctcccaagtcctttatgaatgggagcaaggcttttcccagtccttcacgcaagag

caagtagctgatattgacgggcagtatgcaatgactagggctcagagggtccgagctgccatgttccctgagacgct

agatgagggcatgcagatcccatccacgcagtttgacgctgctcatcccactaatgtccagcgcttggctgaaccat

cacagatgttgaaacatgcagttgtcaatttgattaactatcaggatgacgcggaacttgccacacgtgcaattcct

gagctgacaaaactgctaaacgatgaggaccaggtggtagttaataaagctgctgttatggtccatcagctttccaa

aaaggaagcttccagacatgccatcatgcgctcccctcagatggtgtctgccattgtacgcaccatgcagaatacaa

atgatgtagagacagctcgttgtactgctgggactctgcacaacctttctcaccaccgcgagggcttgctggccatc

tttaagtctggtggcatcccagcgctggtgaaaatgcttgggtcaccagtggattctgtactgttctacgccatcac

gacactgcataatctcctgctccatcaggaaggagctaaaatggcagtgcgcctagctggtggactgcagaaaatgg

ttgctttgctcaacaaaacaaacgtgaaattcttggctattacaacagactgccttcagatcttagcttatggcaat

caagagagcaagctcatcattctggccagtggtggaccccaagccttagtaaacataatgaggacctacacttatga

gaagcttctgtggaccacaagcagagtgctgaaagtgctgtctgtctgctctagcaacaagccggccattgtagaag

ctggtgggatgcaggcactggggcttcatctgacagacccaagtcagcgacttgttcaaaactgtctttggactctc

agaaacctttcagatgcagcgactaagcaggaagggatggaaggcctccttgggactctagtgcagcttctgggttc

cgatgatataaatgtggtcacctgtgcagctggaattctctctaacctcacttgcaataattacaaaaacaagatga

tggtgtgccaagtgggtggcatagaggctcttgtacgcaccgtccttcgtgctggtgacagggaagacatcactgag

cctgccatctgtgctcttcgtcatctgaccagccggcatcaggaagccgagatggcccagaatgccgttcgccttca

ttatggactgcctgttgtggttaaactcctgcacccaccatcccactggcctctgataaaggcaactgttggattga

ttcgaaaccttgccctttgcccagcaaatcatgcgcctttgcgggaacagggtgctattccacgactagttcagctg

cttgtacgagcacatcaggacacccaacggcgcacctccatgggtggaacgcagcagcagtttgtggagggcgtgcg

catggaggagatagtagaagggtgtactggagctctccacatccttgctcgggacgttcacaaccggattgtaatcc

gaggactcaataccattccattgtttgtgcagttgctttattctcccattgaaaatatccaaagagtagctgcaggg

gtcctctgtgaacttgctcaggacaaggaggctgcagaggccattgaagctgagggagccacagctcccctgacaga

gttactccactccaggaatgaaggcgtggcaacatacgcagctgctgtcctattccgaatgtctgaggacaagccac

aggattacaagaagcggctttcagtcgagctgaccagttccctcttcaggacagagccaatggcttggaatgagact

gcagatcttggactggacattggtgcccagggagaagcccttggatatcgccaggatgatcccagctaccgttcttt

tcactctggtggatacggccaggatgccttggggatggaccctatgatggagcatgagatgggtggccaccaccctg

gtgctgactatccagttgatgggctgcctgatctgggacacgcccaggacctcatggatgggctgcccccaggtgat

agcaatcagctggcctggtttgatactgacctg

GSK3β (K85A)

atgtcagggcggcccagaaccacctcctttgcggagagctgcaagccggtgcagcagccttcagcttttggcagcat

gaaagttagcagagacaaggacggcagcaaggtgacaacagtggtggcaactcctgggcagggtccagacaggccac

aagaagtcagctatacagacactaaagtgattggaaatggatcatttggtgtggtatatcaagccaaactttgtgat

tcaggagaactggtcgccatcgcgaaagtattgcaggacaagagatttaagaatcgagagctccagatcatgagaaa

gctagatcactgtaacatagtccgattgcgttatttcttctactccagtggtgagaagaaagatgaggtctatctta

atctggtgctggactatgttccggaaacagtatacagagttgccagacactatagtcgagccaaacagacgctccct

gtgatttatgtcaagttgtatatgtatcagctgttccgaagtttagcctatatccattcctttggaatctgccatcg

ggatattaaaccgcagaacctcttgttggatcctgatactgctgtattaaaactctgtgactttggaagtgcaaagc

agctggtccgaggagaacccaatgtttcgtatatctgttctcggtactatagggcaccagagttgatctttggagcc

actgattatacctctagtatagatgtatggtctgctggctgtgtgttggctgagctgttactaggacaaccaatatt

tccaggggatagtggtgtggatcagttggtagaaataatcaaggtcctgggaactccaacaagggagcaaatcagag

aaatgaacccaaactacacagaatttaaattccctcaaattaaggcacatccttggactaaggtcttccgaccccga

actccaccggaggcaattgcactgtgtagccgtctgctggagtatacaccaactgcccgactaacaccactggaagc

ttgtgcacattcatttttgatgaattacgggacccaaatgtcaaactaccaaatgggcgagacacacctgcactct

tcaacttcaccactcaagaactgtcaagtaatccacctctggctaccatccttattcctcctcatgctcggattcaa

gcagctgcttcaacccccacaaatgccacagcagcgtcagatgctaatactggagaccgtggacagaccaataatgc

tgcttctgcatcagcttccaactccacc

β-catenin (S33Y)

atggctactcaagctgatttgatggagttggacatggccatggaaccagacagaaaagcggctgttagtcactggca

gcaacagtcttacctggactatggaatccattctggtgccactaccacagctccttctctgagtggtaaaggcaatc

ctgaggaagaggatgtggatacctcccaagtcctgtatgagtgggaacagggattttctcagtccttcactcaagaa

caagtagctgatattgatggacagtatgcaatgactcgagctcagagggtacgagctgctatgttccctgagacatt

agatgagggcatgcagatcccatctacacagtttgatgctgctcatcccactaatgtccagcgtttggctgaaccat

cacagatgctgaaacatgcagttgtaaacttgattaactatcaagatgatgcagaacttgccacacgtgcaatccct

gaactgacaaaactgctaaatgacgaggaccaggtggtggttaataaggctgcagttatggtccatcagctttctaa

aaaggaagcttccagacacgctatcatgcgttctcctcagatggtgtctgctattgtacgtaccatgcagaatacaa

atgatgtagaaacagctcgttgtaccgctgggaccttgcataacctttcccatcatcgtgagggcttactggccatc

tttaagtctggaggcattcctgccctggtgaaaatgcttggttcaccagtggattctgtgttgttttatgccattac

aactctccacaacctttattacatcaagaaggagctaaaatggcagtgcgtttagctggtgggctgcagaaaatgg

ttgccttgctcaacaaaacaaatgttaaattcttggctattacgacagactgccttcaaattttagcttatggcaac

caagaaagcaagctcatcatactggctagtggtggaccccaagctttagtaaatataatgaggacctatacttacga

aaaactactgtggaccacaagcagagtgctgaaggtgctatctgtctgctctagtaataagccggctattgtagaag

ctggtggaatgcaagctttaggacttcacctgacagatccaagtcaacgtcttgttcagaactgtctttggactctc

aggaatctttcagatgctgcaactaaacaggaagggatggaaggtctccttgggactcttgttcagcttctgggttc

agatgatataaatgtggtcacctgtgcagctggaattctttctaacctcacttgcaataattataagaacaagatga

tggtctgccaagtgggtggtatagaggctcttgtgcgtactgtccttcgggctggtgacagggaagacatcactgag

cctgccatctgtgctcttcgtcatctgaccagccgacaccaagaagcagagatggcccagaatgcagttcgccttca

ctatggactaccagttgtggttaagctcttacacccaccatcccactggcctctgataaaggctactgttggattga

ttcgaaatcttgccctttgtcccgcaaatcatgcacctttgcgtgagcagggtgccattccacgactagttcagttg

cttgttcgtgcacatcaggatacccagcgccgtacgtccatgggtgggacacagcagcaatttgtggaggggtccg

catggaagaaatagttgaaggttgtaccggagcccttcacatcctagctcgggatgttcacaaccgaattgttatca

gaggactaaataccattccattgtttgtgcagctgctttattctcccattgaaaacatccaaagagtagctgcaggg

gtcctctgtgaacttgctcaggacaaggaagctgcagaagctattgaagctgagggagccacagctcctctgacaga

gttacttcactctaggaatgaaggtgtggcgacatatgcagctgctgttttgttccgaatgtctgaggacaagccac

aagattacaagaaacggctttcagttgagctgaccagctctctcttcagaacagagccaatggcttggaatgagact

gctgatcttggacttgatattggtgcccagggagaaccccttggatatcgccaggatgatcctagctatcgttcttt

tcactctggtggatatggccaggatgccttgggtatggaccccatgatggaacatgagatgggtggccaccaccctg

gtgctgactatccagttgatgggctgccagatctggggcatgcccaggacctcatggatgggctgcctccaggtgac

agcaatcagctggcctggtttgatactgacctg

MAPK9

atgagcgacagtaaatgtgacagtcagttttatagtgtgcaagtggcagactcaaccttcactgtcctaaaacgtta

ccagcagctgaaaccaattggctctggggcccaagggattgtttgtgctgcatttgatacagttcttgggataaatg

ttgcagtcaagaaactaagccgtccttttcagaaccaaactcatgcaaagagagcttatcgtgaacttgtcctctta

aaatgtgtcaatcataaaaatataattagtttgttaaatgtgtttacaccacaaaaaactctagaagaatttcaaga

tgtgtatttggttatggaattaatggatgctaacttatgtcaggttattcacatggagctggatcatgaaagaatgt

cctaccttctttaccagatgctttgtggtattaaacatctgcattcagctggtataattcatagagatttgaagcct

agcaacattgttgtgaaatcagactgcaccctgaagatccttgactttggcctggcccggacagcgtgcactaactt

catgatgacccttacgtggtgacacggtactaccgggcgcccgaagtcatcctgggtatgggctacaaagagaacg

ttgatatctggtcagtgggttgcatcatgggagagctggtgaaaggttgtgtgatattccaaggcactgaccatatt

gatcagtggaataaagttattgagcagctgggaacaccatcagcagagttcatgaagaaacttcagccaactgtgag

gaattatgtcgaaaacagaccaaagtatcctggaatcaaatttgaagaactctttccagattggatattcccatcag

aatctgagcgagacaaaataaaaacaagtcaagccagagatctgttatcaaaaatgttagtgattgatcctgacaag

cggatctctgtagacgaagctctgcgtcacccatacatcactgtttggtatgaccccgccgaagcagaagccccacc

acctcaaatttatgatgcccagttggaagaaagagaacatgcaattgaagaatggaaagagctaatttacaaagaag

tcatggattgggaagaaagaagcaagaatggtgttgtaaaagatcagccttcagatgcagcagtaagtagcaacgcc

actccttctcagtcttcatcgatcaatgacatttcatccatgtccactgagcagacgctggcctcagacacagacag

cagtcttgatgcctcgacgggacccctgaaggctgtcga

FLAG-MKK7-JNK2

atggactacaaagacgatgatgacaaaatggcggcgtcctccctggagcagaagctgtcccgcctggaagccaagct

gaagcaggagaaccgtgaggcccgcaggaggatcgacctcaacttggatatcagcccacagcggcccaggcccaccc

tgcaactcccactggccaacgatgggggcagccgctcaccatcctcagagagctccccacagcaccctacacccccc

acccggccccgccacatgctggggctcccatcaaccttgttcacaccgcgcagtatggagagcatcgagattgacca

gaagctgcaggagatcatgaagcagacagggtacctgactatcgggggccagcgttatcaggcagaaatcaatgact

tggagaacttgggtgagatgggcagtggtacctgtggtcaggtgtggaagatgcggttccggaagacaggccacatc

attgctgttaagcaaatgcggcgctctgggaacaaggaagagaataagcgcattttgatggacctggatgtagtact

caagagccatgactgcccttacatcgttcagtgctttggcaccttcatcaccaacacagacgtctttattgccatgg

agctcatgggcacatgtgcagagaagctgaagaaacgaatgcagggccccattccagagcgaatcctgggcaagatg

actgtggcgattgtgaaagcactgtactatctgaaggagaagcatggcgtcatccatcgcgatgtcaaaccctccaa

catcctgctagatgagcggggccagatcaagctctgtgactttggcatcagtggccgccttgttgactccaaagcca

aaacacggagtgctggctgtgctgcctatatggctcccgagcgcatcgaccctccagatcccaccaagcctgactat

gacatccgagctgatgtgtggagcctgggcatctcactggtggagctggcaacaggacagttcccctataagaactg

caagacggactttgaggtcctcaccaaagtcctacaggaagagcccccactcctgcctggtcacatgggcttctcag

gggacttccagtcatttgtcaaagactgccttactaaagatcacaggaagagaccaaagtataataagctacttgaa

cacagcttcatcaagcactatgagatactcgaggtggatgtcgcgtcctggtttaaggatgtcatggcgaagaccga

gtccccaaggactagtggagtcctgagtcagcaccatctgcccttcttcagtgggagtctggaggagtctcccactt

ccccaccttctcccaagtccttccctctgtcaccagccatccctcacgaaattctgagccgtcgtccgagttaccgt

aaaggatccatgagcgacagtaaatgtgacagtcagttttatagtgtgcaagtggcagactcaaccttcactgtcct

aaaacgttaccagcagctgaaaccaattggctctggggcccaagggattgtttgtgctgcatttgatacagttcttg

ggataaatgttgcagtcaagaaactaagccgtccttttcagaaccaaactcatgcaaagagagcttatcgtgaactt

gtcctcttaaaatgtgtcaatcataaaaatataattagtttgttaaatgtgtttacaccacaaaaaactctagaaga

atttcaagatgtgtatttggttatggaattaatggatgctaacttatgtcaggttattcacatggagctggatcatg

aaagaatgtcctaccttctttaccagatgctttgtggtattaaacatctgcattcagctggtataattcatagagat

ttgaagcctagcaacattgttgtgaaatcagactgcaccctgaagatccttgactttggcctggcccggacagcgtg

cactaacttcatgatgacccccttacgtggtgacacggtactaccgggcgcccgaagtcatcctgggtatgggctaca

aagagaacgttgatatctggtcagtgggttgcatcatgggagagctggtgaaaggttgtgtgatattccaaggcact

gaccatattgatcagtggaataaagttattgagcagctgggaacaccatcagcagagttcatgaagaaacttcagcc

aactgtgaggaattatgtcgaaaacagaccaaagtatcctggaatcaaatttgaagaactctttccagattggatat

tcccatcagaatctgagcgagacaaaataaaaacaagtcaagccagagatctgttatcaaaaatgttagtgattgat

cctgacaagcggatctctgtagacgaagctctgcgtcacccatacatcactgtttggtatgaccccgccgaagcag

MEK5_DD (S311D, T315D)

atgctgtggctagcccttggccccttcctgccatggagaaccaggtgctggtaattcgcatcaagatcccaaatag

tggcgcggtggactggacagtgcactccgggccgcagttactcttcagggatgtgctggatgtgataggccaggttc

tgcctgaagcaacaactacagcatttgaatatgaagatgaagatggtgatcgaattacagtgagaagtgatgaggaa

atgaaggcaatgctgtcatattattattccacagtaatggaacagcaagtaaatggacagttaatagagcctctgca

gatatttccaagagcctgcgagcctcctggggaacggaacatacgtggcctgaaggtgaatactcgggccggaccct

ctcaacacagcagcccagcagtctcagattcacttccaagcaatagcttaaagaagtcttctgctgaactgaaaaaa

atactagccaatggccagatgaatgaacaagacatacgatatcgggacactcttggtcatggcaacggaggcacagt

ctacaaagcatatcatgtcccgagtgggaaaatattagctgtaaaggtcatactactagatattacactggaacttc

agaagcaaattatgtctgaattggaaattctttataagtgcgattcatcatatatcattggattttatggagcattt

tttgtagaaaacaggatttcaatatgtacagaattcatggatgggggatctttggatgtatataggaaaatgccaga

acatgtccttggaagaattgcagtagcagttgttaaaggccttacttatttgtggagtttaaagattttacatagag

acgtgaagccctccaatatgctagtaaacacaagaggacaggttaagctgtgtgattttggagttagcactcagctg

gtgaatgatatagccaaggattatgttggaacaaatgcttatatggcgcctgaaaggatttcaggggagcagtatgg

aattcattctgatgtctggagcttaggaatctcttttatggagcttgctcttgggaggtttccatatcctcagattc

agaaaaaccagggatctttaatgcctctccagcttctgcagtgcattgttgatgaggattcgcccgtccttccaatt

ggagagttctcggagccatttgtacatttcatcactcagtgtatgcgaaaacagccaaaagaaaggccagcacctga

agaattgatgggccacccgttcatcgtgcagttcaatgatggaaatgccgccgtggtgtccatgtgggtgtgccggg

cgctggaggagaggcggagccagcaggggcccccg

myr-FLAG-MEK5

atggggtcttcaaaatctaaaccaaaggaccccagccagcgccggcgcaggatccgaggttaccttgactacaaaga

cgatgacgacaagcaattgacaagtttgtacaaaaaagttggatgctgtggctagcccttggcccctttcctgccat

ggagaaccaggtgctggtaattcgcatcaagatcccaaatagtggcgcggtggactggacagtgcactccgggccgc

agttactcttcagggatgtgctggatgtgataggccaggttctgcctgaagcaacaactacagcatttgaatatgaa

gatgaagatggtgatcgaattacagtgagaagtgatgaggaaatgaaggcaatgctgtcatattattattccacagt

aatggaacagcaagtaaatggacagttaatagagcctctgcagatatttccaagagcctgcaagcctcctggggaac

ggaacatacatggcctgaaggtgaatactcgggccggacgcctctcaacacagcagcccagcagtctcagattcactt

ccaagcaatagcttaaagaagtcttctgctgaactgaaaaaaatactagccaatggccagatgaatgaacaagacat

acgatatcgggacactcttggtcatggcaacggaggcacagtctacaaagcatatcatgtcccgagtgggaaaatat

tagctgtaaaggtcatactactagatattacactggaacttcagaagcaaattatgtctgaattggaaattctttat

aagtgcgattcatcatatatcattggattttatggagcatttttttgtagaaaacaggatttcaatatgtacagaatt

catggatgggggatctttggatgtatataggaaaatgccagaacatgtccttggaagaattgcagtagcagttgtta

aaggccttacttatttgtggagtttaaagattttacatagagacgtgaagccctccaatatgctagtaaacacaaga

ggacaggttaagctgtgtgattttggagttagcactcagctggtgaattctatagccaagacgtatgttggaacaaa

tgcttatatggcgcctgaaaggatttcaggggagcagtatggaattcattctgatgtctggagcttaggaatctctt

ttatggagcttgctcttgggaggtttccatatcctcagattcagaaaaaccagggatctttaatgcctctccagctt

ctgcagtgcattgttgatgaggattcgcccgtccttccagttggagagttctcggagccatttgtacatttcatcac

tcagtgtatgcgaaaacagccaaaagaaaggccagcacctgaagaattgatgggccacccgttcatcgtgcagttca

atgatggaaatgccgccgtggtgtccatgtgggtgtgccgggcgctggaggagaggcggagccagcaggggcccccg

Notch1 ICD

atgcggcggcagcatggccagctctggttccctgagggcttcaaagtgtctgaggccagcaagaagaagcggcggga

gcccctcggcgaggactccgtgggcctcaagcccctgaagaacgcttcagacggtgccctcatggacgacaaccaga

atgagtgggggacgaggacctggagaccaagaagttccggttcgaggagcccgtggttctgcctgacctggacgac

cagacagaccaccggcagtggactcagcagcacctggatgccgctgacctgcgcatgtctgccatggcccccacacc

gccccagggtgaggttgacgccgactgcatggacgtcaatgtccgcgggcctgatggcttcaccccgctcatgatcg

cctcctgcagcgggggcggcctggagacgggcaacagcgaggaagaggaggacgcgccggccgtcatctccgacttc

atctaccagggcgccagcctgcacaaccagacagaccgcacgggcgagaccgccttgcacctggccgcccgctactc

acgctctgattccgccaagcgcctgctggaggccagcgcagatgccaacatccaggacaacatgggccgcaccccgc

tgcatgcggctgtgtctgccgacgcacaaggtgtcttccagatcctgatccggaaccgagccacagacctggatgcc

cgcatgcatgatggcacgacgccactgatcctggctgcccgcctggccgtggagggcatgctggaggacctcatcaa

ctcacacgccgacgtcaacgccgtagatgacctgggcaagtccgccctgcactgggccgccgccgtgaacaatgtgg

atgccgcagttgtgctcctgaagaacggggctaacaaagatatgcagaacaacagggaggagacacccctgtttctg

gccgcccgggagggcagctacgagaccgccaaggtgctgctggaccactttgccaaccgggacatcacggatcatat

ggaccgcctgccgcgcgacatcgcacaggagcgcatgcatcacgacatcgtgaggctgctggacgagtacaacctgg

tgcgcagcccgcagctgcacggagccccgctgggggcacgcccaccctgtcgcccccgctctgctcgcccaacggc

tacctgggcagcctcaagcccggcgtgcagggcaagaaggtccgcaagcccagcagcaaaggcctggcctgtggaag

caaggaggccaaggacctcaaggcacggaggaagaagtcccaggatggcaagggctgcctgctggacagctccggca

tgctctcgcccgtggactccccggagtcaccccatggctacctgtcagacgtggcctcgccgccactgctgccctcc

ccgttccagcagtctccgtccgtgcccctcaaccacctgcctgggatgcccgacacccacctgggcatcgggcacct

gaacgtggcggccaagcccgagatggcggcgctgggtgggggcggccggctggccattgagactggcccacctcgtc

tctcccacctgcctgtggcctctggcaccagcaccgtcctgggctccagcagcggaggggccctgaatttcactgtg

ggcgggtccaccagtttgaatggtcaatgcgagtggctgtcccggctgcagagcggcatggtgccgaaccaatacaa

ccctctgcgggggagtgtggcaccaggcccccctgagcacacaggcccccctccctgcagcatggcatggtaggcccgc

tgcacagtagccttgctgccagcgccctgtcccagatgatgagctaccagggcctgcccagcacccggctggccacc

cagcctcacctggtgcagacccagcaggtgcagccacaaaacttacagatgcagcagcagaacctgcagccagcaaa

catccagcagcagcaaagcctgcagccgccaccaccaccaccacagccgcaccttggcgtgagctcagcagccagcg

gccacctgggccggagcttcctgagtggagagccgagccaggcagacgtgcagccactgggccccagcagcctggcg

gtgcacactattctgccccaggagagccccgccctgcccacgtcgctgccatcctcgctggtcccacccgtgaccgc

agcccagttcctgacgcccccctcgcagcacagctactcctcgcctgtggacaacacccccagccaccagctacagg

tgcctgagcacccccttcctcaccccgtcccctgagtcccctgaccagtggtccagctcgtccccgcattccaacgtc

tccgactggtccgagggcgtctccagccctcccaccagcatgcagtcccagatcgcccgcattccggaggccttcaa

g

Notch3 ICD

atggtggcccggcgcaagcgcgagcacagcaccctctggttccctgagggcttctcactgcacaaggacgtggcctc

tggtcacaagggccggcgggaacccgtgggccaggacgcgctgggcatgaagaacatggccaagggtgagagcctga

tgggggaggtggccacagactggatggacacagagtgcccagaggccaagcggctaaaggtagaggagccaggcatg

ggggctgaggaggctgtggattgccgtcagtggactcaacaccatctggttactgctgacatccgcgtggcaccagc

catggcactgacaccaccacagggcgacgcagatgctgatggcatggatgtcaatgtgcgtggcccagatggcttca

ccccgctaatgctggcttccttctgtgggggggctctggagccaatgccaactgaagaggatgaggcagatgacaca

tcagctagcatcatctccgacctgatctgccaggggctcagcttggggcacggactgaccgtactggcgagactgc

tttgcacctggctgcccgttatgcccgtgctgatgcagccaagcggctgctggatgctggggcagacaccaatgccc

aggaccactcaggccgcactcccctgcacacagctgtcacagccgatgcccagggtgtcttccagattctcatccga

aaccgctctacagacttggatgcccgcatggcagatggctcaacggcactgatcctggcggcccgcctggcagtaga

gggcatggtggaagagctcatcgccagccatgctgatgtcaatgctgtggatgagcttgggaaatcagccttacact

gggctgcggctgtgaacaacgtggaagccactttggccctgctcaaaaatggagccaataaggacatgcaggatagc

aaggaggagacccccctattcctggccgcccgcgagggcagctatgaggctgccaagctgctgttggaccactttgc

caaccgtgagatcaccgaccacctggacaggctgccgcgggacgtagcccaggagagactgcaccaggacatcgtgc

gcttgctggatcaacccagtgggccccgcagcccccccggtccccacggcctggggcctctgctctgtcctccaggg

gccttcctccctggcctcaaagcggcacagtcggggtccaagaagagcaggaggcccccgggaaggcggggctggg

gccgcaggggccccgggggcggggcaagaagctgacgctggcctgcccgggccccctggctgacagctcggtcacgc

tgtcgcccgtggactcgctggactccccgcggcctttcggtgggcccccctgcttcccctggtggcttcccccttgag

gggccctatgcagctgccactgccactgcagtgtctctggcacagcttggtggcccaggccgggcgggtctagggcg

ccagccccctggaggatgtgtactcagcctgggcctgctgaaccctgtggctgtgcccctcgattgggcccggctgc

ccccacctgcccctccaggcccctcgttcctgctgccactggcgccgggaccccagctgctcaacccagggacccccc

gtctccccgcaggagcggcccccgccttacctggcagtcccaggacatggcgaggagtacccggcggctggggcaca

cagcagcccccaaaggcccgcttcctgcgggttcccagtgagcacccttacctgaccccatcccccgaatcccctg

agcactgggccagcccctcacctccctccctctcagactggtccgaatccacgcctagcccagccactgccactggg

gccatggccaccaccactggggcactgcctgcccagccacttcccttgtctgttcccagctcccttgctcaggccca

gacccagctggggccccagccggaagttacccccaagaggcaagtgttggcc

MAPK14

atgtctcaggagaggcccacgttctaccggcaggagctgaacaagacaatctgggaggtgcccgagcgttaccagaa

cctgtctccagtgggctctggcgcctatggctctgtgtgtgctgcttttgacacaaaaacggggttacgtgtggcag

tgaagaagctctccagaccatttcagtccatcattcatgcgaaaagaacctacagagaactgcggttacttaaacat

atgaaacatgaaaatgtgattggtctgttggacgtttttacacctgcaaggtctctggaggaattcaatgatgtgta

tctggtgacccatctcatgggggcagatctgaacaacattgtgaaatgtcagaagcttacagatgaccatgttcagt

tccttatctaccaaattctccgaggtctaaagtatatacattcagctgacataattcacagggacctaaaacctagt

aatctagctgtgaatgaagactgtgagctgaagattctggattttggactggctcggcacacagatgatgaaatgac

aggctacgtggccactaggtggtacagggctcctgagatcatgctgaactggatgcattacaaccagacagttgata

tttggtcagtgggatgcataatggccgagctgttgactggaagaacattgtttcctggtacagaccatattgatcag

ttgaagctcattttaagactcgttggaaccccaggggctgagcttttgaagaaaatctcctcagagtctgcaagaaa

ctatattcagtctttgactcagatgccgaagatgaactttgcgaatgtatttattggtgccaatcccctggctgtcg

acttgctggagaagatgcttgtattggactcagataagagaattacagcggcccaagcccttgcacatgcctacttt

gctcagtaccacgatcctgatgatgaaccagtggccgatccttatgatcagtcctttgaaagcagggacctccttat

agatgagtggaaaagcctgacctatgatgaagtcatcagctttgtgccaccacccccttgaccaagaagagatggagt

cc

FLAG-MKK6 (S207E, T211E)

atggactacaaagacgatgatgacaaatctcagtcgaaaggcaagaagcgaaaccctggccttaaaattccaaaaga

agcatttgaacaacctcagaccagttccacaccacctagagatttagactccaaggcttgcatttctattggaaatc

agaactttgaggtgaaggcagatgacctggagcctataatggaactgggacgaggtgcgtacggggtggtggagaag

atgcggcacgtgcccagcgggcagatcatggcagtgaagcggatccgagccacagtaaatagccaggaacagaaacg

gctactgatggatttggatatttccatgaggacggtggactgtccattcactgtcacctttatggcgcactgtttc

gggagggtgatgtgtggatctgcatggagctcatggatacatcactagataaattctacaaacaagttattgataaa

ggccagacaattccagaggacatcttagggaaaatagcagtttctattgtaaaagcattagaacatttacatagtaa

gctgtctgtcattcacagagacgtcaagccttctaatgtactcatcaatgctctcggtcaagtgaagatgtgcgatt

ttggaatcagtggctacttggtggacgaggttgctaaagaaattgatgcaggttgcaaaccatacatggcccctgaa

agaataaacccagagctcaaccagaagggatacagtgtgaagtctgacatttggagtctgggcatcacgatgattga

gttggccatccttcgatttccctatgattcatggggaactccatttcagcagctcaaacaggtggtagaggagccat

cgccacaactcccagcagacaagttctctgcagagtttgttgactttacctcacagtgcttaaagaagaattccaaa

gaacggcctacatacccagagctaatgcaacatccatttttcaccctacatgaatccaaaggaacagatgtggcatc

ttttgtaaaactgattcttggagac

GLI2 Truncation

atggccctcacctccatcaatgccacgcccacccagctcagcagcagcagcaactgtctgagtgacaccaaccagaa

caagcagagcagtgagtcggccgtcagcagcaccgtcaaccctgtcgccattcacaagcgcagcaaggtcaagaccg

agcctgagggcctgcggccggcctcccctctggcgctgacgcagggccaggtgtctggacacggctcatgtgggtgt

gcccttcccctctcccaggagcagctggctgacctcaaggaagatctggacagggatgactgtaagcaggaggctga

ggtggtcatctatgagaccaactgccactgggaagactgcaccaaggagtacgacacccaggagcagctggtgcatc

acatcaacaacgagcacatccacggggagaagaaggagtttgtgtgccgctggcaggcctgcacgcgggagcagaag

cccttcaaggcgcagtacatgctggtggtgcacatgcggcgacacacgggcgagaagccccacaagtgcacgttcga

gggctgctcgaaggcctactcccgcctggagaacctgaagacacacctgcggtcccacaccggggagaagccatatg

tgtgtgagcacgagggctgcaacaaagccttctccaacgcctcggaccgcgccaagcaccagaatcgcacccactcc

aacgagaaaccctacatctgcaagatcccaggctgcaccaagagatacacagaccccagctctctccggaagcatgt

gaaaacggtccacggcccagatgcccacgtcaccaagaagcagcgcaatgacgtgcacctccgcacaccgctgctca

aagagaatggggacagtgaggccggcacggagcctggcggcccagagagcaccgaggccagcagcaccagccaggcc

gtggaggactgcctgcacgtcagagccatcaagaccgagagctccgggctgtgtcagtccagccccggggcccagtc

gtcctgcagcagcgagccctctcctctgggcagtgcccccaacaatgacagtggcgtggagatgccggggacggggc

ccgggagcctgggagacctgacggcactggatgacacacccccaggggccgacacctcagccctggctgccccctcc

gctggtggcctccagctgcgcaaacacatgaccaccatgcaccggttcgagcagctcaagaaggagaagctcaagtc

actcaaggattcctgctcatgggccgggccgactccacacacgcggaacaccaagctgcctcccctcccgggaagtg

gctccatcctggaaaacttcagtggcagtgggggcggcgggcccgcggggctgctgccgaacccgcggctgtcggag

ctgtccgcgagcgaggtgaccatgctgagccagctgcaggagcgccgcgacagctccaccagcacggtcagctcggc

ctacaccgtgagccgccgctcctccggcatctccccctacttctccagccgccgctccagcgaggcctcgcccctgg

gcgccggccgcccgcacaacgcgagctccgctgactcctacgaccccatctccacggacgcgtcgcggcgctcgagc

gaggccagccagtgcagcggcggctccgggctgctcaacctcacgccggcgcagcagtacagcctgcgggccaagta

cgcggcagccactggcggccccccgcccactccgctgccgggcctggagcgcatgagcctgcggaccaggctggcgc

tgctggacgcgcccgagcgcacgctgcccgccggctgcccacgcccactggggccgcggcgtggcagcgacgggccg

acctatggccacggccacgcgggggctgcgcccgccttcccccacgaggctccaggcggcggagccaggcgggccag

cgaccctgtgcggcggcccgatgccctgtccctgccgcgggtgcagcgcttccacagcacccacaacgtgaaccccg

gcccgctgccgccctgtgccgacaggcgaggcctccgcctgcagagccacccgagcaccgacggcggcctggcccgc

ggcgcctactcgccccggccgcctagcatcagcgagaacgtggcgatggaggccgtggcggcaggagtggacggcgc

ggggcccgaggccgacctggggctgccggaggacgacctggtgcttccagacgacgtggtgcagtacatcaaggcgc

acgccagtggcgctctggacgagggcaccgggcaggtgtatcccacggaaagcactggcttctctgacaaccccaga

ctacccagcccggggctgcacggccagcgcaggatggtggctgcggactccaacgtgggcccctccgcccctatgct

gggaggatgccagttaggctttggggcgccctccagcctgaacaaaaataacatgcctgtgcagtggaatgaggtga

gctccggcaccgtagacgccctggccagccaggtgaagcctccacccttcctcagggcaacctggcggtggtgcag

cagaagcctgcctttggccagtacccgggctacagtccgcaaggcctacaggctagccctgggggcctggacagcac

gcagccacacctgcagccccgcagcggagccccctcccagggcatccccagggtaaactacatgcagcagctgcgac

agccagtggcaggcagccagtgtcctggcatgactaccactatgagcccccatgcctgctatggccaagtccacccc

cagctgagccccagcaccatcagtggggccctcaaccagttcccccaatcctgcagcaacatgccagccaagccagg

gcatctggggcaccctcagcagacagaagtggcacctgaccccaccacgatgggcaatcgccacagggaacttgggg

tccccgattcagccctggctggagtgccaccacctcacccagtccagagctacccacagcagagccatcacctggca

gcctccatgagccaggagggctaccaccaggtccccagccttctgcctgcccgccagcctggcttcatggagcccca

aacaggcccgatgggggtggctacagcaggctttggcctagtgcagccccggcctcccctcgagcccagccccactg

gccgccaccgtggggtacgtgctgtgcagcagcagctggcctacgccagggccacaggccatgccatggctgccatg

ccgtccagtcaggaaacagcagaggctgtgcccaagggagcgatgggcaacatggggtcggtgcctccccagccgcc

tccgcaggacgcaggtggggccccggaccacagcatgctctactactacggccagatccacatgtacgaacaggatg

gaggcctggagaacctcgggagctgccaggtcatgcggtcccagccaccacagccacaggcctgtcaggacagcatc

cagccccagcccttgccctcaccaggggtcaaccaggtgtccagcactgtggactcccagctcctggaggcccccca

gattgacttcgatgccatcatggatgatggcgatcactcgagtttgttctcgggtgctctgagccccagcctcctcc

acagcctctcccagaactcctcccgcctcaccacccccgaaactccttgaccctgccctccatccccgcaggcatc

agcaacatggctgtcggggacatgagctccatgctcaccagcctcgccgaggagagcaagttcctgaacatgatgac

c

SmoM2 (W535L)

atggccgctgcccgcccagcgcgggggccggagctcccgctcctggggctgctgctgctgctgctgggggaccc

gggccgggggggcggcctcgagcgggaacgcgaccgggcctgggcctcggagcgcgggcgggagcgcgaggaggagcg

cggcggtgactggccctccgccgccgctgagccactgcggccgggctgccccctgtgagccgctgcgctacaacgtg

tgcctgggctcggtgctgccctacggggccacctccacactgctggccggagactcggactcccaggaggaagcgca

cggcaagctcgtgctctggtcgggcctccggaatgccccccgctgctgggcagtgatccagcccctgctgtgtgccg

tatacatgcccaagtgtgagaatgaccgggtggagctgcccagccgtaccctctgccaggccacccgaggcccctgt

gccatcgtggagagggagcggggctggcctgacttcctgcgctgcactcctgaccgcttccctgaaggctgcacgaa

tgaggtgcagaacatcaagttcaacagttcaggccagtgcgaagtgcccttggttcggacagacaaccccaagagct

ggtacgaggacgtggagggctgcggcatccagtgccagaacccgctcttcacagaggctgagcaccaggacatgcac

agctacatcgcggccttcggggccgtcacgggcctctgcacgctcttcaccctggccacattcgtggctgactggcg

gaactcgaatcgctaccctgctgttattctcttctacgtcaatgcgtgcttctttgtgggcagcattggctggctgg

cccagttcatggatggtgcccgccgagagatcgtctgccgtgcagatggcaccatgaggcttggggagcccacctcc

aatgagactctgtcctgcgtcatcatctttgtcatcgtgtactacgccctgatggctggtgtggtttggtttgtggt

cctcacctatgcctggcacacttccttcaaagccctgggcaccacctaccagcctctctcgggcaagacctcctact

tccacctgctcacctggtcactccccttttgtcctcactgtggcaatccttgctgtggcgcaggtggatggggactct

gtgagtggcatttgttttgtgggctacaagaactaccgataccgtgcgggcttcgtgctggccccaatcggcctggt

gctcatcgtgggaggctacttcctcatccgaggagtcatgactctgttctccatcaagagcaaccaccccgggctgc

tgagtgagaaggctgccagcaagatcaacgagaccatgctgcgcctgggcatttttggcttcctggcctttggcttt

gtgctcattaccttcagctgccacttctacgacttcttcaaccaggctgagtgggagcgcagcttccgggactatgt

gctatgtcaggccaatgtgaccatcgggctgcccaccaagcagcccatccctgactgtgagatcaagaatcgcccga

gccttctggtggagaagatcaacctgtttgccatgtttggaactggcatcgccatgagcaccttggtctggaccaag

gccacgctgctcatctggaggcgtacctggtgcaggttgactgggcagagtgacgatgagccaaagcggatcaagaa

gagcaagatgattgccaaggccttctctaagcggcacgagctcctgcagaacccaggccaggagctgtccttcagca

tgcacactgtgtcccacgacgggcccgtggcgggcttggcctttgacctcaatgagccctcagctgatgtctcctct

gcctgggcccagcatgtcaccaagatggtggctcggagaggagccatactgccccaggatatttctgtcacccctgt

ggcaactccagtgcccccagaggaacaagccaacctgtggctggttgaggcagagatctccccagagctgcagaagc

gcctgggccggaagaagaagaggaggaagaggaagaaggaggtgtgcccgctggcgccgccccctgagcttcacccc

cctgcccctgcccccagtaccattcctcgactgcctcagctgccccggcagaaatgcctggtggctgcaggtgcctg

gggagctggggactcttgccgacagggagcgtggaccctggtctccaacccattctgcccagagcccagtcccccctc

aggatccatttctgcccagtgcaccggcccc

TGFβR1 (T204D)

atggaggcggcggtcgctgctccgcgtccccggctgctcctcctcgtgctggcggcggcggcggcggcggcggc

gctgctcccgggggcgacggcgttacagtgtttctgccacctctgtacaaaagacaattttacttgtgtgacagatg

ggctctgctttgtctctgtcacagagaccacagacaaagttatacacaacagcatgtgtatagctgaaattgactta

attcctcgagataggccgtttgtatgtgcaccctcttcaaaaactgggtctgtgactacaacatattgctgcaatca

ggaccattgcaataaaatagaacttccaactactgtaaagtcatcacctggccttggtcctgtggaactggcagctg

tcattgctggaccagtgtgcttcgtctgcatctcactcatgttgatggtctatatctgccacaaccgcactgtcatt

caccatcgagtgccaaatgaagaggacccttcattagatcgccctttttatttcagagggtactacgttgaaagactt

aatttatgatatgacaacgtcaggttctggctcaggtttaccattgcttgttcagagaacaattgcgagagatattg

tgttacaagaaagcattggcaaaggtcgatttggagaagtttggagaggaaagtggcggggagaagaagttgctgtt

aagatattctcctctagagaagaacgttcgtggttccgtgaggcagagatttatcaaactgtaatgttacgtcatga

aaacatcctgggatttatagcagcagacaataaagacaatggtacttggactcagctctggttggtgtcagattatc

atgagcatggatctctttttgattacttaaacagatacacagttactgtggaaggaatgataaaacttgctctgtcc

acggcgagcggtcttgcccatcttcacatggagattgttggtacccaaggaaagccagccattgctcatagagattt

gaaatcaaagaatatcttggtaaagaagaatggaacttgctgtattgcagacttaggactggcagtaagacatgatt

cagccacagataccattgatattgctccaaaccacagagtgggaacaaaaaggtacatggcccctgaagttctcgat

gattccataaatatgaaacattttgaatccttcaaacgtgctgacatctatgcaatgggcttagtattctgggaaat

tgctcgacgatgttccattggtggaattcatgaagattaccaactgccttattatgatcttgtaccttctgacccat

cagttgaagaaatgagaaaagttgtttgtgaacagaagttaaggccaaatatcccaaacagatggcagagctgtgaa

gccttgagagtaatggctaaaattatgagagaatgttggtatgccaatggagcagctaggcttacagcattgcggat

taagaaaacattatcgcaactcagtcaacaggaaggcatcaaaatg

BCL2

atggcgcacgctgggagaacagggtacgataaccgggagatagtgatgaagtacatccattataagctgtcgcagag

gggctacgagtgggatgcgggagatgtgggcgccgcgcccccgggggccgcccccgcaccgggcatcttctcctccc

agcccgggcacacgccccatccagccgcatcccgggacccggtcgccaggacctcgccgctgcagaccccggctgcc

cccggcgccgccgcggggcctgcgctcagcccggtgccacctgtggtccacctgaccctccgccaggccggcgacga

cttctcccgccgctaccgccgcgacttcgccgagatgtccagccagctgcacctgacgcccttcaccgcgcggggac

gctttgccacggtggtggaggagctcttcagggacggggtgaactgggggaggattgtggccttctttgagttcggt

ggggtcatgtgtgtggagagcgtcaaccgggagatgtcgcccctggtggacaacatcgccctgtggatgactgagta

cctgaaccggcacctgcacacctggatccaggataacggaggctgggatgcctttgtggaactgtacggccccagca

tgcggcctctgtttgatttctcctggctgtctctgaagactctgctcagtttggccctggtgggagcttgcatcacc

ctgggtgcctatctgggccacaag

BCL-XL

atgtctcagagcaaccgggagctggtggttgactttctctcctacaagctttcccagaaaggatacagctggagtca

gtttagtgatgtggaagagaacaggactgaggccccagaagggactgaatcggagatggagacccccagtgccatca

atggcaacccatcctggcacctggcagacagccccgcggtgaatggagccactggccacagcagcagtttggatgcc

cgggaggtgatccccatggcagcagtaaagcaagcgctgagggaggcaggcgacgagtttgaactgcggtaccggcg

ggcattcagtgacctgacatcccagctccacatcaccccagggacagcatatcagagctttgaacaggtagtgaatg

aactcttccgggatggggtaaactggggtcgcattgtggccttttttctccttcggcggggcactgtgcgtggaaagc

gtagacaaggagatgcaggtattggtgagtcggatcgcagcttggatggccacttacctgaatgaccacctagagcc

ttggatccaggagaacggcggctgggatactttgtggaactctatgggaacaatgcagcagccgagagccgaaagg

gccaggaacgcttcaaccgctggttcctgacgggcatgactgtggccggcgtggttctgctgggctcactcttcagt

cggaaa

Caspase8 (C360A)

atgttggaggaaagcaatctgtccttcctgaaggagctgctcttccgaattaatagactggatttgctgattaccta

cctaaacactagaaaggaggagatggaaagggaacttcagacaccaggcagggctcaaatttctgcctacagggtca

tgctctatcagatttcagaagaagtgagcagatcagaattgaggtcttttaagtttcttttgcaagaggaaatctcc

aaatgcaaactggatgatgacatgaacctgctggatattttcatagagatggagaagagggtcatcctgggagaagg

aaagttggacatcctgaaaagagtctgtgcccaaatcaacaagagcctgctgaagataatcaacgactatgaagaat

tcagcaaagagagaagcagcagccttgaaggaagtcctgatgaattttcaaatggggaggagttgtgtggggtaatg

acaatctcggactctccaagagaacaggatagtgaatcacagactttggacaaagtttaccaaatgaaaagcaaacc

tcggggatactgtctgatcatcaacaatcacaattttgcaaaagcacgggagaaagtgcccaaacttcacagcatta

gggacaggaatggaacacacttggatgcaggggctttgaccacgacctttgaagagcttcattttgagatcaagccc

cacgatgactgcacagtagagcaaatctatgagattttgaaaatctaccaactcatggaccacagtaacatggactg

cttcatctgctgtatcctctcccatggagacaagggcatcatctatggcactgatggacaggaggcccccatctatg

agctgacatctcagttcactggtttgaagtgcccttcccttgctggaaaacccaaagtgtttttttattcaggctgct

caggggataactaccagaaaggtatacctgttgagactgattcagaggagcaaccctatttagaaatggatttatc

atcacctcaaacgagatatatcccggatgaggctgactttctgctggggatggccactgtgaataactgtgtttcct

accgaaaccctgcagagggaacctggtacatccagtcactttgccagagcctgagagagcgatgtcctcgaggcgat

gatattctcaccatcctgactgaagtgaactatgaagtaagcaacaaggatgacaagaaaaacatggggaaacagat

gcctcagcctactttcacactaagaaaaaaacttgtcttcccttctgat

Caspase3 (C163A)

atggagaacactgaaaactcagtggattcaaaatccattaaaaatttggaaccaaagatcatacatggaagcgaatc

aatggactctggaatatccctggacaacagttataaaatggattatcctgagatgggtttatgtataataattaata

ataagaattttcataaaagcactggaatgacatctcggtctggtacagatgtcgatgcagcaaacctcagggaaaca

ttcagaaacttgaaatatgaagtcaggaataaaaatgatcttacacgtgaagaaattgtggaattgatgcgtgatgt

ttctaaagaagatcacagcaaaaggagcagttttgtttgtgtgcttctgagccatggtgaagaaggaataatttttg

gaacaaatggacctgttgacctgaaaaaaataacaaacttttcagaggggatcgttgtagaagtctaactggaaaa

cccaaacttttcattattcaggccgcccggggtacagaactggactgtggcattgagacagacagtggtgttgatga

tgacatggcgtgtcataaaataccagtggaggccgacttcttgtatgcatactccacagcacctggttattattctt

ggcgaaattcaaaggatggctcctggttcatccagtcgctttgtgccatgctgaaacagtatgccgacaagcttgaa

tttatgcacattcttacccgggttaaccgaaaggtggcaacagaatttgagtccttttccttgacgctacttttca

tgcaaagaaacagattccatgtattgtttccatgctcacaaaagaactctatttttatcac

ERα (Y537S)

atgaccatgaccctccacaccaaagcatctgggatggccctactgcatcagatccaagggaacgagctggagcccct

gaaccgtccgcagctcaagatccccctggagcggcccctgggcgaggtgtacctggacagcagcaagcccgccgtgt

acaactaccccgagggcgccgcctacgagttcaacgccgcggccgccgccaacgcgcaggtctacggtcagaccggc

ctcccctacggccccgggtctgaggctgcggcgttcggctccaacggcctgggggtttccccccactcaacagcgt

gtctccgagcccgctgatgctactgcacccgccgccgcagctgtcgcctttcctgcagccccacggccagcaggtgc

cctactacctggagaacgagcccagcggctacacggtgcgcgaggccggcccgccggcattctacaggccaaattca

gataatcgacgccagggtggcagagaaagattggccagtaccaatgacaagggaagtatggctatggaatctgccaa

ggagactcgctactgtgcagtgtgcaatgactatgcttcaggctaccattatggagtctggtcctgtgagggctgca

aggccttcttcaagagaagtattcaaggacataacgactatatgtgtccagccaccaaccagtgcaccattgataaa

aacaggaggaagagctgccaggcctgccggctccgcaaatgctacgaagtgggaatgatgaaaggtgggatacgaaa

agaccgaagaggagggagaatgttgaaacacaagcgccagagagatgatggggagggcaggggtgaagtggggtctg

ctggagacatgagagctgccaacctttggccaagcccgctcatgatcaaacgctctaagaagaacagcctggccttg

tccctgacggccgaccagatggtcagtgccttgttggatgctgagccccccatactctattccgagtatgatcctac

cagacccttcagtgaagcttcgatgatgggcttactgaccaacctggcagacagggagctggttcacatgatcaact

gggcgaagagggtgccaggctttgtggatttgaccctccatgatcaggtccaccttctagaatgtgcctggctagag

atcctgatgattggtctcgtctggcgctccatggagcacccagggaagctactgtttgctcctaacttgctcttgga

caggaaccagggaaaatgtgtagagggcatggtggagatcttcgacatgctgctggctacatcatctcggttccgca

tgatgaatctgcagggagaggagtttgtgtgcctcaaatctattattttgcttaattctggagtgtacacatttctg

tccagcaccctgaagtctctggaagagaaggaccatatccaccgagtcctggacaagatcacagacactttgatcca

cctgatggccaaggcaggcctgaccctgcagcagcagcaccagcggctggcccagctcctcctcatcctctcccaca

tcaggcacatgagtaacaaaggcatggagcatctgtacagcatgaagtgcaagaacgtggtgcccctctcagatctg

ctgctggagatgctggacgcccaccgcctacatgcgcccactagccgtggaggggcatccgtggaggagacggacca

aagccacttggccactgcgggctctacttcatcgcattccttgcaaaagtattacatcacgggggaggcagagggtt

tccctgccacggtc

AR-V7

atggaagtgcagttagggctgggaagggtctaccctcggccgccgtccaagacctaccgaggagctttccagaatct

gttccagagcgtgcgcgaagtgatccagaacccgggccccaggcacccagaggccgcgagcgcagcacctcccggcg

ccagtttgctgctgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag

cagcagcagcagcagcaagagactagccccaggcagcagcagcagcagcagggtgaggatggttctccccaagccca

tcgtagaggccccacaggctacctggtcctggatgaggaacagcaaccttcacagccgcagtcggccctggagtgcc

accccgagagaggttgcgtcccagagcctggagccgccgtggccgccagcaaggggctgccgcagcagctgccagca

cctccggacgaggatgactcagctgccccatccacgttgtccctgctgggccccactttccccggcttaagcagctg

ctccgctgaccttaaagacatcctgagcgaggccagcaccatgcaactccttcagcaacagcagcaggaagcagtat

ccgaaggcagcagcagcgggagagcgagggaggcctcgggggctcccacttcctccaaggacaattacttaggggc

acttcgaccatttctgacaacgccaaggagttgtgtaaggcagtgtcggtgtccatgggcctgggtgtggaggcgtt

ggagcatctgagtccaggggaacagcttcgggggggattgcatgtacgccccacttttgggagttccacccgctgtgc

gtcccactccttgtgccccattggccgaatgcaaaggttctctgctagacgacagcgcaggcaagagcactgaagat

actgctgagtattccccttcaagggaggttacaccaaagggctagaaggcgagagcctaggctgctctggcagcgc

tgcagcagggagctccgggacacttgaactgccgtctaccctgtctctctacaagtccggagcactggacgaggcag

ctgcgtaccagagtcgcgactactacaactttccactggctctggccggaccgccgccccctccgccgcctccccat

ccccacgctcgcatcaagctggagaacccgctggactacggcagcgcctgggcggctgcggcggcgcagtgccgcta

tggggacctggcgagcctgcatggcgcgggtgcagcgggacccggttctgggtcaccctcagccgccgcttcctcat

cctggcacactctcttcacagccgaagaaggccagttgtatggaccgtgtggtggtggtgggggtggtggcggcggc

ggcggcggcggcggcggcggcggcggcgaggcgggagctgtagccccctacggctacactcggcccccctcaggg

gctggcgggccaggaaagcgacttcaccgcacctgatgtgtggtaccctggcggcatggtgagcagagtgccctatc

ccagtcccacttgtgtcaaaagcgaaatgggcccctggatggatagctactccggaccttacggggacatgcgtttg

gagactgccagggaccatgttttgcccattgactattactttccaccccagaagacctgcctgatctgtggagatga

agcttctgggtgtcactatggagctctcacatgtggaagctgcaaggtcttcttcaaaagagccgctgaagggaaac

agaagtacctgtgcgccagcagaaatgattgcactattgataaattccgaaggaaaaattgtccatcttgtcgtctt

cggaaatgttatgaagcagggatgactctgggagaaaaattccgggttggcaattgcaagcatctcaaaatgaccag

accc

FLAG-YAP2 (8SA)

atggactacaaagacgatgacgataaagcaaggctcgaatcggtacctaaggatcccgggcagcagccgccgcctca

accggcccccagggccaagggcagccgccttcgcagcccccgcaggggcagggcccgccgtccggacccgggcaac

cggcacccgcggcgacccaggcggcgccgcaggcacccccccgccgggcatcagatcgtgcacgtccgcggggacgcg

gagaccgacctggaggcgctcttcaacgccgtcatgaaccccaagacggccaacgtgccccagaccgtgcccatgag

gctccggaagctgcccgactccttcttcaagccgccggagcccaaatcccactcccgacaggccgctactgatgcag

gcactgcaggagccctgactccacagcatgttcgagctcatgccgctccagctgctctgcagttgggagctgtttct

cctgggacactgacccccactggagtagtctctggcccagcagctacacccacagctcagcatcttcgacaggctgc

ttttgagatacctgatgatgtacctctgccagcaggttgggagatggcaaagacatcttctggtcagagatacttct

taaatcacatcgatcagacaacaacatggcaggaccccaggaaggccatgctgtcccagatgaacgtcacagccccc

accagtccaccagtgcagcagaatatgatgaactcggcttcaggtcctcttcctgatggatgggaacaagccatgac

tcaggatggagaaatttactatataaaccataagaacaagaccacctcttggctagacccaaggcttgaccctcgtt

ttgccatgaaccagagaatcagtcagagtgctccagtgaaacagccaccacccctggctccccagagcccacaggga

ggcgtcatgggtggcagcaactccaaccagcagcaacagatgcgactgcagcaactgcagatggagaaggagaggct

gcggctgaaacagcaagaactgcttcggcaggagttagccctgtgtagccagttaccaacactggagcaggatggtg

ggactcaaaatccagtgtcttctcccgggatgtctcaggaattgagaacaatgacgaccaatagctcagatcctttc

cttaacagtggcacctatcactctcgagatgaggctacagacagtggactaagcatgagcagctacagtgtccctcg

aaccccagatgacttcctgaacagtgtggatgagatggatacaggtgatactatcaaccaaagcaccctgccctcac

agcagaaccgtttcccagactaccttgaagccattcctgggacaaatgtggaccttggaacactggaaggagatgga

atgaacatagaaggagaggagctgatgccaagtctgcaggaagctttgagttctgacatccttaatgacatggagtc

tgttttggctgccaccaagctagataaagaaagctttcttacatggtta

p53 (R175H)

atggaggagccgcagtcagatcctagcgtcgagccccctctgagtcaggaaacattttcagacctatggaaactact

tcctgaaaacaacgttctgtccccttgccgtcccaagcaatggatgatttgatgctgtccccggacgatattgaac

aatggttcactgaagacccaggtccagatgaagctcccagaatgccagaggctgctccccgcgtggcccctgcacca

gcagctcctacaccggcggcccctgcaccagccccctcctggcccctgtcatcttctgtcccttcccagaaaacctа

ccagggcagctacggtttccgtctgggcttcttgcattctgggacagccaagtctgtgacttgcacgtactcccctg

ccctcaacaagatgttttgccaactggccaagacctgccctgtgcagctgtgggttgattccacacccccgcccggc

acccgcgtccgcgccatggccatctacaagcagtcacagcacatgacggaggttgtgaggcactgcccccaccatga

gcgctgctcagatagcgatggtctggcccctcctcagcatcttatccgagtggaaggaaatttgcgtgtggagtatt

tggatgacagaaacacttttcgacatagtgtggtggtgccctatgagccgcctgaggttggctctgactgtaccacc

atccactacaactacatgtgtaacagttcctgcatgggcggcatgaaccggaggcccatcctcaccatcatcacact

ggaagactccagtggtaatctactgggacggaacagctttgaggtgcgtgtttgtgcctgtcctgggagagaccggc

gcacagaggaagagaatctccgcaagaaaggggagcctcaccacgagctgcccccagggagcactaagcgagcactg

cccaacaacaccagctcctctccccagccaaagaagaaaccactggatggagaatatttcacccttcaggaccagac

cagctttcaaaaagaaaattgt

Hras (G12V, E37G)

atgacggaatataagctggtggtggtgggcgccgtcggtgtgggcaagagtgcgctgaccatccagctgatccagaa
ccattttgtggacgaatacgaccccactataggggattcctaccggaagcaggtggtcattgatggggagacgtgcc
tgttggacatcctggataccgccggccaggaggagtacagcgccatgcgggaccagtacatgcgcaccggggagggc
ttcctgtgtgtgtttgccatcaacaacaccaagtcttttgaggacatccaccagtacagggagcagatcaaacgggt
gaaggactcggatgacgtgcccatggtgctggtggggaacaagtgtgacctggctgcacgcactgtggaatctcggc
aggctcaggacctcgcccgaagctacggcatcccctacatcgagacctcggccaagacccggcagggagtggaggat
gccttctacacgttggtgcgtgagatccggcagcacaagctgcggaagctgaaccctcctgatgagagtggccccgg
ctgcatgagctgcaagtgtgtgctctcc

Rgl2-CAAX

atgctcccgcggcccctgcggctgcttttggacacgacccccccgggggagtcgtgctgagcagcttccggagccg
ggaccccgaagagggtggggacccaggtggccgggccgtgggcgggggggcaggaggaagaggatgaggaagaagaag
aggcttctgtgtcagtctgggacgaggaggaggatggtgcgacctttactgtcacaagccgccagtacaggcctctt
gacccctggctcccttgcctccacctcgctcctcccgacggctccgcgctggcactctggaggccctggtcagaca
cctcttggatgccaggacagcaggggctgacatgatgttcactccggccttgctggccacccaccgggccttcacct
ccactcctgccctgtttgggcttgtggctgacaggctggaagcccttgaatctcatcctcccggtgagctagagagg
accacaggggtagccatctctgtactttcaacttggctggcctctcaccctgaggattttggctctgaggtcaaggg
tcaacttgaccggcttgagagcttcttgcttcggacagggtatgcagcacgggagggtgttgtgggggggcagtgctg
acctcatccgaaacctccgggcccgggtggacccccgggcccccgaccttcctaagcccctggcccttcctggcgat
tcccctgctgaccccacggatgtcctggtgttcctcgctgaccacttggccgaacagctgaccctgctagatgcgga
actgtttcttaatctgatcccctctcagtgtttgggaggcctctggggtcacagagaccggccaggacattctcacc
tctgcccgtctgtccgagctaccgtcacacagttcaacaaggtggcgggggcggtagttagctctgtcttgggggcc
acctcaattggagaggggccaagagaggtgactgtgagaccactgcggcccccacagagggcccggctcctagagaa
gtggatccgtgtggccgaggagtgccgcctgcttcggaacttctcctcagtgtatgctgttgtgtccgctctgcagt
ccagccctatccacaggcttcgggcagcctgggggggagacaaccagggacagcctccgagtcttttccagcctgtgc
cagattttctcagaggaggataattattcccagagcagagagctcctcatgcaggaagtgaagccgcagcccctgt
ggagccacactccaagaaggccccaaggtctggcttcaggggtgggggtgtggttccctacctgggaaccttcctga

aggacctcgtgatgctggatgcggcctccaaggatgagctggagaatggctacatcaattttgacaagcggaggaag

gagtttgctatcctttcggagctgttgcgcctccagaaagaatgtcgtggctacgacctccgacctaactctgatat

ccagcaatggctccagggcctccagcctttaactgaagctcagagtcaccgtgtatcctgtgaagtggagccaccag

ggaccagtgactcccccgctgcaaggacacctcggccaacactagtgatcacacagtggacggaagttctgggctct

gttggaggccccactccgcttgtgtcctgggatcggcccagtgtcgggggagatgaggtgcctggaaccccagcacc

tctgctgactcgcctcgcccagcacatgaagtggccatcagtctcatctctggactctgccctggaaagcagcccct

ccttgcacagccctgctgaccctggccacctctctcctccagcctcctcccctaggccttcccggggtcaccgtcgc

tcagcctcctgtgggtctccgttgagtggaaacacaggagaagggacctctaggagtgctggatgtgggggcggggt

atctgggccagggtcctctgattgccgaatcatccgagtccagatggagctgggggaggatggcagcgtctacaaga

gcatcctggtgacaagccaggacaaagctccaagtgttattagtcgtgtccttaagaaaaacaatcgtgattctgct

gtggcttcagagttcgagctggtgcagctgctacctggggatcgagagctgaccatcccacactcagctaacgtctt

ctatgccatggatggtgcatctcatgacttcctcctgcggcagcgcagaagaccctctgctgccaccccgggttccc

acagcggcccctctgcctcaggaactcctccgagcgaggggggaggggctcctttcccaggatcaaggccacgggg

aggaagattgcacgggcactgttcacgcgtctaagcaaagatggtaaaaagaagaaaaagaagtcaaagacaaagtg

tgtaattatg

RalA (G23V, mature peptide)

atggctgcaaataagcccaagggtcaaaattctttggccttacacaaagtcatcatggtgggcagtgttggtgtggg

caagtcagctctgactctacagttcatgtacgatgagtttgtggaggactatgagcctaccaaagcagacagctacc

ggaagaaggtagtgctggatggggaggaagtacagattgatatcttagatacagctgggcaggaggactatgctgca

attagagacaactacttccgaagcggggagggcttcctctgtgtcttctctattacagaaatggaatcctttgcagc

tacagctgacttcagggagcagattttaagagtaaaagaagacgagaatgttccatttctactggttggtaacaaat

cagatttagaagataaaaggcaggtttctgtagaagaggcaaaaaacagagctgaccagtggaatgttaactacgtg

gaaacgtctgctaaaacacgagctaatgttgacaaggtatttttgatttaatgagagaaattcgagcaagaaagat

ggaagacagcaaagaaaagaatggaaaaaagaagaggaaaagtttagccaagagaatcagagaaagatgc

RalA (G23V, full)

atggctgcaaataagcccaagggtcaaaattctttggccttacacaaagtcatcatggtgggcagtgttggtgtggg

caagtcagctctgactctacagttcatgtacgatgagtttgtggaggactatgagcctaccaaagcagacagctacc

ggaagaaggtagtgctggatggggaggaagtacagattgatatcttagatacagctgggcaggaggactatgctgca

attagagacaactacttccgaagcggggagggcttcctctgtgtcttctctattacagaaatggaatcctttgcagc

tacagctgacttcagggagcagattttaagagtaaaagaagacgagaatgttccatttctactggttggtaacaaat

cagatttagaagataaaaggcaggtttctgtagaagaggcaaaaaacagagctgaccagtggaatgttaactacgtg

gaaacgtctgctaaaacacgagctaatgttgacaaggtatttttttgatttaatgagagaaattcgagcaagaaagat

ggaagacagcaaagaaaagaatggaaaaaagaagaggaaaagtttagccaagagaatcagagaaagatgctgcattt

ta

SV40 Large T Antigen

atggataaagttttaaacagagaggaatctttgcagctaatggaccttctaggtcttgaaaggagtgcctgggggaa

tattcctctgatgagaaaggcatatttaaaaaaatgcaaggagtttcatcctgataaaggaggagatgaagaaaaaa

tgaagaaaatgaatactctgtacaagaaaatggaagatggagtaaaatatgctcatcaacctgactttggaggcttc

tgggatgcaactgagattccaacctatggaactgatgaatgggagcagtggtggaatgcctttaatgaggaaaacct

gttttgctcagaagaaatgccatctagtgatgatgaggctactgctgactctcaacattctactcctccaaaaaaga

agagaaaggtagaagaccccaaggactttccttcagaattgctaagtttttgagtcatgctgtgtttagtaataga

actcttgcttgctttgctatttacaccacaaaggaaaaagctgcactgctatacaagaaaattatggaaaaatattc

tgtaacctttataagtaggcataacagttataatcataacatactgttttttcttactccacacaggcatagagtgt

ctgctattaataactatgctcaaaaattgtgtacctttagctttttaatttgtaaaggggttaataaggaatatttg

atgtatagtgccttgactagagatccattttctgttattgaggaaagtttgccaggtgggttaaaggagcatgattt

taatccagaagaagcagaggaaactaaacaagtgtcctggaagcttgtaacagagtatgcaatggaaacaaaatgtg

atgatgtgttgttattgcttgggatgtacttggaatttcagtacagttttgaaatgtgtttaaaatgtattaaaaaa

gaacagcccagccactataagtaccatgaaaagcattatgcaaatgctgctatatttgctgacagcaaaaaccaaaa

aaccatatgccaacaggctgttgatactgttttagctaaaaagcgggttgatagcctacaattaactagagaacaaa

tgttaacaaacagatttaatgatcttttggataggatggatataatgtttggttctacaggctctgctgacatagaa

gaatggatggctggagttgcttggctacactgtttgttgcccaaaatggattcagtggtgtatgacttttaaaatg

catggtgtacaacattcctaaaaaaagatactggctgtttaaaggaccaattgatagtggtaaaactacattagcag

ctgctttgcttgaattatgtgggggaaagctttaaatgttaatttgcccttggacaggctgaactttgagctagga

gtagctattgaccagtttttagtagttttgaggatgtaaagggcactggaggggagtccagagatttgccttcagg

tcagggaattaataacctggacaatttaagggattatttggatggcagtgttaaggtaaacttagaaaagaaacacc

taaataaaagaactcaaatatttccccctggaatagtcaccatgaatgagtacagtgtgcctaaaacactgcaggcc

agatttgtaaaacaaatagattttaggcccaaagattatttaaagcattgcctggaacgcagtgagtttttgttaga

aaagagaataattcaaagtggcattgctttgcttcttatgttaatttggtacagacctgtggctgagtttgctcaaa

gtattcagagcagaattgtggagtggaaagagagattggacaaagagtttagtttgtcagtgtatcaaaaaatgaag

tttaatgtggctatgggaattggagttttagattggctaagaaacagtgatgatgatgatgaagacagccaggaaaa

tgctgataaaaatgaagatggtggggagaagaacatggaagactcagggcatgaaacaggcattgattcacagtccc

aaggctcatttcaggcccctcagtcctcacagtctgttcatgatcataatcagccataccacatttgtagaggtttt

acttgctttaaaaaacctcccacacctccccctgaacctgaaaca

hTERT

atgccgcgcgctccccgctgccgagccgtgcgctccctgctgcgcagccactaccgcgaggtgctgccgctggccac

gttcgtgcggcgcctggggccccagggctggcggctggtgcagcgcggggacccggcggctttccgcgcgctggtgg

cccagtgcctggtgtgcgtgccctgggacgcacggccgccccccgccgccccctccttccgccaggtgtcctgcctg

aaggagctggtggcccgagtgctgcagaggctgtgcgagcgcggcgcgaagaacgtgctggccttcggcttcgcgct

gctggacggggcccgcgggggcccccccgaggccttcaccaccagcgtgcgcagctacctgcccaacacggtgaccg

acgcactgcgggggagcggggcgtgggggctgctgctgcgccgcgtgggcgacgacgtgctggttcacctgctggca

cgctgcgcgctctttgtgctggtggctcccagctgcgcctaccaggtgtgcgggccgccgctgtaccagctcggcgc

tgccactcaggcccggcccccgccacacgctagtggaccccgaaggcgtctgggatgcgaacgggcctggaaccata

gcgtcagggaggccggggtccccctgggcctgccagccccgggtgcgaggaggcgcgggggcagtgccagccgaagt

ctgccgttgcccaagaggcccaggcgtggcgctgcccctgagccggagcggacgcccgttgggcaggggtcctgggc

ccacccgggcaggacgcgtggaccgagtgaccgtggtttctgtgtggtgtcacctgccagacccgccgaagaagcca

cctctttggagggtgcgctctctggcacgcgccactcccacccatccgtgggccgccagcaccacgcgggcccccca

tccacatcgcggccaccacgtccctgggacacgccttgtcccccggtgtacgccgagaccaagcacttcctctactc

ctcaggcgacaaggagcagctgcggccctccttcctactcagctctctgaggcccagcctgactggcgctcggaggc

tcgtggagaccatctttctgggttccaggccctggatgccagggactccccgcaggttgccccgcctgccccagcgc

tactggcaaatgcggcccctgtttctggagctgcttgggaaccacgcgcagtgcccctacggggtgctcctcaagac

gcactgcccgctgcgagctgcggtcaccccagcagccggtgtctgtgcccgggagaagccccagggctctgtggcgg

cccccgaggaggaggacacagacccccgtcgcctggtgcagctgctccgccagcacagcagcccctggcaggtgtac

ggcttcgtgcgggcctgcctgcgccggctggtgcccccaggcctctggggctccaggcacaacgaacgccgcttcct

caggaacaccaagaagttcatctccctggggaagcatgccaagctctcgctgcaggagctgacgtggaagatgagcg

tgcggggctgcgcttggctgcgcaggagcccaggggttggctgtgttccggccgcagagcaccgtctgcgtgaggag

atcctggccaagttcctgcactggctgatgagtgtgtacgtcgtcgagctgctcaggtctttcttttatgtcacgga

gaccacgtttcaaaagaacaggctctttttctaccggaagagtgtctggagcaagttgcaaagcattggaatcagac

agcacttgaagagggtgcagctgcgggagctgtcggaagcagaggtcaggcagcatcgggaagccaggcccgccctg

ctgacgtccagactccgcttcatccccaagcctgacgggctgcggccgattgtgaacatggactacgtcgtgggagc

cagaacgttccgcagagaaaagagggccgagcgtctcacctcgagggtgaaggcactgttcagcgtgctcaactacg

agcgggcgcggcgccccggcctcctgggcgcctctgtgctgggcctggacgatatccacagggcctggcgcaccttc

gtgctgcgtgtgcgggcccaggacccgccgcctgagctgtactttgtcaaggtggatgtgacgggcgcgtacgacac

catcccccaggacaggctcacggaggtcatcgccagcatcatcaaaccccagaacacgtactgcgtgcgtcggtatg

ccgtggtccagaaggccgcccatgggcacgtccgcaaggccttcaagagccacgtctctaccttgacagacctccag

ccgtacatgcgacagttcgtggctcacctgcaggagaccagcccgctgagggatgccgtcgtcatcgagcagagctc

ctccctgaatgaggccagcagtggcctcttcgacgtcttcctacgcttcatgtgccaccacgccgtgcgcatcaggg

gcaagtcctacgtccagtgccaggggatcccgcagggctccatcctctccacgctgctctgcagcctgtgctacggc

gacatggagaacaagctgtttgcggggattcggcgggacgggctgctcctgcgtttggtggatgatttcttgttggt

gacacctcacctcacccacgcgaaaaccttcctcaggaccctggtccgaggtgtccctgagtatggctgcgtggtga

acttgcggaagacagtggtgaacttccctgtagaagacgaggccctgggtggcacggcttttgttcagatgccggcc

cacggcctattcccctggtgcggcctgctgctggatacccggaccctggaggtgcagagcgactactccagctatgc

ccggacctccatcagagccagtctcaccttcaaccgcggcttcaaggctgggaggaacatgcgtcgcaaactctttg

gggtcttgcggctgaagtgtcacagcctgtttctggatttgcaggtgaacagcctccagacggtgtgcaccaacatc

tacaagatcctcctgctgcaggcgtacaggtttcacgcatgtgtgctgcagctcccatttcatcagcaagtttggaa

gaaccccacatttttcctgcgcgtcatctctgacacggcctccctctgctactccatcctgaaagccaagaacgcag

ggatgtcgctgggggccaagggcgccgccggccctctgccctccgaggccgtgcagtggctgtgccaccaagcattc

ctgctcaagctgactcgacaccgtgtcacctacgtgccactcctggggtcactcaggacagcccagacgcagctgag

tcggaagctcccggggacgacgctgactgccctggaggccgcagccaacccggcactgccctcagacttcaagacca

tcctggac

CCNB1

ATGGCGCTCCGAGTCACCAGGAACTCGAAAATTAATGCTGAAAATAAGGCGAAGATCAACATGGCAGGCGCAAAGCG

CGTTCCTACGGCCCCTGCTGCAACCTCCAAGCCCGGACTGAGGCCAAGAACAGCTCTTGGGGACATTGGTAACAAAG

TCAGTGAACAACTGCAGGCCAAAATGCCTATGAAGAAGGAAGCAAAACCTTCAGCTACTGGAAAAGTCATTGATAAA

AAACTACCAAAACCTCTTGAAAAGGTACCTATGCTGGTGCCAGTGCCAGTGTCTGAGCCAGTGCCAGAGCCAGAACC

TGAGCCAGAACCTGAGCCTGTTAAAGAAGAAAAACTTTCGCCTGAGCCTATTTTGGTTGATACTGCCTCTCCAAGCC

CAATGGAAACATCTGGATGTGCCCCTGCAGAAGAAGACCTGTGTCAGGCTTTCTCTGATGTAATTCTTGCAGTAAAT

GATGTGGATGCAGAAGATGGAGCTGATCCAAACCTTTGTAGTGAATATGTGAAAGATATTTATGCTTATCTGAGACA

ACTTGAGGAAGAGCAAGCAGTCAGACCAAAATACCTACTGGGTCGGGAAGTCACTGGAAACATGAGAGCCATCCTAA

TTGACTGGCTAGTACAGGTTCAAATGAAATTCAGGTTGTTGCAGGAGACCATGTACATGACTGTCTCCATTATTGAT

CGGTTCATGCAGAATAATTGTGTGCCCAAGAAGATGCTGCAGCTGGTTGGTGTCACTGCCATGTTTATTGCAAGCAA

ATATGAAGAAATGTACCCTCCAGAAATTGGTGACTTTGCTTTTGTGACTGACAACACTTATACTAAGCACCAAATCA

GACAGATGGAAATGAAGATTCTAAGAGCTTTAAACTTTGGTCTGGGTCGGCCTCTACCTTTGCACTTCCTTCGGAGA

GCATCTAAGATTGGAGAGGTTGATGTCGAGCAACATACTTTGGCCAAATACCTGATGGAACTAACTATGTTGGACTA

TGACATGGTGCACTTTCCTCCTTCTCAAATTGCAGCAGGAGCTTTTTGCTTAGCACTGAAAATTCTGGATAATGGTG

AATGGACACCAACTCTACAACATTACCTGTCATATACTGAAGAATCTCTTCTTCCAGTTATGCAGCACCTGGCTAAG

AATGTAGTCATGGTAAATCAAGGACTTACAAAGCACATGACTGTCAAGAACAAGTATGCCACATCGAAGCATGCTAA

GATCAGCACTCTACCACAGCTGAATTCTGCACTAGTTCAAGATTTAGCCAAGGCTGTGGCAAAGGTG

CCND1

ATGGAACACCAGCTCCTGTGCTGCGAAGTGGAAACCATCCGCCGCGCGTACCCCGATGCCAACCTCCTCAACGACCG

GGTGCTGCGGGCCATGCTGAAGGCGGAGGAGACCTGCGCGCCCTCGGTGTCCTACTTCAAATGTGTGCAGAAGGAGG

TCCTGCCGTCCATGCGGAAGATCGTCGCCACCTGGATGCTGGAGGTCTGCGAGGAACAGAAGTGCGAGGAGGAGGTC

TTCCCGCTGGCCATGAACTACCTGGACCGCTTCCTGTCGCTGGAGCCCGTGAAAAAGAGCCGCCTGCAGCTGCTGGG

GGCCACTTGCATGTTCGTGGCCTCTAAGATGAAGGAGACCATCCCCCTGACGGCCGAGAAGCTGTGCATCTACACCG

ACAACTCCATCCGGCCCGAGGAGCTGCTGCAAATGGAGCTGCTCCTGGTGAACAAGCTCAAGTGGAACCTGGCCGCA

ATGACCCCGCACGATTTCATTGAACACTTCCTCTCCAAAATGCCAGAGGCGGAGGAGAACAAACAGATCATCCGCAA

ACACGCGCAGACCTTCGTTGCCCTCTGTGCCACAGATGTGAAGTTCATTTCCAATCCGCCCTCCATGGTGGCAGCGG

GGAGCGTGGTGGCCGCAGTGCAAGGCCTGAACCTGAGGAGCCCCAACAACTTCCTGTCCTACTACCGCCTCACACGC

TTCCTCTCCAGAGTGATCAAGTGTGACCCGGACTGCCTCCGGGCCTGCCAGGAGCAGATCGAAGCCCTGCTGGAGTC

AAGCCTGCGCCAGGCCCAGCAGAACATGGACCCCAAGGCCGCCGAGGAGGAGGAAGAGGAGGAGGAGGAGGTGGACC

TGGCTTGCACACCCACCGACGTGCGGGACGTGGACATC

CDK1

ATGGAAGATTATACCAAAATAGAGAAAATTGGAGAAGGTACCTATGGAGTTGTGTATAAGGGTAGACACAAAACTAC

AGGTCAAGTGGTAGCCATGAAAAAAATCAGACTAGAAAGTGAAGAGGAAGGGGTTCCTAGTACTGCAATTCGGGAAA

TTTCTCTATTAAAGGAACTTCGTCATCCAAATATAGTCAGTCTTCAGGATGTGCTTATGCAGGATTCCAGGTTATAT

CTCATCTTTGAGTTTCTTTCCATGGATCTGAAGAAATACTTGGATTCTATCCCTCCTGGTCAGTACATGGATTCTTC

ACTTGTTAAGAGTTATTTATACCAAATCCTACAGGGGATTGTGTTTTGTCACTCTAGAAGAGTTCTTCACAGAGACT

TAAAACCTCAAAATCTCTTGATTGATGACAAAGGAACAATTAAACTGGCTGATTTTGGCCTTGCCAGAGCTTTTGGA

ATACCTATCAGAGTATATACACATGAGGTAGTAACACTCTGGTACAGATCTCCAGAAGTATTGCTGGGGTCAGCTCG

TTACTCAACTCCAGTTGACATTTGGAGTATAGGCACCATATTTGCTGAACTAGCAACTAAGAAACCACTTTTCCATG

GGGATTCAGAAATTGATCAACTCTTCAGGATTTTCAGAGCTTTGGGCACTCCCAATAATGAAGTGTGGCCAGAAGTG

GAATCTTTACAGGACTATAAGAATACATTTCCCAAATGGAAACCAGGAAGCCTAGCATCCCATGTCAAAAACTTGGA

TGAAAATGGCTTGGATTTGCTCTCGAAAATGTTAATCTATGATCCAGCCAAACGAATTTCTGGCAAAATGGCACTGA

ATCATCCATATTTTAATGATTTGGACAATCAGATTAAGAAGATG

CDK1 (T14A, Y15F)

ATGGAAGATTATACCAAAATAGAGAAAATTGGAGAAGGTGCCTTTGGAGTTGTGTATAAGGGTAGACACAAAACTAC

AGGTCAAGTGGTAGCCATGAAAAAAATCAGACTAGAAAGTGAAGAGGAAGGGGTTCCTAGTACTGCAATTCGGGAAA

TTTCTCTATTAAAGGAACTTCGTCATCCAAATATAGTCAGTCTTCAGGATGTGCTTATGCAGGATTCCAGGTTATAT

CTCATCTTTGAGTTTCTTTCCATGGATCTGAAGAAATACTTGGATTCTATCCCTCCTGGTCAGTACATGGATTCTTC

ACTTGTTAAGAGTTATTTATACCAAATCCTACAGGGGATTGTGTTTTGTCACTCTAGAAGAGTTCTTCACAGAGACT

TAAAACCTCAAAATCTCTTGATTGATGACAAAGGAACAATTAAACTGGCTGATTTTGGCCTTGCCAGAGCTTTTGGA

ATACCTATCAGAGTATATACACATGAGGTAGTAACACTCTGGTACAGATCTCCAGAAGTATTGCTGGGGTCAGCTCG

TTACTCAACTCCAGTTGACATTTGGAGTATAGGCACCATATTTGCTGAACTAGCAACTAAGAAACCACTTTTCCATG

GGGATTCAGAAATTGATCAACTCTTCAGGATTTTCAGAGCTTTGGGCACTCCCAATAATGAAGTGTGGCCAGAAGTG

GAATCTTTACAGGACTATAAGAATACATTTCCCAAATGGAAACCAGGAAGCCTAGCATCCCATGTCAAAAACTTGGA

TGAAAATGGCTTGGATTTGCTCTCGAAAATGTTAATCTATGATCCAGCCAAACGAATTTCTGGCAAAATGGCACTGA

ATCATCCATATTTTAATGATTTGGACAATCAGATTAAGAAGATG

CDK4

ATGGCTACCTCTCGATATGAGCCAGTGGCTGAAATTGGTGTCGGTGCCTATGGGACAGTGTACAAGGCCCGTGATCC

CCACAGTGGCCACTTTGTGGCCCTCAAGAGTGTGAGAGTCCCCAATGGAGGAGGAGGTGGAGGAGGCCTTCCCATCA

GCACAGTTCGTGAGGTGGCTTTACTGAGGCGACTGGAGGCTTTTGAGCATCCCAATGTTGTCCGGCTGATGGACGTC

TGTGCCACATCCCGAACTGACCGGGAGATCAAGGTAACCCTGGTGTTTGAGCATGTAGACCAGGACCTAAGGACATA

TCTGGACAAGGCACCCCCACCAGGCTTGCCAGCCGAAACGATCAAGGATCTGATGCGCCAGTTTCTAAGAGGCCTAG

ATTTCCTTCATGCCAATTGCATCGTTCACCGAGATCTGAAGCCAGAGAACATTCTGGTGACAAGTGGTGGAACAGTC

AAGCTGGCTGACTTTGGCCTGGCCAGAATCTACAGCTACCAGATGGCACTTACACCCGTGGTTGTTACACTCTGGTA

CCGAGCTCCCGAAGTTCTTCTGCAGTCCACATATGCAACACCTGTGGACATGTGGAGTGTTGGCTGTATCTTTGCAG

AGATGTTTCGTCGAAAGCCTCTCTTCTGTGGAAACTCTGAAGCCGACCAGTTGGGCAAAATCTTTGACCTGATTGGG

CTGCCTCCAGAGGATGACTGGCCTCGAGATGTATCCCTGCCCCGTGGAGCCTTTCCCCCCAGAGGGCCCCGCCCAGT

GCAGTCGGTGGTACCTGAGATGGAGGAGTCGGGAGCACAGCTGCTGCTGGAAATGCTGACTTTTAACCCACACAAGC

GAATCTCTGCCTTTCGAGCTCTGCAGCACTCTTATCTACATAAGGATGAAGGTAATCCGGAG

CDK4 (R24C)

ATGGCTACCTCTCGATATGAGCCAGTGGCTGAAATTGGTGTCGGTGCCTATGGGACAGTGTACAAGGCCTGTGATCC

CCACAGTGGCCACTTTGTGGCCCTCAAGAGTGTGAGAGTCCCCAATGGAGGAGGAGGTGGAGGAGGCCTTCCCATCA

GCACAGTTCGTGAGGTGGCTTTACTGAGGCGACTGGAGGCTTTTGAGCATCCCAATGTTGTCCGGCTGATGGACGTC

TGTGCCACATCCCGAACTGACCGGGAGATCAAGGTAACCCTGGTGTTTGAGCATGTAGACCAGGACCTAAGGACATA

TCTGGACAAGGCACCCCCACCAGGCTTGCCAGCCGAAACGATCAAGGATCTGATGCGCCAGTTTCTAAGAGGCCTAG

ATTTCCTTCATGCCAATTGCATCGTTCACCGAGATCTGAAGCCAGAGAACATTCTGGTGACAAGTGGTGGAACAGTC

AAGCTGGCTGACTTTGGCCTGGCCAGAATCTACAGCTACCAGATGGCACTTACACCCGTGGTTGTTACACTCTGGTA

CCGAGCTCCCGAAGTTCTTCTGCAGTCCACATATGCAACACCTGTGGACATGTGGAGTGTTGGCTGTATCTTTGCAG

AGATGTTTCGTCGAAAGCCTCTCTTCTGTGGAAACTCTGAAGCCGACCAGTTGGGCAAAATCTTTGACCTGATTGGG

CTGCCTCCAGAGGATGACTGGCCTCGAGATGTATCCCTGCCCCGTGGAGCCTTTCCCCCCAGAGGGCCCCGCCCAGT

GCAGTCGGTGGTACCTGAGATGGAGGAGTCGGGAGCACAGCTGCTGCTGGAAATGCTGACTTTTAACCCACACAAGC

GAATCTCTGCCTTTCGAGCTCTGCAGCACTCTTATCTACATAAGGATGAAGGTAATCCGGAG

CDK6

ATGGAGAAGGACGGCCTGTGCCGCGCTGACCAGCAGTACGAATGCGTGGCGGAGATCGGGGAGGGCGCCTATGGGAA

GGTGTTCAAGGCCCGCGACTTGAAGAACGGAGGCCGTTTCGTGGCGTTGAAGCGCGTGCGGGTGCAGACCGGCGAGG

AGGGCATGCCGCTCTCCACCATCCGCGAGGTGGCGGTGCTGAGGCACCTGGAGACCTTCGAGCACCCCAACGTGGTC

AGGTTGTTTGATGTGTGCACAGTGTCACGAACAGACAGAGAAACCAAACTAACTTTAGTGTTTGAACATGTCGATCA

AGACTTGACCACTTACTTGGATAAAGTTCCAGAGCCTGGAGTGCCCACTGAAACCATAAAGGATATGATGTTTCAGC

TTCTCCGAGGTCTGGACTTTCTTCATTCACACCGAGTAGTGCATCGCGATCTAAAACCACAGAACATTCTGGTGACC

AGCAGCGGACAAATAAAACTCGCTGACTTCGGCCTTGCCCGCATCTATAGTTTCCAGATGGCTCTAACCTCAGTGGT

CGTCACGCTGTGGTACAGAGCACCCGAAGTCTTGCTCCAGTCCAGCTACGCCACCCCCGTGGATCTCTGGAGTGTTG

GCTGCATATTTGCAGAAATGTTTCGTAGAAAGCCTCTTTTTCGTGGAAGTTCAGATGTTGATCAACTAGGAAAAATC

TTGGACGTGATTGGACTCCCAGGAGAAGAAGACTGGCCTAGAGATGTTGCCCTTCCCAGGCAGGCTTTTCATTCAAA

ATCTGCCCAACCAATTGAGAAGTTTGTAACAGATATCGATGAACTAGGCAAAGACCTACTTCTGAAGTGTTTGACAT

TTAACCCAGCCAAAAGAATATCTGCCTACAGTGCCCTGTCTCACCCATACTTCCAGGACCTGGAAAGGTGCAAAGAA

AACCTGGATTCCCACCTGCCGCCCAGCCAGAACACCTCGGAGCTGAATACAGCC

RHOJ

ATGAACTGCAAAGAGGGAACTGACAGCAGCTGCGGCTGCAGGGGCAACGACGAGAAGAAGATGTTGAAGTGTGTGGT

GGTGGGGGACGGTGCCGTGGGGAAAACCTGCCTGCTGATGAGCTACGCCAACGACGCCTTCCCAGAGGAATACGTGC

CCACTGTGTTTGACCACTATGCAGTTACTGTGACTGTGGGAGGCAAGCAACACTTGCTCGGACTGTATGACACCGCG

GGACAGGAGGACTACAACCAGCTGAGGCCACTCTCCTACCCCAACACGGATGTGTTTTTGATCTGCTTCTCTGTCGT

AAACCCTGCCTCTTACCACAATGTCCAGGAGGAATGGGTCCCCGAGCTCAAGGACTGCATGCCTCACGTGCCTTATG

TCCTCATAGGGACCCAGATTGATCTCCGTGATGACCCAAAAACCTTGGCCCGTTTGCTGTATATGAAAGAGAAACCT

CTCACTTACGAGCATGGTGTGAAGCTCGCAAAAGCGATCGGAGCACAGTGCTACTTGGAATGTTCAGCTCTGACTCA

GAAAGGTCTCAAAGCGGTTTTTGATGAAGCAATCCTCACCATTTTCCACCCCAAGAAAAAGAAGAAACGCTGTTCTG

AGGGTCACAGCTGCTGTTCAATTATC

PTEN (C124S)

ATGACAGCCATCATCAAAGAGATCGTTAGCAGAAACAAAAGGAGATATCAAGAGGATGGATTCGACTTAGACTTGAC

CTATATTTATCCAAACATTATTGCTATGGGATTTCCTGCAGAAAGACTTGAAGGCGTATACAGGAACAATATTGATG

ATGTAGTAAGGTTTTTGGATTCAAAGCATAAAAACCATTACAAGATATACAATCTTTGTGCTGAAAGACATTATGAC

ACCGCCAAATTTAATTGCAGAGTTGCACAATATCCTTTTGAAGACCATAACCCACCACAGCTAGAACTTATCAAACC

CTTTTGTGAAGATCTTGACCAATGGCTAAGTGAAGATGACAATCATGTTGCAGCAATTCACTCCAAGGCTGGAAAGG

GACGAACTGGTGTAATGATATGTGCATATTTATTACATCGGGGCAAATTTTTAAAGGCACAAGAGGCCCTAGATTTC

TATGGGGAAGTAAGGACCAGAGACAAAAAGGGAGTAACTATTCCCAGTCAGAGGCGCTATGTGTATTATTATAGCTA

CCTGTTAAAGAATCATCTGGATTATAGACCAGTGGCACTGTTGTTTCACAAGATGATGTTTGAAACTATTCCAATGT

TCAGTGGCGGAACTTGCAATCCTCAGTTTGTGGTCTGCCAGCTAAAGGTGAAGATATATTCCTCCAATTCAGGACCC

ACACGACGGGAAGACAAGTTCATGTACTTTGAGTTCCCTCAGCCGTTACCTGTGTGTGGTGATATCAAAGTAGAGTT

CTTCCACAAACAGAACAAGATGCTAAAAAAGGACAAAATGTTTCACTTTTGGGTAAATACATTCTTCATACCAGGAC

CAGAGGAAACCTCAGAAAAAGTAGAAAATGGAAGTCTATGTGATCAAGAAATCGATAGCATTTGCAGTATAGAGCGT

GCAGATAATGACAAGGAATATCTAGTACTTACTTTAACAAAAAATGATCTTGACAAAGCAAATAAAGACAAAGCCAA

CCGATACTTTTCTCCAAATTTTAAGGTGAAGCTGTACTTCACAAAAACAGTAGAGGAGCCGTCAAATCCAGAGGCTA

GCAGTTCAACTTCTGTAACACCAGATGTTAGTGACAATGAACCTGATCATTATAGATATTCTGACACCACTGACTCT

GATCCAGAGAATGAACCTTTTGATGAAGATCAGCATACACAAATTACAAAAGTC

c-MYC

ATGCCCCTCAACGTTAGCTTCACCAACAGGAACTATGACCTCGACTACGACTCGGTGCAGCCGTATTTCTACTGCGA

CGAGGAGGAGAACTTCTACCAGCAGCAGCAGCAGAGCGAGCTGCAGCCCCCGGCGCCCAGCGAGGATATCTGGAAGA

AATTCGAGCTGCTGCCCACCCCGCCCCTGTCCCCTAGCCGCCGCTCCGGGCTCTGCTCGCCCTCCTACGTTGCGGTC

ACACCCTTCTCCCTTCGGGGAGACAACGACGGCGGTGGCGGGAGCTTCTCCACGGCCGACCAGCTGGAGATGGTGAC

CGAGCTGCTGGGAGGAGACATGGTGAACCAGAGTTTCATCTGCGACCCGGACGACGAGACCTTCATCAAAAACATCA

TCATCCAGGACTGTATGTGGAGCGGCTTCTCGGCCGCCGCCAAGCTCGTCTCAGAGAAGCTGGCCTCCTACCAGGCT

GCGCGCAAAGACAGCGGCAGCCCGAACCCCGCCCGCGGCCACAGCGTCTGCTCCACCTCCAGCTTGTACCTGCAGGA

TCTGAGCGCCGCCGCCTCAGAGTGCATCGACCCCTCGGTGGTCTTCCCCTACCCTCTCAACGACAGCAGCTCGCCCA

AGTCCTGCGCCTCGCAAGACTCCAGCGCCTTCTCTCCGTCCTCGGATTCTCTGCTCTCCTCGACGGAGTCCTCCCCG

CAGGGCAGCCCCGAGCCCCTGGTGCTCCATGAGGAGACACCGCCCACCACCAGCAGCGACTCTGAGGAGGAACAAGA

AGATGAGGAAGAAATCGATGTTGTTTCTGTGGAAAAGAGGCAGGCTCCTGGCAAAAGGTCAGAGTCTGGATCACCTT

CTGCTGGAGGCCACAGCAAACCTCCTCACAGCCCACTGGTCCTCAAGAGGTGCCACGTCTCCACACATCAGCACAAC

TACGCAGCGCCTCCCTCCACTCGGAAGGACTATCCTGCTGCCAAGAGGGTCAAGTTGGACAGTGTCAGAGTCCTGAG

ACAGATCAGCAACAACCGAAAATGCACCAGCCCCAGGTCCTCGGACACCGAGGAGAATGTCAAGAGGCGAACACACA

ACGTCTTGGAGCGCCAGAGGAGGAACGAGCTAAAACGGAGCTTTTTTGCCCTGCGTGACCAGATCCCGGAGTTGGAA

AACAATGAAAAGGCCCCCAAGGTAGTTATCCTTAAAAAAGCCACAGCATACATCCTGTCCGTCCAAGCAGAGGAGCA

AAAGCTCATTTCTGAAGAGGACTTGTTGCGGAAACGACGAGAACAGTTGAAACACAAACTTGAACAGCTACGGAACT

CTTGTGCG

KLF4

ATGGCTGTCAGCGACGCGCTGCTCCCATCTTTCTCCACGTTCGCGTCTGGCCCGGCGGGAAGGGAGAAGACACTGCG

TCAAGCAGGTGCCCCGAATAACCGCTGGCGGGAGGAGCTCTCCCACATGAAGCGACTTCCCCCAGTGCTTCCCGGCC

GCCCCTATGACCTGGCGGCGGCGACCGTGGCCACAGACCTGGAGAGCGGCGGAGCCGGTGCGGCTTGCGGCGGTAGC

AACCTGGCGCCCCTACCTCGGAGAGAGACCGAGGAGTTCAACGATCTCCTGGACCTGGACTTTATTCTCTCCAATTC

GCTGACCCATCCTCCGGAGTCAGTGGCCGCCACCGTGTCCTCGTCAGCGTCAGCCTCCTCTTCGTCGTCGCCGTCGA

GCAGCGGCCCTGCCAGCGCGCCCTCCACCTGCAGCTTCACCTATCCGATCCGGGCCGGGAACGACCCGGGCGTGGCG

CCGGGCGGCACGGGCGGAGGCCTCCTCTATGGCAGGGAGTCCGCTCCCCCTCCGACGGCTCCCTTCAACCTGGCGGA

CATCAACGACGTGAGCCCCTCGGGCGGCTTCGTGGCCGAGCTCCTGCGGCCAGAATTGGACCCGGTGTACATTCCGC

CGCAGCAGCCGCAGCCGCCAGGTGGCGGGCTGATGGGCAAGTTCGTGCTGAAGGCGTCGCTGAGCGCCCCTGGCAGC

GAGTACGGCAGCCCGTCGGTCATCAGCGTCAGCAAAGGCAGCCCTGACGGCAGCCACCCGGTGGTGGTGGCGCCCTA

CAACGGCGGGCCGCCGCGCACGTGCCCCAAGATCAAGCAGGAGGCGGTCTCTTCGTGCACCCACTTGGGCGCTGGAC

CCCCTCTCAGCAATGGCCACCGGCCGGCTGCACACGACTTCCCCCTGGGGCGGCAGCTCCCCAGCAGGACTACCCCG

ACCCTGGGTCTTGAGGAAGTGCTGAGCAGCAGGGACTGTCACCCTGCCCTGCCGCTTCCTCCCGGCTTCCATCCCCA

CCCGGGGCCCAATTACCCATCCTTCCTGCCCGATCAGATGCAGCCGCAAGTCCCGCCGCTCCATTACCAAGAGCTCA

TGCCACCCGGTTCCTGCATGCCAGAGGAGCCCAAGCCAAAGAGGGGAAGACGATCGTGGCCCCGGAAAAGGACCGCC

ACCCACACTTGTGATTACGCGGGCTGCGGCAAAACCTACACAAAGAGTTCCCATCTCAAGGCACACCTGCGAACCCA

CACAGGTGAGAAACCTTACCACTGTGACTGGGACGGCTGTGGATGGAAATTCGCCCGCTCAGATGAACTGACCAGGC

ACTACCGTAAACACACGGGGCACCGCCCGTTCCAGTGCCAAAAATGCGACCGAGCATTTTCCAGGTCGGACCACCTC

GCCTTACACATGAAGAGGCATTTT

POU5F1

ATGGCGGGACACCTGGCTTCAGATTTTGCCTTCTCGCCCCCTCCAGGTGGTGGAGGTGATGGGCCAGGGGGGCCGGA

GCCGGGCTGGGTTGATCCTCGGACCTGGCTAAGCTTCCAAGGCCCTCCTGGAGGGCCAGGAATCGGGCCGGGGGTTG

GGCCAGGCTCTGAGGTGTGGGGGATTCCCCCATGCCCCCCGCCGTATGAGTTCTGTGGGGGGATGGCGTACTGTGGG

CCCCAGGTTGGAGTGGGGCTAGTGCCCCAAGGCGGCTTGGAGACCTCTCAGCCTGAGGGCGAAGCAGGAGTCGGGGT

GGAGAGCAACTCCGATGGGGCCTCCCCGGAGCCCTGCACCGTCACCCCTGGTGCCGTGAAGCTGGAGAAGGAGAAGC

TGGAGCAAAACCCGGAGGAGTCCCAGGACATCAAAGCTCTGCAGAAAGAACTCGAGCAATTTGCCAAGCTCCTGAAG

CAGAAGAGGATCACCCTGGGATATACACAGGCCGATGTGGGGCTCACCCTGGGGGTTCTATTTGGGAAGGTATTCAG

CCAAACGACCATCTGCCGCTTTGAGGCTCTGCAGCTTAGCTTCAAGAACATGTGTAAGCTGCGGCCCTTGCTGCAGA

AGTGGGTGGAGGAAGCTGACAACAATGAAAATCTTCAGGAGATATGCAAAGCAGAAACCCTCGTGCAGGCCCGAAAG

AGAAAGCGAACCAGTATCGAGAACCGAGTGAGAGGCAACCTGGAGAATTTGTTCCTGCAGTGCCCGAAACCCACACT

GCAGCAGATCAGCCACATCGCCCAGCAGCTTGGGCTCGAGAAGGATGTGGTCCGAGTGTGGTTCTGTAACCGGCGCC

AGAAGGGCAAGCGATCAAGCAGCGACTATGCACAACGAGAGGATTTTGAGGCTGCTGGGTCTCCTTTCTCAGGGGGA

CCAGTGTCCTTTCCTCTGGCCCCAGGGCCCCATTTTGGTACCCCAGGCTATGGGAGCCCTCACTTCACTGCACTGTA

CTCCTCGGTCCCTTTCCCTGAGGGGGAAGCCTTTCCCCCTGTCTCTGTCACCACTCTGGGCTCTCCCATGCATTCAA

AC

SOX2

ATGTACAACATGATGGAGACGGAGCTGAAGCCGCCGGGCCCGCAGCAAACTTCGGGGGGCGGCGGCGGCAACTCCAC

CGCGGCGGCGGCCGGCGGCAACCAGAAAAACAGCCCGGACCGCGTCAAGCGGCCCATGAATGCCTTCATGGTGTGGT

CCCGCGGGCAGCGGCGCAAGATGGCCCAGGAGAACCCCAAGATGCACAACTCGGAGATCAGCAAGCGCCTGGGCGCC

GAGTGGAAACTTTTGTCGGAGACGGAGAAGCGGCCGTTCATCGACGAGGCTAAGCGGCTGCGAGCGCTGCACATGAA

GGAGCACCCGGATTATAAATACCGGCCCCGGCGGAAAACCAAGACGCTCATGAAGAAGGATAAGTACACGCTGCCCG

GCGGGCTGCTGGCCCCCGGCGGCAATAGCATGGCGAGCGGGGTCGGGGTGGGCGCCGGCCTGGGCGCGGGCGTGAAC

CAGCGCATGGACAGTTACGCGCACATGAACGGCTGGAGCAACGGCAGCTACAGCATGATGCAGGACCAGCTGGGCTA

CCCGCAGCACCCGGGCCTCAATGCGCACGGCGCAGCGCAGATGCAGCCCATGCACCGCTACGACGTGAGCGCCCTGC

AGTACAACTCCATGACCAGCTCGCAGACCTACATGAACGGCTCGCCCACCTACAGCATGTCCTACTCGCAGCAGGGC

ACCCCTGGCATGGCTCTTGGCTCCATGGGTTCGGTGGTCAAGTCCGAGGCCAGCTCCAGCCCCCCTGTGGTTACCTC

TTCCTCCCACTCCAGGGCGCCCTGCCAGGCCGGGGACCTCCGGGACATGATCAGCATGTATCTCCCCGGCGCCGAGG

TGCCGGAACCCGCCGCCCCCCAGCAGACTTCACATGTCCCAGCACTACCAGAGCGGCCCGGTGCCCGGCACGGCCATT

AACGGCACACTGCCCCTCTCACACATG

Barcode Sequences

Gene	Barcode Sequence
mCherry	TTCTCATTTTGGAATATAAA
KRAS (G12V)	CTTATCCACTATAAGCGATA
HRAS (G12V)	GTACTTTCTCGCAATTTGTG
MEK1 (S218D, S222D)	TTTAGTATCCAGTCGTAGAA
myr-FLAG-PIK3CA	CAGGACAGGGGGCAGGGAAA
myr-FLAG-AKT1	TCACTTGTCAACAGGTCCAA
FLAG-Rheb (Q64L)	CGTTGTTTTTCGGGAAAACT
IKKα (S176E, S180E)	TGTACCTACAAGCTAAAAAG
FLAG-IKKβ (S177E, S181E)	CTAGTCCATTTAGTATCGGC
STAT3 (A662C, N664C, V667L)	ACATTACGTTCCAGTTACCC
β-catenin (S33A, S37A, T41A, S45A)	ATAAGAACCCTATTTTCAAT

GSK3β (K85A)	TTAGGACTCGAGAGGTTGTT
β-catenin (S33Y)	TACCATTTGTCTTCCGATAG
MAPK9	ATCACTAAGTTCCGTAACCC
FLAG-MKK7-JNK2	ACCCATGCAAATGTAATCTC
MEK5_DD (S311D, T315D)	TTTTGATTATCTTTATAATA
myr-FLAG-MEK5	AAAGGATTACAAATGACCCA
Notch1 ICD	CATTGCTCTTTGTCGATCAT
Notch3 ICD	CAAGTTATTATTTTACTTTT
MAPK14	AGATCAGTCGTTTTGTTCTT
FLAG-MKK6 (S207E, T211E)	GTACAGTAAGGAATATTCCA
GLI2 Truncation	CTAACGTTTGCCACATATTG
SmoM2 (W535L)	TGTCCGTTCGGCTGTATGCC
TGFβR1 (T204D)	CCAATAATGTAAATATCATA
BCL2	AAGTGTACCCCGATAGATTT
BCL-XL	TTCTTCAATGACACTTCGCA
Caspase8 (C360A)	CCTATCACTTCCTCATTATC
Caspase3 (C163A)	TTAGTAGGACTCATGCTCCT
ERα (Y537S)	TACCCCACATGGATTGAATA
AR-V7	AAACCTGGTTGTATTTTTGC
FLAG-YAP2 (8SA)	TTAAAACTAATAAGATTCAC
p53 (R175H)	TGCATTCATGTGAAAGATTT
HRAS (G12V, E37G)	CCGCACCGTTGTGGTGGGGG
Rgl2-CAAX	TAAATGATTCCTGTTTAAGC
RalA (G23V, mature peptide)	TATAAATTTCTATTTACTTA

RalA (G23V, full)	TGTTTCACGTATATGTGATT
SV40 Large T Antigen	TGGCATGGATAACCTATGAT
hTERT	CTTAATTTATACATTCATCT
CCNB1	AAGGTTCTTGCATATGGCAA
CCND1	AGAACCCCTGGGGTACTATG
CDK1	CAACCGTACGCGCTCGGTTA
CDK1 (T14A, Y15F)	TTTATTATGTACTTTTTAGC
CDK4	GACGCCTAGTGTACAGCGCA
CDK4 (R24C)	TGATTGTCAATAGCGTTATT
CDK6	GCCAGCTATAGTACTATTCG
RHOJ	TTGTGTCATTAGCTTTTGAG
PTEN (C124S)	CAACCTTACAGATTTTGTGT
c-MYC	CGACATACATGTAGCTTGCG
KLF4	AAGACTCTCATAATTACCTG
POU5F1	CAGACAGTGCGGCCTGCTCG
SOX2	CCTGTTGACACAAATGTGCT

DNA sequences from Chapter 4

EF1α + TetOn-3G + T2A + Puromycin Resistance

gggcagagcgcacatcgcccacagtccccgagaagttgggggggaggggtcggcaattgaacgggtgcctagagaagg

tggcgcggggtaaactgggaaagtgatgtcgtgtactggctccgccttttttcccgagggtgggggagaaccgtatat

aagtgcagtagtcgccgtgaacgttctttttcgcaacgggtttgccgccagaacacagctgaagcttcgaggggctc

gcatctctccttcacgcgcccgccgccctacctgaggccgccatccacgccggttgagtcgcgttctgccgcctccc

gcctgtggtgcctcctgaactgcgtccgccgtctaggtaagtttaaagctcaggtcgagaccgggcctttgtccggc

gctcccttggagcctacctagactcagccggctctccacgctttgcctgaccctgcttgctcaactctacgtctttg

tttcgttttctgttctgcgccgttacagatctaagctgtgaccggcgcctactctagagccgccatgtctagactgg

acaagagcaaagtcataaactctgctctggaattactcaatggagtcggtatcgaaggcctgacgacaaggaaactc

gctcaaaagctgggagttgagcagcctaccctgtactggcacgtgaagaacaagcgggccctgctcgatgccctgcc

aatcgagatgctggacaggcatcatacccactcctgccccctggaaggcgagtcatggcaagactttctgcggaaca

acgccaagtcataccgctgtgctctcctctcacatcgcgacggggctaaagtgcatctcggcacccgcccaacagag

aaacagtacgaaaccctggaaaatcagctcgcgttcctgtgtcagcaaggcttctccctggagaacgcactgtacgc

tctgtccgccgtgggccactttacactgggctgcgtattggaggaacaggagcatcaagtagcaaaagaggaaagag

agacacctaccaccgattctatgcccccacttctgaaacaagcaattgagctgttcgaccggcagggagccgaacct

gccttccttttcggcctggaactaatcatatgtggcctggagaaacagctaaagtgcgaaagcggcgggccgaccga

cgcccttgacgattttgacttagacatgctcccagccgatgcccttgacgactttgaccttgatatgctgcctgctg

acgctcttgacgattttgaccttgacatgctccccggggagggccgaggaagtcttctaacatgcggtgacgtggag

gagaatcccggccctatgaccgagtacaagcccacggtgcgcctcgccacccgcgacgacgtccccagggccgtacg

caccctcgccgccgcgttcgccgactaccccgccacgcgccacaccgtcgatccggaccgccacatcgagcgggtca

ccgagctgcaagaactcttcctcacgcgcgtcgggctcgacatcggcaaggtgtgggtcgcggacgacggcgccgcg

gtggcggtctggaccacgccggagagcgtcgaagcgggggcggtgttcgccgagatcggcccgcgcatggccgagtt

gagcggttcccggctggccgcgcagcaacagatggaaggcctcctggcgccgcaccggcccaaggagcccgcgtggt

tcctggccaccgtcggcgtctcgcccgaccaccagggcaagggtctgggcagcgccgtcgtgctccccggagtggag

gcggccgagcgcgccggggtgcccgccttcctggagacctccgcgccccgcaacctccccttctacgagcggctcgg

cttcaccgtcaccgccgacgtcgaggtgcccgaaggaccgcgcacctggtgcatgacccgcaagcccggtgcctga

TRE3G Promoter + Kozak Sequence

Gagtttactccctatcagtgatagagaacgtatgaagagtttactccctatcagtgatagagaacgtatgcagactt
tactccctatcagtgatagagaacgtataaggagtttactccctatcagtgatagagaacgtatgaccagtttactc
cctatcagtgatagagaacgtatctacagtttactccctatcagtgatagagaacgtatatccagtttactccctat
cagtgatagagaacgtataagctttaggcgtgtacggtgggcgcctataaaagcagagctcgtttagtgaaccgtca
gatcgcctggagcaattccacaacacttttgtcttataccaactttccgtaccacttcctaccctcgtaaaggtcta
gagctagccacc

GFP

Atggtgagcaagggcgaggagctgttcaccggggtggtgcccatcctggtcgagctggacggcgacgtaaacggcca
caagttcagcgtgtccggcgagggcgagggcgatgccacctacggcaagctgaccctgaagttcatctgcaccaccg
gcaagctgcccgtgccctggcccaccctcgtgaccaccctgacctacggcgtgcagtgcttcagccgctaccccgac
cacatgaagcagcacgacttcttcaagtccgccatgcccgaaggctacgtccaggagcgcaccatcttcttcaagga
cgacggcaactacaagacccgcgccgaggtgaagttcgagggcgacaccctggtgaaccgcatcgagctgaagggca
tcgacttcaaggaggacggcaacatcctggggcacaagctggagtacaactacaacagccacaacgtctatatcatg
gccgacaagcagaagaacggcatcaaggtgaacttcaagatccgccacaacatcgaggacggcagcgtgcagctcgc
cgaccactaccagcagaacaccccatcggcgacggccccgtgctgctgcccgacaaccactacctgagcacccagt
ccgccctgagcaaagaccccaacgagaagcgcgatcacatggtcctgctggagttcgtgaccgccgccgggatcact
ctcggcatggacgagctgtacaagtaa

mCherry

atggtgagcaagggcgaggaggataacatggccatcatcaaggagttcatgcgcttcaaggtgcacatggagggctc
cgtgaacggccacgagttcgagatcgagggcgagggcgagggccgcccctacgagggcacccagaccgccaagctga
aggtgaccaagggtggcccccctgcccttcgcctgggacatcctgtcccctcagttcatgtacggctccaaggcctac
gtgaagcaccccgccgacatccccgactacttgaagctgtccttccccgagggcttcaagtgggagcgcgtgatgaa
cttcgaggacggcggcgtggtgaccgtgacccaggactcctccctgcaggacggcgagttcatctacaaggtgaagc

tgcgcggcaccaacttcccctccgacggccccgtaatgcagaagaagaccatgggctgggaggcctcctccgagcgg

atgtaccccgaggacggcgccctgaagggcgagatcaagcagaggctgaagctgaaggacggcggccactacgacgc

tgaggtcaagaccacctacaaggccaagaagcccgtgcagctgcccggcgcctacaacgtcaacatcaagttggaca

tcacctcccacaacgaggactacaccatcgtggaacagtacgaacgcgccgagggccgccactccaccggcggcatg

gacgagctgtacaagtaa

NEUROD1

ATGACCAAATCGTACAGCGAGAGTGGGCTGATGGGCGAGCCTCAGCCCCAAGGTCCTCCAAGCTGGACAGACGAGTG

TCTCAGTTCTCAGGACGAGGAGCACGAGGCAGACAAGAAGGAGGACGACCTCGAAGCCATGAACGCAGAGGAGGACT

CACTGAGGAACGGGGGAGAGGAGGAGGACGAAGATGAGGACCTGGAAGAGGAGGAAGAAGAGGAAGAGGAGGATGAC

GATCAAAAGCCCAAGAGACGCGGCCCCAAAAAGAAGAAGATGACTAAGGCTCGCCTGGAGCGTTTTAAATTGAGACG

CATGAAGGCTAACGCCCGGGAGCGGAACCGCATGCACGGACTGAACGCGGCGCTAGACAACCTGCGCAAGGTGGTGC

CTTGCTATTCTAAGACGCAGAAGCTGTCCAAAATCGAGACTCTGCGCTTGGCCAAGAACTACATCTGGGCTCTGTCG

GAGATCCTGCGCTCAGGCAAAAGCCCAGACCTGGTCTCCTTCGTTCAGACGCTTTGCAAGGGCTTATCCCAACCCAC

CACCAACCTGGTTGCGGGCTGCCTGCAACTCAATCCTCGGACTTTTCTGCCTGAGCAGAACCAGGACATGCCCCCCC

ACCTGCCGACGGCCAGCGCTTCCTTCCCTGTACACCCCTACTCCTACCAGTCGCCTGGGCTGCCCAGTCCGCCTTAC

GGTACCATGGACAGCTCCCATGTCTTCCACGTTAAGCCTCCGCCGCACGCCTACAGCGCAGCGCTGGAGCCCTTCTT

TGAAAGCCCTCTGACTGATTGCACCAGCCCTTCCTTTGATGGACCCCTCAGCCCGCCGCTCAGCATCAATGGCAACT

TCTCTTTCAAACACGAACCGTCCGCCGAGTTTGAGAAAAATTATGCCTTTACCATGCACTATCCTGCAGCGACACTG

GCAGGGGCCCAAAGCCACGGATCAATCTTCTCAGGCACCGCTGCCCCTCGCTGCGAGATCCCCATAGACAATATTAT

GTCCTTCGATAGCCATTCACATCATGAGCGAGTCATGAGTGCCCAGCTCAATGCCATATTTCATGAT

MYOD1

ATGGAGCTACTGTCGCCACCGCTCCGCGACGTAGACCTGACGGCCCCCGACGGCTCTCTCTGCTCCTTTGCCACAAC

GGACGACTTCTATGACGACCCGTGTTTCGACTCCCCGGACCTGCGCTTCTTCGAAGACCTGGACCCGCGCCTGATGC

ACGTGGGCGCGCTCCTGAAACCCGAAGAGCACTCGCACTTCCCCGCGGCGGTGCACCCGGCCCCGGGCGCACGTGAG

GACGAGCATGTGCGCGCGCCCAGCGGGCACCACCAGGCGGGCCGCTGCCTACTGTGGGCCTGCAAGGCGTGCAAGCG

CAAGACCACCAACGCCGACCGCCGCAAGGCCGCCACCATGCGCGAGCGGCGCCGCCTGAGCAAAGTAAATGAGGCCT

TTGAGACACTCAAGCGCTGCACGTCGAGCAATCCAAACCAGCGGTTGCCCAAGGTGGAGATCCTGCGCAACGCCATC

CGCTATATCGAGGGCCTGCAGGCTCTGCTGCGCGACCAGGACGCCGCGCCCCCTGGCGCCGCAGCCGCCTTCTATGC

GCCGGGCCCGCTGCCCCCGGGCCGCGGCGGCGAGCACTACAGCGGCGACTCCGACGCGTCCAGCCCGCGCTCCAACT

GCTCCGACGGCATGATGGACTACAGCGGCCCCCCGAGCGGCGCCCGGCGGCGGAACTGCTACGAAGGCGCCTACTAC

AACGAGGCGCCCAGCGAACCCAGGCCCGGGAAGAGTGCGGCGGTGTCGAGCCTAGACTGCCTGTCCAGCATCGTGGA

GCGCATCTCCACCGAGAGCCCTGCGGCGCCCGCCCTCCTGCTGGCGGACGTGCCTTCTGAGTCGCCTCCGCGCAGGC

AAGAGGCTGCCGCCCCCAGCGAGGGAGAGAGCAGCGGCGACCCCACCCAGTCACCGGACGCCGCCCCGCAGTGCCCT

GCGGGTGCGAACCCCAACCCGATATACCAGGTGCTC

BAF60C

ATGGCCGCGGACGAAGTTGCCGGAGGGGCGCGCAAAGCCACGAAAAGCAAACTTTTTGAGTTTCTGGTCCATGGGGT

GCGCCCCGGGATGCCGTCTGGAGCCCGGATGCCCCACCAGGGGGCGCCCATGGGCCCCCCGGGCTCCCCGTACATGG

GCAGCCCCGCCGTGCGACCCGGCCTGGCCCCCGCGGGCATGGAGCCCGCCCGCAAGCGAGCAGCGCCCCCGCCCGGG

CAGAGCCAGGCACAGAGCCAGGGCCAGCCGGTGCCCACCGCCCCCGCGCGGAGCCGCAGTGCCAAGAGGAGGAAGAT

GGCTGACAAAATCCTCCCTCAAAGGATTCGGGAGCTGGTCCCCGAGTCCCAGGCTTACATGGACCTCTTGGCATTTG

AGAGGAAACTGGATCAAACCATCATGCGGAAGCGGGTGGACATCCAGGAGGCTCTGAAGAGGCCCATGAAGCAAAAG

CGGAAGCTGCGACTCTATATCTCCAACACTTTTAACCCTGCGAAGCCTGATGCTGAGGATTCCGACGGCAGCATTGC

CTCCTGGGAGCTACGGGTGGAGGGGAAGCTCCTGGATGATCCCAGCAAACAGAAGCGGAAGTTCTCTTCTTTCTTCA

AGAGTTTGGTCATCGAGCTGGACAAAGATCTTTATGGCCCTGACAACCACCTCGTTGAGTGGCATCGGACACCCACG

ACCCAGGAGACGGACGGCTTCCAGGTGAAACGGCCTGGGGACCTGAGTGTGCGCTGCACGCTGCTCCTCATGCTGGA

CTACCAGCCTCCCCAGTTCAAACTGGATCCCCGCCTAGCCCGGCTGCTGGGGCTGCACACACAGAGCCGCTCAGCCA

TTGTCCAGGCCCTGTGGCAGTATGTGAAGACCAACAGGCTGCAGGACTCCCATGACAAGGAATACATCAATGGGGAC

AAGTATTTCCAGCAGATTTTTGATTGTCCCCGGCTGAAGTTTTCTGAGATTCCCCAGCGCCTCACAGCCCTGCTATT

GCCCCCTGACCCAATTGTCATCAACCATGTCATCAGCGTGGACCCTTCAGACCAGAAGAAGACGGCGTGCTATGACA

TTGACGTGGAGGTGGAGGAGCCATTAAAGGGGCAGATGAGCAGCTTCCTCCTATCCACGGCCAACCAGCAGGAGATC

AGTGCTCTGGACAGTAAGATCCATGAGACGATTGAGTCCATAAACCAGCTCAAGATCCAGAGGGACTTCATGCTAAG

CTTCTCCAGAGACCCCAAAGGCTATGTCCAAGACCTGCTCCGCTCCCAGAGCCGGGACCTCAAGGTGATGACAGATG

TAGCCGGCAACCCTGAAGAGGAGCGCCGGGCTGAGTTCTACCACCAGCCCTGGTCCCAGGAGGCCGTCAGTCGCTAC

TTCTACTGCAAGATCCAGCAGCGCAGGCAGGAGCTGGAGCAGTCGCTGGTTGTGCGCAACACC

ASCL1

ATGGAGTCTTCTGCTAAAATGGAGTCCGGAGGCGCGGGACAACAACCACAACCGCAACCACAACAACCCTTCCTGCC

GCCGGCCGCATGTTTTTTCGCGACCGCTGCTGCTGCTGCAGCGGCGGCGGCTGCTGCCGCCGCGCAATCCGCCCAAC

AGCAACAACAACAACAGCAGCAGCAGCAACAAGCGCCTCAACTTCGACCCGCTGCAGACGGGCAGCCCTCAGGGGGA

GGGCACAAGAGCGCTCCGAAGCAGGTTAAAAGGCAGAGGAGCAGTAGTCCCGAACTGATGCGATGTAAGAGGCGCCT

CAATTTTAGCGGTTTTGGTTACTCTTTGCCCCAGCAGCAGCCGGCTGCCGTAGCTCGCCGAAATGAGCGGGAAAGGA

ACCGCGTTAAACTTGTGAATCTCGGTTTCGCGACACTTCGAGAGCACGTACCAAATGGGGCAGCTAACAAGAAAATG

AGTAAAGTTGAGACACTGCGGTCTGCAGTGGAGTATATTAGAGCTCTTCAACAATTGCTTGACGAGCACGATGCCGT

ATCAGCCGCATTTCAAGCCGGGGTGCTGTCCCCAACAATATCTCCGAACTACAGCAATGATCTTAATAGCATGGCGG

GAAGTCCCGTTTCCTCCTACTCCTCTGATGAGGGCAGCTACGACCCTCTCAGTCCCGAGGAGCAAGAGCTTCTTGAC

TTCACTAACTGGTTC

DLX2

atgactggagtctttgacagtctagtggctgatatgcactcgacccagatcgccgcctccagcacgtaccaccagca

ccagcagcccccgagcggcggcggcgccggcccgggtggcaacagcagcagcagcagcctccacaagccccagg

agtcgcccacccttccggtgtccaccgccaccgacagcagctactacaccaaccagcagcacccggcgggcggcggc

ggcggcgggggctcgccctacgcgcacatgggttcctaccagtaccaagccagcggcctcaacaacgtcccttactc

cgccaagagcagctatgacctgggctacaccgccgcctacacctcctacgctccctatggaaccagttcgtccccag

ccaacaacgagcctgagaaggaggaccttgagcctgaaattcggatagtgaacgggaagccaaagaaagtccggaaa

ccccgcaccatctactccagtttccagctggcggctcttcagcggcgtttccaaaagactcaatacttggccttgcc

ggagcgagccgagctggcggcctctctgggcctcacccagactcaggtcaaaatctggttccagaaccgccggtcca

agttcaagaagatgtggaaaagtggtgagatcccctcggagcagcaccctggggccagcgcttctccaccttgtgct

tcgccgccagtctcagcgccggcctcctgggactttggtgtgccgcagcggatggcgggcggcggtggtccgggcag

tggcggcagcggcgccggcagctcgggctccagcccgagcagcgcggcctcggcttttctgggcaactacccctggt

accaccagacctcgggatccgcctcacacctgcaggccacggcgccgctgctgcaccccactcagaccccgcagccg

catcaccaccaccaccatcacggcggcgggggcgccccggtgagcgcggggacgattttc

ETV2

ATGGATCTTTGGAACTGGGATGAAGCTTCCCCTCAAGAAGTTCCCCCCGGAAATAAACTCGCGGGGCTTGGAAGACT

CCCTCGCCTTCCGCAACGCGTCTGGGGCGGATGCCCTGGTGGAGCCTCAGCGGACCCAAACCCTTTGTCTCCAGCGG

AGGGGGCAAAGTTGGGTTTCTGCTTCCCGGATCTTGCTTTGCAAGGCGATACTCCAACGGCGACGGCAGAGACCTGT

TGGAAAGGCACCAGTAGCTCCCTGGCCAGCTTTCCGCAGCTCGATTGGGGGTCAGCCCTTCTCCATCCCGAAGTTCC

CTGGGGGGCGGAACCCGACTCCCAAGCCCTTCCCTGGAGTGGTGATTGGACAGATATGGCATGCACAGCCTGGGACA

GTTGGTCCGGGGCGTCACAGACATTGGGACCAGCCCCACTTGGACCGGGGCCTATCCCCGCAGCAGGAAGCGAAGGA

GCTGCTGGTCAGAACTGTGTGCCCGTGGCTGGTGAGGCTACCAGTTGGTCCAGGGCCCAGGCAGCAGGCAGTAACAC

CAGCTGGGATTGCTCAGTGGGGCCTGACGGGGATACTTATTGGGGCTCTGGTCTTGGTGGAGAACCGAGAACGGACT

GTACGATAAGTTGGGGCGGTCCAGCTGGGCCTGATTGTACTACGTCATGGAATCCTGGCTTGCACGCCGGCGGCACG

ACAAGCCTTAAGAGATATCAAAGTTCAGCCCTTACAGTTTGCTCAGAACCTTCCCCGCAAAGTGACCGAGCGTCACT

GGCGCGATGTCCTAAAACTAATCATCGAGGGCCGATCCAGTTGTGGCAGTTTTTGCTTGAACTCCTTCACGATGGCG

CGAGGAGCAGTTGCATCAGATGGACCGGTAACAGCAGGGAGTTCCAATTGTGTGACCCCAAGGAAGTGGCTCGACTG

TGGGGTGAGCGCAAACGGAAGCCTGGTATGAATTACGAAAAGTTGAGTAGGGGTTTGCGATATTACTATAGGCGCGA

CATCGTTCGAAAGTCCGGTGGTCGAAAGTACACATACAGATTCGGCGGTCGCGTACCATCTCTTGCATACCCTGATT

GCGCAGGCGGGGGTAGGGGTGCGGAAACACAA

Bibliography

(1) Gould, S. E.; Junttila, M. R.; De Sauvage, F. J. Translational Value of Mouse Models in Oncology Drug Development. *Nature Medicine*. Nature Publishing Group May 7, 2015, pp 431–439.

(2) Cheon, D.-J.; Orsulic, S. Mouse Models of Cancer. *Annu. Rev. Pathol. Mech. Dis.* **2011**, *6* (1), 95–119.

(3) Jucker, M. The Benefits and Limitations of Animal Models for Translational Research in Neurodegenerative Diseases. *Nature Medicine*. Nature Publishing Group September 21, 2010, pp 1210–1214.

(4) Greek, R.; Hansen, L. A. The Strengths and Limits of Animal Models as Illustrated by the Discovery and Development of Antibacterials. *Biol Syst* **2013**, *2*, 109.

(5) Fischer, M. Mice Are Not Humans: The Case of P53. *Trends in Cancer*. Cell Press January 1, 2021, pp 12–14.

(6) Rangarajan, A.; Weinberg, R. A. Comparative Biology of Mouse versus Human Cells: Modelling Human Cancer in Mice. *Nature Reviews Cancer*. European Association for Cardio-Thoracic Surgery 2003, pp 952–959.

(7) Wilding, J. L.; Bodmer, W. F. Cancer Cell Lines for Drug Discovery and Development. *Cancer Research*. American Association for Cancer Research Inc. January 5, 2014, pp 2377–2384.

(8) Gillet, J. P.; Varma, S.; Gottesman, M. M. The Clinical Relevance of Cancer Cell Lines. *Journal of the National Cancer Institute*. Oxford University Press April 3, 2013, pp 452–458.

(9) Chappell, L.; Russell, A. J. C.; Voet, T. Single-Cell (Multi)Omics Technologies. *Annual Review of Genomics and Human Genetics*. Annual Reviews Inc. August 31, 2018, pp 15–41.

(10) Nam, A. S.; Chaligne, R.; Landau, D. A. Integrating Genetic and Non-Genetic Determinants of Cancer Evolution by Single-Cell Multi-Omics. *Nature Reviews Genetics*. Nature Research January 1, 2021, pp 3–18.

(11) Wilbrey-Clark, A.; Roberts, K.; Teichmann, S. A. Cell Atlas Technologies and Insights into Tissue Architecture. *Biochemical Journal*. Portland Press Ltd April 1, 2020, pp 1427–1442.

(12) Regev, A.; Teichmann, S. A.; Lander, E. S.; Amit, I.; Benoist, C.; Birney, E.; Bodenmiller, B.; Campbell, P.; Carninci, P.; Clatworthy, M.; et al. The Human Cell Atlas. *Elife* **2017**, *6*.

(13) Lake, B. B.; Ai, R.; Kaeser, G. E.; Salathia, N. S.; Yung, Y. C.; Liu, R.; Wildberg, A.; Gao, D.; Fung, H. L.; Chen, S.; et al. Neuronal Subtypes and Diversity Revealed by Single-Nucleus RNA Sequencing of the Human Brain. *Science (80-.).* **2016**, *352* (6293), 1586–1590.

(14) Villani, A. C.; Satija, R.; Reynolds, G.; Sarkizova, S.; Shekhar, K.; Fletcher, J.; Griesbeck, M.; Butler, A.; Zheng, S.; Lazo, S.; et al. Single-Cell RNA-Seq Reveals New Types of Human Blood Dendritic Cells, Monocytes, and Progenitors. *Science (80-.).* **2017**, *356* (6335).

(15) Darmanis, S.; Sloan, S. A.; Zhang, Y.; Enge, M.; Caneda, C.; Shuer, L. M.; Gephart, M. G. H.; Barres, B. A.; Quake, S. R. A Survey of Human Brain Transcriptome Diversity at the Single Cell Level. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (23), 7285–7290.

(16) Rozenblatt-Rosen, O.; Stubbington, M. J. T.; Regev, A.; Teichmann, S. A. The Human Cell Atlas: From Vision to Reality. *Nature*. Nature Publishing Group October 18, 2017, pp 451–453.

(17) Atala, A. Regenerative Medicine Strategies. *Journal of Pediatric Surgery*. W.B. Saunders January 1, 2012, pp 17–28.

(18) Shafiee, A.; Atala, A. Tissue Engineering: Toward a New Era of Medicine. *Annual Review of Medicine*. Annual Reviews Inc. January 14, 2017, pp 29–40.

(19) *Principles of Tissue Engineering - 5th Edition*, 5th Edition.; Lanza, R., Langer, R., Vacanti, J., Atala, A., Eds.; Academic Press, 2020.

(20) Persidis, A. Tissue Engineering. *Nature Biotechnology*. Nature Publishing Group 1999, pp 508–510.

(21) Aguilar, C. A.; Greising, S. M.; Watts, A.; Goldman, S. M.; Peragallo, C.; Zook, C.; Larouche, J.; Corona, B. T. Multiscale Analysis of a Regenerative Therapy for Treatment of Volumetric Muscle Loss Injury. *Cell Death Discov.* **2018**, *4* (1), 1–11.

(22) Piao, J.; Zabierowski, S.; Dubose, B. N.; Hill, E. J.; Navare, M.; Claros, N.; Rosen, S.; Ramnarine, K.; Horn, C.; Fredrickson, C.; et al. Preclinical Efficacy and Safety of a Human Embryonic Stem Cell-Derived Midbrain Dopamine Progenitor Product, MSK-DA01. *Cell Stem Cell* **2021**, *28* (2), 217-229.e7.

(23) Doi, D.; Magotani, H.; Kikuchi, T.; Ikeda, M.; Hiramatsu, S.; Yoshida, K.; Amano, N.; Nomura, M.; Umekage, M.; Morizane, A.; et al. Pre-Clinical Study of Induced Pluripotent Stem Cell-Derived Dopaminergic Progenitor Cells for Parkinson's Disease. *Nat. Commun.* **2020**, *11* (1), 1–14.

(24) Kim, T. W.; Piao, J.; Koo, S. Y.; Kriks, S.; Chung, S. Y.; Betel, D.; Socci, N. D.; Choi, S. J.; Zabierowski, S.; Dubose, B. N.; et al. Biphasic Activation of WNT Signaling Facilitates the Derivation of Midbrain Dopamine Neurons from HESCs for Translational Use. *Cell Stem Cell* **2021**, *28* (2), 343-355.e5.

(25) Mandai, M.; Watanabe, A.; Kurimoto, Y.; Hirami, Y.; Morinaga, C.; Daimon, T.; Fujihara, M.; Akimaru, H.; Sakai, N.; Shibata, Y.; et al. Autologous Induced Stem-Cell–Derived Retinal Cells for Macular Degeneration. *N. Engl. J. Med.* **2017**, *376* (11), 1038–1046.

(26) Sharma, R.; Khristov, V.; Rising, A.; Jha, B. S.; Dejene, R.; Hotaling, N.; Li, Y.; Stoddard, J.; Stankewicz, C.; Wan, Q.; et al. Clinical-Grade Stem Cell-Derived Retinal Pigment Epithelium Patch Rescues Retinal Degeneration in Rodents and Pigs. *Sci. Transl. Med.*

2019, *11* (475).

(27) Yoshihara, E.; O'Connor, C.; Gasser, E.; Wei, Z.; Oh, T. G.; Tseng, T. W.; Wang, D.; Cayabyab, F.; Dai, Y.; Yu, R. T.; et al. Immune-Evasive Human Islet-like Organoids Ameliorate Diabetes. *Nature* **2020**, *586* (7830), 606–611.

(28) Feins, S.; Kong, W.; Williams, E. F.; Milone, M. C.; Fraietta, J. A. An Introduction to Chimeric Antigen Receptor (CAR) T-Cell Immunotherapy for Human Cancer. *American Journal of Hematology*. Wiley-Liss Inc. May 1, 2019, pp S3–S9.

(29) Maldini, C. R.; Ellis, G. I.; Riley, J. L. CAR T Cells for Infection, Autoimmunity and Allotransplantation. *Nature Reviews Immunology*. Nature Publishing Group October 1, 2018, pp 605–616.

(30) Seif, M.; Einsele, H.; Löffler, J. CAR T Cells Beyond Cancer: Hope for Immunomodulatory Therapy of Infectious Diseases. *Frontiers in Immunology*. Frontiers Media S.A. November 21, 2019, p 2711.

(31) Sottoriva, A.; Kang, H.; Ma, Z.; Graham, T. A.; Salomon, M. P.; Zhao, J.; Marjoram, P.; Siegmund, K.; Press, M. F.; Shibata, D.; et al. A Big Bang Model of Human Colorectal Tumor Growth. *Nat. Genet.* **2015**, *47* (3), 209–216.

(32) Takahashi, K.; Yamanaka, S. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell* **2006**, *126* (4), 663–676.

(33) Takahashi, K.; Tanabe, K.; Ohnuki, M.; Narita, M.; Ichisaka, T.; Tomoda, K.; Yamanaka, S. Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell* **2007**, *131* (5), 861–872.

(34) Shi, Y.; Inoue, H.; Wu, J. C.; Yamanaka, S. Induced Pluripotent Stem Cell Technology: A Decade of Progress. *Nature Reviews Drug Discovery*. Nature Publishing Group February 2, 2017, pp 115–130.

(35) Hockemeyer, D.; Jaenisch, R. Induced Pluripotent Stem Cells Meet Genome Editing. *Cell Stem Cell* **2016**, *18* (5), 573–586.

(36) Li, Y.; Muffat, J.; Javed, A. O.; Keys, H. R.; Lungjangwa, T.; Bosch, I.; Khan, M.; Virgilio, M. C.; Gehrke, L.; Sabatini, D. M.; et al. Genome-Wide CRISPR Screen for Zika Virus Resistance in Human Neural Cells. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116* (19), 9527–9532.

(37) Tian, R.; Gachechiladze, M. A.; Ludwig, C. H.; Laurie, M. T.; Hong, J. Y.; Nathaniel, D.; Prabhu, A. V.; Fernandopulle, M. S.; Patel, R.; Abshari, M.; et al. CRISPR Interference-Based Platform for Multimodal Genetic Screens in Human IPSC-Derived Neurons. *Neuron* **2019**, *104* (2), 239-255.e12.

(38) Ihry, R. J.; Salick, M. R.; Ho, D. J.; Sondey, M.; Kommineni, S.; Paula, S.; Raymond, J.; Henry, B.; Frias, E.; Wang, Q.; et al. Genome-Scale CRISPR Screens Identify Human Pluripotency-Specific Genes. *Cell Rep.* **2019**, *27* (2), 616-630.e6.

(39) Hwang, N. S.; Varghese, S.; Elisseeff, J. Controlled Differentiation of Stem Cells.

Advanced Drug Delivery Reviews. NIH Public Access January 14, 2008, pp 199–214.

(40) Musah, S.; Wrighton, P. J.; Zaltsman, Y.; Zhong, X.; Zorn, S.; Parlato, M. B.; Hsiao, C.; Palecek, S. P.; Chang, Q.; Murphy, W. L.; et al. Substratum-Induced Differentiation of Human Pluripotent Stem Cells Reveals the Coactivator YAP Is a Potent Regulator of Neuronal Specification. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (38), 13805–13810.

(41) Davis, R. L.; Weintraub, H.; Lassar, A. B. Expression of a Single Transfected CDNA Converts Fibroblasts to Myoblasts. *Cell* **1987**, *51* (6), 987–1000.

(42) Xu, J.; Du, Y.; Deng, H. Direct Lineage Reprogramming: Strategies, Mechanisms, and Applications. *Cell Stem Cell* **2015**, *16* (2), 119–134.

(43) Pang, Z. P.; Yang, N.; Vierbuchen, T.; Ostermeier, A.; Fuentes, D. R.; Yang, T. Q.; Citri, A.; Sebastiano, V.; Marro, S.; Südhof, T. C.; et al. Induction of Human Neuronal Cells by Defined Transcription Factors. *Nature* **2011**, *476* (7359), 220–223.

(44) Zhang, Y.; Pak, C.; Han, Y.; Ahlenius, H.; Zhang, Z.; Chanda, S.; Marro, S.; Patzke, C.; Acuna, C.; Covy, J.; et al. Rapid Single-Step Induction of Functional Neurons from Human Pluripotent Stem Cells. *Neuron* **2013**, *78* (5), 785–798.

(45) Yang, N.; Chanda, S.; Marro, S.; Ng, Y.-H.; Janas, J. A.; Haag, D.; Ang, C. E.; Tang, Y.; Flores, Q.; Mall, M.; et al. Generation of Pure GABAergic Neurons by Transcription Factor Programming. *Nat. Methods* **2017**, *14* (6), 621–628.

(46) Chanda, S.; Ang, C. E.; Davila, J.; Pak, C.; Mall, M.; Lee, Q. Y.; Ahlenius, H.; Jung, S. W.; Südhof, T. C.; Wernig, M. Generation of Induced Neuronal Cells by the Single Reprogramming Factor ASCL1. *Stem Cell Reports* **2014**, *3* (2), 282–296.

(47) Abujarour, R.; Bennett, M.; Valamehr, B.; Lee, T. T.; Robinson, M.; Robbins, D.; Le, T.; Lai, K.; Flynn, P. Myogenic Differentiation of Muscular Dystrophy-Specific Induced Pluripotent Stem Cells for Use in Drug Discovery. *Stem Cells Transl. Med.* **2014**, *3* (2), 149–160.

(48) Lancaster, M. a.; Knoblich, J. a. Organogenesis in a Dish: Modeling Development and Disease Using Organoid Technologies. *Science* **2014**, *345* (6194), 1247125.

(49) Schutgens, F.; Clevers, H. Human Organoids: Tools for Understanding Biology and Treating Diseases. *Annual Review of Pathology: Mechanisms of Disease*. Annual Reviews Inc. January 24, 2020, pp 211–234.

(50) Sato, T.; Vries, R. G.; Snippert, H. J.; Van De Wetering, M.; Barker, N.; Stange, D. E.; Van Es, J. H.; Abo, A.; Kujala, P.; Peters, P. J.; et al. Single Lgr5 Stem Cells Build Crypt-Villus Structures in Vitro without a Mesenchymal Niche. *Nature* **2009**, *459* (7244), 262–265.

(51) Lancaster, M. A.; Renner, M.; Martin, C. A.; Wenzel, D.; Bicknell, L. S.; Hurles, M. E.; Homfray, T.; Penninger, J. M.; Jackson, A. P.; Knoblich, J. A. Cerebral Organoids Model Human Brain Development and Microcephaly. *Nature* **2013**, *501* (7467), 373–379.

(52) Takasato, M.; Er, P. X.; Chiu, H. S.; Maier, B.; Baillie, G. J.; Ferguson, C.; Parton, R. G.;

Wolvetang, E. J.; Roost, M. S.; De Sousa Lopes, S. M. C.; et al. Kidney Organoids from Human IPS Cells Contain Multiple Lineages and Model Human Nephrogenesis. *Nature* **2015**, *526* (7574), 564–568.

(53) Clevers, H. Modeling Development and Disease with Organoids. *Cell* **2016**, *165* (7), 1586–1597.

(54) Drost, J.; Van Jaarsveld, R. H.; Ponsioen, B.; Zimberlin, C.; Van Boxtel, R.; Buijs, A.; Sachs, N.; Overmeer, R. M.; Offerhaus, G. J.; Begthel, H.; et al. Sequential Cancer Mutations in Cultured Human Intestinal Stem Cells. *Nature* **2015**, *521* (7550), 43–47.

(55) Matano, M.; Date, S.; Shimokawa, M.; Takano, A.; Fujii, M.; Ohta, Y.; Watanabe, T.; Kanai, T.; Sato, T. Modeling Colorectal Cancer Using CRISPR-Cas9-Mediated Engineering of Human Intestinal Organoids. *Nat. Med.* **2015**, *21* (3), 256–262.

(56) Neal, J. T.; Li, X.; Zhu, J.; Giangarra, V.; Grzeskowiak, C. L.; Ju, J.; Liu, I. H.; Chiou, S. H.; Salahudeen, A. A.; Smith, A. R.; et al. Organoid Modeling of the Tumor Immune Microenvironment. *Cell* **2018**, *175* (7), 1972-1988.e16.

(57) Di Lullo, E.; Kriegstein, A. R. The Use of Brain Organoids to Investigate Neural Development and Disease. *Nat. Rev. Neurosci.* **2017**, *18* (10), 573–584.

(58) Smith, R. C.; Tabar, V. Constructing and Deconstructing Cancers Using Human Pluripotent Stem Cells and Organoids. *Cell Stem Cell*. Cell Press January 3, 2019, pp 12–24.

(59) Bian, S.; Repic, M.; Guo, Z.; Kavirayani, A.; Burkard, T.; Bagley, J. A.; Krauditsch, C.; Knoblich, J. A. Genetically Engineered Cerebral Organoids Model Brain Tumor Formation. *Nat. Methods* **2018**, *15* (8), 631–639.

(60) Daniel, E.; Cleaver, O. Vascularizing Organogenesis: Lessons from Developmental Biology and Implications for Regenerative Medicine. In *Current Topics in Developmental Biology*; Academic Press Inc., 2019; Vol. 132, pp 177–220.

(61) Petrova, T. V.; Koh, G. Y. Organ-Specific Lymphatic Vasculature: From Development to Pathophysiology. *Journal of Experimental Medicine*. Rockefeller University Press January 1, 2018, pp 35–49.

(62) Grebenyuk, S.; Ranga, A. Engineering Organoid Vascularization. *Frontiers in Bioengineering and Biotechnology*. Frontiers Media S.A. March 19, 2019, p 39.

(63) Lancaster, M. A. Brain Organoids Get Vascularized. *Nature Biotechnology*. Nature Publishing Group May 9, 2018, pp 407–408.

(64) Kolesky, D. B.; Homan, K. A.; Skylar-Scott, M. A.; Lewis, J. A. Three-Dimensional Bioprinting of Thick Vascularized Tissues. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (12), 3179–3184.

(65) Skylar-Scott, M. A.; Uzel, S. G. M.; Nam, L. L.; Ahrens, J. H.; Truby, R. L.; Damaraju, S.; Lewis, J. A. Biomanufacturing of Organ-Specific Tissues with High Cellular Density and Embedded Vascular Channels. *Sci. Adv.* **2019**, *5* (9), eaaw2459.

(66) Mansour, A. A.; Gonçalves, J. T.; Bloyd, C. W.; Li, H.; Fernandes, S.; Quang, D.; Johnston, S.; Parylak, S. L.; Jin, X.; Gage, F. H. An in Vivo Model of Functional and Vascularized Human Brain Organoids. *Nat. Biotechnol.* **2018**, *36* (5), 432–441.

(67) Pham, M. T.; Pollock, K. M.; Rose, M. D.; Cary, W. A.; Stewart, H. R.; Zhou, P.; Nolta, J. A.; Waldau, B. Generation of Human Vascularized Brain Organoids. *Neuroreport* **2018**, *29* (7), 588–593.

(68) Takebe, T.; Sekine, K.; Enomura, M.; Koike, H.; Kimura, M.; Ogaeri, T.; Zhang, R.-R.; Ueno, Y.; Zheng, Y.-W.; Koike, N.; et al. Vascularized and Functional Human Liver from an IPSC-Derived Organ Bud Transplant. *Nature* **2013**, *499* (7459), 481–484.

(69) Takebe, T.; Enomura, M.; Yoshizawa, E.; Kimura, M.; Koike, H.; Ueno, Y.; Matsuzaki, T.; Yamazaki, T.; Toyohara, T.; Osafune, K.; et al. Vascularized and Complex Organ Buds from Diverse Tissues via Mesenchymal Cell-Driven Condensation. *Cell Stem Cell* **2015**, *16* (5), 556–565.

(70) Cakir, B.; Xiang, Y.; Tanaka, Y.; Kural, M. H.; Parent, M.; Kang, Y. J.; Chapeton, K.; Patterson, B.; Yuan, Y.; He, C. S.; et al. Engineering of Human Brain Organoids with a Functional Vascular-like System. *Nat. Methods* **2019**, *16* (11), 1169–1175.

(71) Wu, J.; Platero-Luengo, A.; Sakurai, M.; Sugawara, A.; Gil, M. A.; Yamauchi, T.; Suzuki, K.; Bogliotti, Y. S.; Cuello, C.; Morales Valencia, M.; et al. Interspecies Chimerism with Mammalian Pluripotent Stem Cells. *Cell* **2017**, *168* (3), 473-486.e15.

(72) Kobayashi, T.; Yamaguchi, T.; Hamanaka, S.; Kato-Itoh, M.; Yamazaki, Y.; Ibata, M.; Sato, H.; Lee, Y. S.; Usui, J. ichi; Knisely, A. S.; et al. Generation of Rat Pancreas in Mouse by Interspecific Blastocyst Injection of Pluripotent Stem Cells. *Cell* **2010**, *142* (5), 787–799.

(73) Matsunari, H.; Nagashima, H.; Watanabe, M.; Umeyama, K.; Nakano, K.; Nagaya, M.; Kobayashi, T.; Yamaguchi, T.; Sumazaki, R.; Herzenberg, L. A.; et al. Blastocyst Complementation Generates Exogenic Pancreas in Vivo in Apancreatic Cloned Pigs. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (12), 4557–4562.

(74) McDonald, D.; Wu, Y.; Dailamy, A.; Tat, J.; Parekh, U.; Zhao, D.; Hu, M.; Tipps, A.; Zhang, K.; Mali, P. Defining the Teratoma as a Model for Multi-Lineage Human Development. *Cell* **2020**, *183* (5), 1402-1419.e18.

(75) Chan, S. S. K.; Arpke, R. W.; Filareto, A.; Xie, N.; Pappas, M. P.; Penaloza, J. S.; Perlingeiro, R. C. R.; Kyba, M. Skeletal Muscle Stem Cells from PSC-Derived Teratomas Have Functional Regenerative Capacity. *Cell Stem Cell* **2018**, *23* (1), 74-85.e6.

(76) Lensch, M. W.; Schlaeger, T. M.; Zon, L. I.; Daley, G. Q. Teratoma Formation Assays with Human Embryonic Stem Cells: A Rationale for One Type of Human-Animal Chimera. *Cell Stem Cell*. Elsevier Inc. September 13, 2007, pp 253–258.

(77) Feiner, R.; Dvir, T. Engineering Smart Hybrid Tissues with Built-In Electronics. *iScience*. Elsevier Inc. February 21, 2020, p 100833.

(78) Kim, D. H.; Ghaffari, R.; Lu, N.; Rogers, J. A. Flexible and Stretchable Electronics for

Biointegrated Devices. *Annual Review of Biomedical Engineering*. Annual Reviews August 18, 2012, pp 113–128.

(79) Hong, G.; Lieber, C. M. Novel Electrode Technologies for Neural Recordings. *Nature Reviews Neuroscience*. Nature Publishing Group June 1, 2019, pp 330–345.

(80) Gu, Y.; Zhang, T.; Chen, H.; Wang, F.; Pu, Y.; Gao, C.; Li, S. Mini Review on Flexible and Wearable Electronics for Monitoring Human Health Information. *Nanoscale Research Letters*. Springer New York LLC December 1, 2019, pp 1–15.

(81) Liu, Y.; Pharr, M.; Salvatore, G. A. Lab-on-Skin: A Review of Flexible and Stretchable Electronics for Wearable Health Monitoring. *ACS Nano*. American Chemical Society October 24, 2017, pp 9614–9635.

(82) Li, Q.; Zhang, J.; Li, Q.; Li, G.; Tian, X.; Luo, Z.; Qiao, F.; Wu, X.; Zhang, J. Review of Printed Electrodes for Flexible Devices. *Frontiers in Materials*. Frontiers Media S.A. January 30, 2019, p 77.

(83) Chen, Y.; Zhang, Y.; Liang, Z.; Cao, Y.; Han, Z.; Feng, X. Flexible Inorganic Bioelectronics. *npj Flexible Electronics*. Nature Research December 1, 2020, pp 1–20.

(84) Inzelberg, L.; Hanein, Y. Electrophysiology Meets Printed Electronics: The Beginning of a Beautiful Friendship. *Front. Neurosci.* **2019**, *12*, 992.

(85) Choi, H. K.; Lee, J. H.; Lee, T.; Lee, S. N.; Choi, J. W. Flexible Electronics for Monitoring in Vivo Electrophysiology and Metabolite Signals. *Frontiers in Chemistry*. Frontiers Media S.A. November 19, 2020, p 547591.

(86) Xu, B.; Akhtar, A.; Liu, Y.; Chen, H.; Yeo, W. H.; Park, S.; Boyce, B.; Kim, H.; Yu, J.; Lai, H. Y.; et al. An Epidermal Stimulation and Sensing Platform for Sensorimotor Prosthetic Control, Management of Lower Back Exertion, and Electrical Muscle Activation. *Adv. Mater.* **2016**, *28* (22), 4462–4471.

(87) Lega, B. C.; Halpern, C. H.; Jaggi, J. L.; Baltuch, G. H. Deep Brain Stimulation in the Treatment of Refractory Epilepsy: Update on Current Data and Future Directions. *Neurobiology of Disease*. Academic Press June 1, 2010, pp 354–360.

(88) Benabid, A. L. Deep Brain Stimulation for Parkinson's Disease. *Current Opinion in Neurobiology*. Elsevier Ltd December 1, 2003, pp 696–706.

(89) Ud-Din, S.; Bayat, A.; Ud-Din, S.; Bayat, A. Electrical Stimulation and Cutaneous Wound Healing: A Review of Clinical Evidence. *Healthcare* **2014**, *2* (4), 445–467.

(90) Feiner, R.; Engel, L.; Fleischer, S.; Malki, M.; Gal, I.; Shapira, A.; Shacham-Diamand, Y.; Dvir, T. Engineered Hybrid Cardiac Patches with Multifunctional Electronics for Online Monitoring and Regulation of Tissue Function. *Nat. Mater.* **2016**, *15* (March), 1–8.

(91) Wang, S.; Xu, J.; Wang, W.; Wang, G. J. N.; Rastak, R.; Molina-Lopez, F.; Chung, J. W.; Niu, S.; Feig, V. R.; Lopez, J.; et al. Skin Electronics from Scalable Fabrication of an Intrinsically Stretchable Transistor Array. *Nature* **2018**, *555* (7694), 83–88.

(92) Liu, J.; Zhang, X.; Liu, Y.; Rodrigo, M.; Loftus, P. D.; Aparicio-Valenzuela, J.; Zheng, J.; Pong, T.; Cyr, K. J.; Babakhanian, M.; et al. Intrinsically Stretchable Electrode Array Enabled in Vivo Electrophysiological Mapping of Atrial Fibrillation at Cellular Resolution. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117* (26), 14769–14778.

(93) Dai, X.; Zhou, W.; Gao, T.; Liu, J.; Lieber, C. M. Three-Dimensional Mapping and Regulation of Action Potential Propagation in Nanoelectronics-Innervated Tissues. *Nat. Nanotechnol.* **2016**, *11* (9), 776–782.

(94) Ronaldson-bouchard, K.; Ma, S. P.; Yeager, K.; Chen, T.; Song, L.; Sirabella, D.; Morikawa, K.; Teles, D.; Yazawa, M.; Vunjak-Novakovic, G. Advanced Maturation of Human Cardiac Tissue Grown from Pluripotent Stem Cells. *Nature* **2018**, *556* (2), 239.

(95) Namsheer, K.; Rout, C. S. Conducting Polymers: A Comprehensive Review on Recent Advances in Synthesis, Properties and Applications. *RSC Advances*. Royal Society of Chemistry February 3, 2021, pp 5659–5697.

(96) Nezakati, T.; Seifalian, A.; Tan, A.; Seifalian, A. M. Conductive Polymers: Opportunities and Challenges in Biomedical Applications. *Chemical Reviews*. American Chemical Society July 25, 2018, pp 6766–6843.

(97) Alcácer, L. Orbitals and Bands in Organic Semiconductors. In *Electronic Structure of Organic Semiconductors: Polymers and small molecules*; IOP Publishing, 2018.

(98) Ouyang, J. “Secondary Doping” Methods to Significantly Enhance the Conductivity of PEDOT:PSS for Its Application as Transparent Electrode of Optoelectronic Devices. *Displays* **2013**, *34* (5), 423–436.

(99) Yu, J.; Vodyanik, M. A.; Smuga-Otto, K.; Antosiewicz-Bourget, J.; Frane, J. L.; Tian, S.; Nie, J.; Jonsdottir, G. A.; Ruotti, V.; Stewart, R.; et al. Induced Pluripotent Stem Cell Lines Derived from Human Somatic Cells. *Science (80-.).* **2007**, *318* (5858), 1917–1920.

(100) Wernig, M.; Meissner, A.; Foreman, R.; Brambrink, T.; Ku, M.; Hochedlinger, K.; Bernstein, B. E.; Jaenisch, R. In Vitro Reprogramming of Fibroblasts into a Pluripotent ES-Cell-like State. *Nature* **2007**, *448* (7151), 318–324.

(101) Maherali, N.; Sridharan, R.; Xie, W.; Utikal, J.; Eminli, S.; Arnold, K.; Stadtfeld, M.; Yachechko, R.; Tchieu, J.; Jaenisch, R.; et al. Directly Reprogrammed Fibroblasts Show Global Epigenetic Remodeling and Widespread Tissue Contribution. *Cell Stem Cell* **2007**, *1* (1), 55–70.

(102) Park, I.-H.; Zhao, R.; West, J. A.; Yabuuchi, A.; Huo, H.; Ince, T. A.; Lerou, P. H.; Lensch, M. W.; Daley, G. Q. Reprogramming of Human Somatic Cells to Pluripotency with Defined Factors. *Nature* **2008**, *451* (7175), 141–146.

(103) Sugimura, R.; Jha, D. K.; Han, A.; Soria-Valles, C.; da Rocha, E. L.; Lu, Y.-F.; Goettel, J. A.; Serrao, E.; Rowe, R. G.; Malleshaiah, M.; et al. Haematopoietic Stem and Progenitor Cells from Human Pluripotent Stem Cells. *Nature* **2017**, *545* (7655), 432–438.

(104) Kolodziejczyk, A. A.; Kim, J. K.; Svensson, V.; Marioni, J. C.; Teichmann, S. A. The Technology and Biology of Single-Cell RNA Sequencing. *Mol. Cell* **2015**, *58* (4), 610–

620.

(105) Mohr, S.; Bakal, C.; Perrimon, N. Genomic Screening with RNAi: Results and Challenges. *Annu. Rev. Biochem.* **2010**, *79*, 37–64.

(106) Shalem, O.; Sanjana, N. E.; Zhang, F. High-Throughput Functional Genomics Using CRISPR–Cas9. *Nat. Rev. Genet.* **2015**, *16* (5), 299–311.

(107) Adamson, B.; Norman, T. M.; Jost, M.; Cho, M. Y.; Nuñez, J. K.; Chen, Y.; Villalta, J. E.; Gilbert, L. A.; Horlbeck, M. A.; Hein, M. Y.; et al. A Multiplexed Single-Cell CRISPR Screening Platform Enables Systematic Dissection of the Unfolded Protein Response. *Cell* **2016**, *167* (7), 1867-1882.e21.

(108) Dixit, A.; Parnas, O.; Li, B.; Chen, J.; Fulco, C. P.; Jerby-Arnon, L.; Marjanovic, N. D.; Dionne, D.; Burks, T.; Raychowdhury, R.; et al. Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens. *Cell* **2016**, *167* (7), 1853-1866.e17.

(109) Jaitin, D. A.; Weiner, A.; Yofe, I.; Lara-Astiaso, D.; Keren-Shaul, H.; David, E.; Salame, T. M.; Tanay, A.; van Oudenaarden, A.; Amit, I. Dissecting Immune Circuits by Linking CRISPR-Pooled Screens with Single-Cell RNA-Seq. *Cell* **2016**, *167* (7), 1883-1896.e15.

(110) Xie, S.; Duan, J.; Li, B.; Zhou, P.; Hon, G. C. Multiplexed Engineering and Analysis of Combinatorial Enhancer Activity in Single Cells. *Mol. Cell* **2017**, *66* (2), 285-299.e5.

(111) Datlinger, P.; Rendeiro, A. F.; Schmidl, C.; Krausgruber, T.; Traxler, P.; Klughammer, J.; Schuster, L. C.; Kuchler, A.; Alpar, D.; Bock, C. Pooled CRISPR Screening with Single-Cell Transcriptome Readout. *Nat. Methods* **2017**, *14* (3), 297–301.

(112) La Russa, M. F.; Qi, L. S. The New State of the Art: Cas9 for Gene Activation and Repression. *Mol. Cell. Biol.* **2015**, *35* (22), 3800–3809.

(113) Dominguez, A. A.; Lim, W. A.; Qi, L. S. Beyond Editing: Repurposing CRISPR-Cas9 for Precision Genome Regulation and Interrogation. *Nat. Rev. Mol. Cell Biol.* **2016**, *17* (1), 5–15.

(114) Macosko, E. Z.; Basu, A.; Satija, R.; Nemesh, J.; Shekhar, K.; Goldman, M.; Tirosh, I.; Bialas, A. R.; Kamitaki, N.; Martersteck, E. M.; et al. Highly Parallel Genome-Wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* **2015**, *161* (5), 1202–1214.

(115) Blondel, V. D.; Guillaume, J.-L.; Lambiotte, R.; Lefebvre, E. Fast Unfolding of Communities in Large Networks. *arXiv* **2008**, 1–12.

(116) Li, W.; Xu, H.; Xiao, T.; Cong, L.; Love, M. I.; Zhang, F.; Irizarry, R. A.; Liu, J. S.; Brown, M.; Liu, X. S. MAGeCK Enables Robust Identification of Essential Genes from Genome-Scale CRISPR/Cas9 Knockout Screens. *Genome Biol.* **2014**, *15* (12), 554.

(117) Subramanian, A.; Tamayo, P.; Mootha, V. K.; Mukherjee, S.; Ebert, B. L.; Gillette, M. a; Paulovich, A.; Pomeroy, S. L.; Golub, T. R.; Lander, E. S.; et al. Gene Set Enrichment Analysis: A Knowledge-Based Approach for Interpreting Genome-Wide Expression

Profiles. *Proc. Natl. Acad. Sci.* **2005**, *102* (43), 15545–15550.

(118) Dobin, A.; Davis, C. A.; Schlesinger, F.; Drenkow, J.; Zaleski, C.; Jha, S.; Batut, P.; Chaisson, M.; Gingeras, T. R. STAR: Ultrafast Universal RNA-Seq Aligner. *Bioinformatics* **2013**, *29* (1), 15–21.

(119) Gao, Z.; Ure, K.; Ables, J. L.; Lagace, D. C.; Nave, K.-A.; Goebbels, S.; Eisch, A. J.; Hsieh, J. Neurod1 Is Essential for the Survival and Maturation of Adult-Born Neurons. *Nat. Neurosci.* **2009**, *12* (9), 1090–1092.

(120) Pataskar, A.; Jung, J.; Smialowski, P.; Noack, F.; Calegari, F.; Straub, T.; Tiwari, V. K. NeuroD1 Reprograms Chromatin and Transcription Factor Landscapes to Induce the Neuronal Program. *EMBO J.* **2016**, *35* (1), 24–45.

(121) Busskamp, V.; Lewis, N. E.; Guye, P.; Ng, A. H. M.; Shipman, S. L.; Byrne, S. M.; Sanjana, N. E.; Murn, J.; Li, Y.; Li, S.; et al. Rapid Neurogenesis through Transcriptional Activation in Human Stem Cells. *Mol. Syst. Biol.* **2014**, *10* (11), 760.

(122) Zhang, Y.; Pak, C.; Han, Y.; Ahlenius, H.; Zhang, Z.; Chanda, S.; Marro, S.; Patzke, C.; Acuna, C.; Covy, J.; et al. Rapid Single-Step Induction of Functional Neurons from Human Pluripotent Stem Cells. *Neuron* **2013**, *78* (5), 785–798.

(123) D'Angelo, A.; Bluteau, O.; Garcia-Gonzalez, M. A.; Gresh, L.; Doyen, A.; Garbay, S.; Robine, S.; Pontoglio, M.; Barra, J. Hepatocyte Nuclear Factor 1alpha and Beta Control Terminal Differentiation and Cell Fate Commitment in the Gut Epithelium. *Development* **2010**, *137* (9), 1573–1582.

(124) Servitja, J.-M.; Pignatelli, M.; Maestro, M. A.; Cardalda, C.; Boj, S. F.; Lozano, J.; Blanco, E.; Lafuente, A.; McCarthy, M. I.; Sumoy, L.; et al. Hnf1alpha (MODY3) Controls Tissue-Specific Transcriptional Programs and Exerts Opposed Effects on Cell Growth in Pancreatic Islets and Liver. *Mol. Cell. Biol.* **2009**, *29* (11), 2945–2959.

(125) Maestro, M. A.; Cardalda, C.; Boj, S. F.; Luco, R. F.; Servitja, J. M.; Ferrer, J. Distinct Roles of HNF1beta, HNF1alpha, and HNF4alpha in Regulating Pancreas Development, Beta-Cell Function and Growth. *Endocr. Dev.* **2007**, *12*, 33–45.

(126) Si-Tayeb, K.; Lemaigre, F. P.; Duncan, S. A. Organogenesis and Development of the Liver. *Dev. Cell* **2010**, *18* (2), 175–189.

(127) Clotman, F.; Jacquemin, P.; Plumb-Rudewiez, N.; Pierreux, C. E.; Van der Smissen, P.; Dietz, H. C.; Courtoy, P. J.; Rousseau, G. G.; Lemaigre, F. P. Control of Liver Cell Fate Decision by a Gradient of TGF Beta Signaling Modulated by Onecut Transcription Factors. *Genes Dev.* **2005**, *19* (16), 1849–1854.

(128) Pierreux, C. E.; Poll, A. V.; Kemp, C. R.; Clotman, F.; Maestro, M. A.; Cordi, S.; Ferrer, J.; Leyns, L.; Rousseau, G. G.; Lemaigre, F. P. The Transcription Factor Hepatocyte Nuclear Factor-6 Controls the Development of Pancreatic Ducts in the Mouse. *Gastroenterology* **2006**, *130* (2), 532–541.

(129) Jacquemin, P.; Lemaigre, F. P.; Rousseau, G. G. The Onecut Transcription Factor HNF-6 (OC-1) Is Required for Timely Specification of the Pancreas and Acts Upstream of Pdx-1

in the Specification Cascade. *Dev. Biol.* **2003**, *258* (1), 105–116.

(130) Jacquemin, P.; Durviaux, S. M.; Jensen, J.; Godfraind, C.; Gradwohl, G.; Guillemot, F.; Madsen, O. D.; Carmeliet, P.; Dewerchin, M.; Collen, D.; et al. Transcription Factor Hepatocyte Nuclear Factor 6 Regulates Pancreatic Endocrine Cell Differentiation and Controls Expression of the Proendocrine Gene Ngn3. *Mol. Cell. Biol.* **2000**, *20* (12), 4445–4454.

(131) Niwa, H.; Toyooka, Y.; Shimosato, D.; Strumpf, D.; Takahashi, K.; Yagi, R.; Rossant, J. Interaction between Oct3/4 and Cdx2 Determines Trophectoderm Differentiation. *Cell* **2005**, *123* (5), 917–929.

(132) Strumpf, D.; Mao, C.-A.; Yamanaka, Y.; Ralston, A.; Chawengsaksophak, K.; Beck, F.; Rossant, J. Cdx2 Is Required for Correct Cell Fate Specification and Differentiation of Trophectoderm in the Mouse Blastocyst. *Development* **2005**, *132* (9), 2093–2102.

(133) Nishiyama, A.; Xin, L.; Sharov, A. A.; Thomas, M.; Mowrer, G.; Meyers, E.; Piao, Y.; Mehta, S.; Yee, S.; Nakatake, Y.; et al. Uncovering Early Response of Gene Regulatory Networks in ESCs by Systematic Induction of Transcription Factors. *Cell Stem Cell* **2009**, *5*, 420–433.

(134) Liu, Y.-N.; Abou-Kheir, W.; Yin, J. J.; Fang, L.; Hynes, P.; Casey, O.; Hu, D.; Wan, Y.; Seng, V.; Sheppard-Tillman, H.; et al. Critical and Reciprocal Regulation of KLF4 and SLUG in Transforming Growth Factor β-Initiated Prostate Cancer Epithelial-Mesenchymal Transition. *Mol. Cell. Biol.* **2012**, *32* (5), 941–953.

(135) Li, R.; Liang, J.; Ni, S.; Zhou, T.; Qing, X.; Li, H.; He, W.; Chen, J.; Li, F.; Zhuang, Q.; et al. A Mesenchymal-to-Epithelial Transition Initiates and Is Required for the Nuclear Reprogramming of Mouse Fibroblasts. *Cell Stem Cell* **2010**, *7* (1), 51–63.

(136) Barrallo-Gimeno, A.; Nieto, M. A.; Ip, Y. T. The Snail Genes as Inducers of Cell Movement and Survival: Implications in Development and Cancer. *Development* **2005**, *132* (14), 3151–3161.

(137) Morita, R.; Suzuki, M.; Kasahara, H.; Shimizu, N.; Shichita, T.; Sekiya, T.; Kimura, A.; Sasaki, K.; Yasukawa, H.; Yoshimura, A. ETS Transcription Factor ETV2 Directly Converts Human Fibroblasts into Functional Endothelial Cells. *Proc. Natl. Acad. Sci.* **2015**, *112* (1), 160–165.

(138) Lindgren, A. G.; Veldman, M. B.; Lin, S. ETV2 Expression Increases the Efficiency of Primitive Endothelial Cell Derivation from Human Embryonic Stem Cells. *Cell Regen.* **2015**, *4* (1), 4:1.

(139) Tsang, K. M.; Hyun, J. S.; Cheng, K. T.; Vargas, M.; Mehta, D.; Ushio-Fukai, M.; Zou, L.; Pajcini, K. V; Rehman, J.; Malik, A. B. Embryonic Stem Cell Differentiation to Functional Arterial Endothelial Cells through Sequential Activation of ETV2 and NOTCH1 Signaling by HIF1α. *Stem Cell Reports* **2017**, *9* (3), 796–806.

(140) Pelengaris, S.; Khan, M.; Evan, G. C-MYC: More than Just a Matter of Life and Death. *Nat. Rev. Cancer* **2002**, *2* (10), 764–776.

(141) McConnell, B. B.; Yang, V. W. Mammalian Krüppel-like Factors in Health and Diseases. *Physiol. Rev.* **2010**, *90* (4), 1337–1381.

(142) Tiwari, N.; Meyer-Schaller, N.; Arnold, P.; Antoniadis, H.; Pachkov, M.; van Nimwegen, E.; Christofori, G. Klf4 Is a Transcriptional Regulator of Genes Critical for EMT, Including Jnk1 (Mapk8). *PLoS One* **2013**, *8* (2), e57329.

(143) Zhang, B.; Zhang, Z.; Xia, S.; Xing, C.; Ci, X.; Li, X.; Zhao, R.; Tian, S.; Ma, G.; Zhu, Z.; et al. KLF5 Activates MicroRNA 200 Transcription to Maintain Epithelial Characteristics and Prevent Induced Epithelial-Mesenchymal Transition in Epithelial Cells. *Mol. Cell. Biol.* **2013**, *33* (24), 4919–4935.

(144) Gumireddy, K.; Li, A.; Gimotty, P. A.; Klein-Szanto, A. J.; Showe, L. C.; Katsaros, D.; Coukos, G.; Zhang, L.; Huang, Q. KLF17 Is a Negative Regulator of Epithelial-Mesenchymal Transition and Metastasis in Breast Cancer. *Nat. Cell Biol.* **2009**, *11* (11), 1297–1304.

(145) Sack, L. M.; Davoli, T.; Xu, Q.; Li, M. Z.; Elledge, S. J. Sources of Error in Mammalian Genetic Screens. *G3* **2016**, *6* (9), 2781–2790.

(146) Hill, A. J.; McFaline-Figueroa, J. L.; Starita, L. M.; Gasperini, M. J.; Matreyek, K. A.; Packer, J.; Jackson, D.; Shendure, J.; Trapnell, C. On the Design of CRISPR-Based Single-Cell Molecular Screens. *Nat. Methods* **2018**, *15* (4), 271–274.

(147) Xie, S.; Cooley, A.; Armendariz, D.; Zhou, P.; Hon, G. C. Frequent SgRNA-Barcode Recombination in Single-Cell Perturbation Assays. *PLoS One* **2018**, *13* (6), e0198635.

(148) Adamson, B.; Norman, T. M.; Jost, M.; Weissman, J. S. Approaches to Maximize SgRNA-Barcode Coupling in Perturb-Seq Screens. *bioRxiv* **2018**, 298349.

(149) Santoni de Sio, F. R.; Gritti, A.; Cascio, P.; Neri, M.; Sampaolesi, M.; Galli, C.; Luban, J.; Naldini, L. Lentiviral Vector Gene Transfer Is Limited by the Proteasome at Postentry Steps in Various Types of Stem Cells. *Stem Cells* **2008**, *26* (8), 2142–2152.

(150) Geis, F. K.; Galla, M.; Hoffmann, D.; Kuehle, J.; Zychlinski, D.; Maetzig, T.; Schott, J. W.; Schwarzer, A.; Goffinet, C.; Goff, S. P.; et al. Potent and Reversible Lentiviral Vector Restriction in Murine Induced Pluripotent Stem Cells. *Retrovirology* **2017**, *14* (1), 34.

(151) Cao, J.; Packer, J. S.; Ramani, V.; Cusanovich, D. A.; Huynh, C.; Daza, R.; Qiu, X.; Lee, C.; Furlan, S. N.; Steemers, F. J.; et al. Comprehensive Single-Cell Transcriptional Profiling of a Multicellular Organism. *Science (80-.).* **2017**, *357* (6352), 661–667.

(152) Rosenberg, A. B.; Roco, C. M.; Muscat, R. A.; Kuchina, A.; Sample, P.; Yao, Z.; Graybuck, L. T.; Peeler, D. J.; Mukherjee, S.; Chen, W.; et al. Single-Cell Profiling of the Developing Mouse Brain and Spinal Cord with Split-Pool Barcoding. *Science (80-.).* **2018**, *360* (6385), 176–182.

(153) Merlo, L. M. F.; Pepper, J. W.; Reid, B. J.; Maley, C. C. Cancer as an Evolutionary and Ecological Process. *Nature Reviews Cancer*. Nature Publishing Group December 9, 2006, pp 924–935.

(154) Greaves, M.; Maley, C. C. Clonal Evolution in Cancer. *Nature*. Nature Publishing Group January 19, 2012, pp 306–313.

(155) Hagis, K. M.; Cichowski, K.; Elledge, S. J. Tissue-Specificity in Cancer: The Rule, Not the Exception. *Science (80-.).* **2019**, *363* (6432), 1150–1151.

(156) Schneider, G.; Schmidt-Supprian, M.; Rad, R.; Saur, D. Tissue-Specific Tumorigenesis: Context Matters. *Nature Reviews Cancer*. Nature Publishing Group April 1, 2017, pp 239–253.

(157) Beroukhim, R.; Mermel, C. H.; Porter, D.; Wei, G.; Raychaudhuri, S.; Donovan, J.; Barretina, J.; Boehm, J. S.; Dobson, J.; Urashima, M.; et al. The Landscape of Somatic Copy-Number Alteration across Human Cancers. *Nature* **2010**, *463* (7283), 899–905.

(158) Matthew Bailey, A. H.; Tokheim, C.; Porta-Pardo, E.; Mills, G. B.; Karchin, R.; Ding, L.; Bailey, M. H.; Sengupta, S.; Bertrand, D.; Weerasinghe, A.; et al. Comprehensive Characterization of Cancer Driver Genes and Mutations Article Comprehensive Characterization of Cancer Driver Genes and Mutations. *Cell* **2018**, *173*, 371-376.e18.

(159) Alexandrov, L. B.; Nik-Zainal, S.; Wedge, D. C.; Aparicio, S. A. J. R.; Behjati, S.; Biankin, A. V.; Bignell, G. R.; Bolli, N.; Borg, A.; Børresen-Dale, A. L.; et al. Signatures of Mutational Processes in Human Cancer. *Nature* **2013**, *500* (7463), 415–421.

(160) Forbes, S. A.; Beare, D.; Boutselakis, H.; Bamford, S.; Bindal, N.; Tate, J.; Cole, C. G.; Ward, S.; Dawson, E.; Ponting, L.; et al. COSMIC: Somatic Cancer Genetics at High-Resolution. *Nucleic Acids Res.* **2017**, *45* (D1), D777–D783.

(161) Zack, T. I.; Schumacher, S. E.; Carter, S. L.; Cherniack, A. D.; Saksena, G.; Tabak, B.; Lawrence, M. S.; Zhang, C. Z.; Wala, J.; Mermel, C. H.; et al. Pan-Cancer Patterns of Somatic Copy Number Alteration. *Nat. Genet.* **2013**, *45* (10), 1134–1140.

(162) Sanchez-Vega, F.; Mina, M.; Armenia, J.; Chatila, W. K.; Luna, A.; La, K. C.; Dimitriadoy, S.; Liu, D. L.; Kantheti, H. S.; Saghafinia, S.; et al. Oncogenic Signaling Pathways in The Cancer Genome Atlas. *Cell* **2018**, *173* (2), 321-337.e10.

(163) Etzioni, R.; Urban, N.; Ramsey, S.; McIntosh, M.; Schwartz, S.; Reid, B.; Radich, J.; Anderson, G.; Hartwell, L. The Case for Early Detection. *Nature Reviews Cancer*. European Association for Cardio-Thoracic Surgery 2003, pp 243–252.

(164) Balani, S.; Nguyen, L. V.; Eaves, C. J. Modeling the Process of Human Tumorigenesis. *Nature Communications*. Nature Publishing Group May 25, 2017, pp 1–10.

(165) Richmond, A.; Yingjun, S. Mouse Xenograft Models vs GEM Models for Human Cancer Therapeutics. *DMM Disease Models and Mechanisms*. Company of Biologists 2008, pp 78–82.

(166) Martz, C. A.; Ottina, K. A.; Singleton, K. R.; Jasper, J. S.; Wardell, S. E.; Peraza-penton, A.; Anderson, G. R.; Winter, P. S.; Wang, T.; Alley, H. M.; et al. Systematic Identification of Signaling Pathways with Potential to Confer Anticancer Drug Resistance. *Sci. Signal.* **2014**, *7* (357), 1–14.

(167) Hahn, W. C.; Counter, C. M.; Lundberg, A. S.; Beijersbergen, R. L.; Brooks, M. W.; Weinberg, R. A. Creation of Human Tumour Cells with Defined Genetic Elements. *Nature* **1999**, *400* (6743), 464–468.

(168) Elenbaas, B.; Spirio, L.; Koerner, F.; Fleming, M. D.; Zimonjic, D. B.; Donaher, J. L.; Popescu, N. C.; Hahn, W. C.; Weinberg, R. A. Human Breast Cancer Cells Generated by Oncogenic Transformation of Primary Mammary Epithelial Cells. *Genes Dev.* **2001**, *15* (1), 50–65.

(169) Daley, G. Q.; Mclaughlin, J.; Witte, O. N.; Baltimore, D. The CML-Specific P210 Bcr/Abl Protein, Unlike v-Abl, Does Not Transform NIH/3T3 Fibroblasts. *Science (80-.).* **1987**, *237* (4814), 532–535.

(170) Clark, R.; Stampfer, M. R.; Milley, R.; O'rourke, E.; Walen, K. H.; Kriegler, M.; Kopplin, J.; Mccormick, F. *Transformation of Human Mammary Epithelial Cells by Oncogenic Retroviruses '*; 1988; Vol. 48.

(171) Sasaki, R.; Narisawa-Saito, M.; Yugawa, T.; Fujita, M.; Tashiro, H.; Katabuchi, H.; Kiyono, T. Oncogenic Transformation of Human Ovarian Surface Epithelial Cells with Defined Cellular Oncogenes. *Carcinogenesis* **2009**, *30* (3), 423–431.

(172) Boehm, J. S.; Hession, M. T.; Bulmer, S. E.; Hahn, W. C. Transformation of Human and Murine Fibroblasts without Viral Oncoproteins. *Mol. Cell. Biol.* **2005**, *25* (15), 6464–6474.

(173) Geder, L.; Lausch, R.; O'Neill, F.; Rapp, F. Oncogenic Transformation of Human Embryo Lung Cells by Human Cytomegalovirus. *Science (80-.).* **1976**, *192* (4244), 1134–1137.

(174) Park, J. W.; Lee, J. K.; Sheu, K. M.; Wang, L.; Balanis, N. G.; Nguyen, K.; Smith, B. A.; Cheng, C.; Tsai, B. L.; Cheng, D.; et al. Reprogramming Normal Human Epithelial Tissues to a Common, Lethal Neuroendocrine Cancer Lineage. *Science* **2018**, *362* (6410), 91–95.

(175) Sack, L. M.; Davoli, T.; Li, M. Z.; Li, Y.; Xu, Q.; Naxerova, K.; Wooten, E. C.; Bernardi, R. J.; Martin, T. D.; Chen, T.; et al. Profound Tissue Specificity in Proliferation Control Underlies Cancer Drivers and Aneuploidy Patterns. *Cell* **2018**, *0* (0), 1–16.

(176) Puisieux, A.; Pommier, R. M.; Morel, A. P.; Lavial, F. Cellular Pliancy and the Multistep Process of Tumorigenesis. *Cancer Cell*. Cell Press February 12, 2018, pp 164–172.

(177) Drost, J.; Van Boxtel, R.; Blokzijl, F.; Mizutani, T.; Sasaki, N.; Sasselli, V.; De Ligt, J.; Behjati, S.; Grolleman, J. E.; Van Wezel, T.; et al. Use of CRISPR-Modified Human Stem Cell Organoids to Study the Origin of Mutational Signatures in Cancer. *Science (80-.).* **2017**, *358* (6360), 234–238.

(178) Fumagalli, A.; Drost, J.; Suijkerbuijk, S. J. E.; Van Boxtel, R.; De Ligt, J.; Offerhaus, G. J.; Begthel, H.; Beerling, E.; Tan, E. H.; Sansom, O. J.; et al. Genetic Dissection of Colorectal Cancer Progression by Orthotopic Transplantation of Engineered Cancer Organoids. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (12), E2357–E2364.

(179) Rosenbluth, J. M.; Schackmann, R. C. J.; Gray, G. K.; Selfors, L. M.; Li, C. M. C.; Boedicker, M.; Kuiken, H. J.; Richardson, A.; Brock, J.; Garber, J.; et al. Organoid

Cultures from Normal and Cancer-Prone Human Breast Tissues Preserve Complex Epithelial Lineages. *Nat. Commun.* **2020**, *11* (1), 1–14.

(180) Li, X.; Nadauld, L.; Ootani, A.; Corney, D. C.; Pai, R. K.; Gevaert, O.; Cantrell, M. A.; Rack, P. G.; Neal, J. T.; Chan, C. W. M.; et al. Oncogenic Transformation of Diverse Gastrointestinal Tissues in Primary Organoid Culture. *Nat. Med.* **2014**, *20* (7), 769–777.

(181) Lannagan, T. R. M.; Lee, Y. K.; Wang, T.; Roper, J.; Bettington, M. L.; Fennell, L.; Vrbanac, L.; Jonavicius, L.; Somashekar, R.; Gieniec, K.; et al. Genetic Editing of Colonic Organoids Provides a Molecularly Distinct and Orthotopic Preclinical Model of Serrated Carcinogenesis. *Gut* **2019**, *68* (4), 684–692.

(182) Hanahan, D.; Weinberg, R. A. Hallmarks of Cancer: The next Generation. *Cell* **2011**, *144* (5), 646–674.

(183) Koga, T.; Chaim, I. A.; Benitez, J. A.; Markmiller, S.; Parisian, A. D.; Hevner, R. F.; Turner, K. M.; Hessenauer, F. M.; D'Antonio, M.; Nguyen, N. phuong D.; et al. Longitudinal Assessment of Tumor Development Using Cancer Avatars Derived from Genetically Engineered Pluripotent Stem Cells. *Nat. Commun.* **2020**, *11* (1), 1–14.

(184) Duan, S.; Yuan, G.; Liu, X.; Ren, R.; Li, J.; Zhang, W.; Wu, J.; Xu, X.; Fu, L.; Li, Y.; et al. PTEN Deficiency Reprogrammes Human Neural Stem Cells towards a Glioblastoma Stem Cell-like Phenotype. *Nat. Commun.* **2015**, *6* (1), 1–14.

(185) Pei, Y.; Moore, C. E.; Wang, J.; Tewari, A. K.; Eroshkin, A.; Cho, Y. J.; Witt, H.; Korshunov, A.; Read, T. A.; Sun, J. L.; et al. An Animal Model of MYC-Driven Medulloblastoma. *Cancer Cell* **2012**, *21* (2), 155–167.

(186) Pei, Y.; Liu, K. W.; Wang, J.; Garancher, A.; Tao, R.; Esparza, L. A.; Maier, D. L.; Udaka, Y. T.; Murad, N.; Morrissy, S.; et al. HDAC and PI3K Antagonists Cooperate to Inhibit Growth of MYC-Driven Medulloblastoma. *Cancer Cell* **2016**, *29* (3), 311–323.

(187) Lensch, M. W.; Schlaeger, T. M.; Zon, L. I.; Daley, G. Q. Teratoma Formation Assays with Human Embryonic Stem Cells: A Rationale for One Type of Human-Animal Chimera. *Cell Stem Cell*. Elsevier Inc. September 13, 2007, pp 253–258.

(188) Parekh, U.; Wu, Y.; Zhao, D.; Worlikar, A.; Shah, N.; Zhang, K.; Mali, P. Mapping Cellular Reprogramming via Pooled Overexpression Screens with Paired Fitness and Single-Cell RNA-Sequencing Readout. *Cell Syst.* **2018**, *7* (5), 548-555.e8.

(189) Hayer, A.; Shao, L.; Chung, M.; Joubert, L. M.; Yang, H. W.; Tsai, F. C.; Bisaria, A.; Betzig, E.; Meyer, T. Engulfed Cadherin Fingers Are Polarized Junctional Structures between Collectively Migrating Endothelial Cells. *Nat. Cell Biol.* **2016**, *18* (12), 1311–1323.

(190) Kim, E.; Ilic, N.; Shrestha, Y.; Zou, L.; Kamburov, A.; Zhu, C.; Yang, X.; Lubonja, R.; Tran, N.; Nguyen, C.; et al. Systematic Functional Interrogation of Rare Cancer Variants Identifies Oncogenic Alleles. *Cancer Discov.* **2016**, *6* (7), 714–726.

(191) Hagting, A.; Karlsson, C.; Clute, P.; Jackman, M.; Pines, J. MPF Localization Is Controlled by Nuclear Export. *EMBO J.* **1998**, *17* (14), 4127–4138.

(192) Johannessen, C. M.; Boehm, J. S.; Kim, S. Y.; Thomas, S. R.; Wardwell, L.; Johnson, L. A.; Emery, C. M.; Stransky, N.; Cogdill, A. P.; Barretina, J.; et al. COT Drives Resistance to RAF Inhibition through MAP Kinase Pathway Reactivation. *Nature* **2010**, *468* (7326), 968–972.

(193) Ramaswamy, S.; Nakamura, N.; Vazquez, F.; Batt, D. B.; Perera, S.; Roberts, T. M.; Sellers, W. R. Regulation of G1 Progression by the PTEN Tumor Suppressor Protein Is Linked to Inhibition of the Phosphatidylinositol 3-Kinase/Akt Pathway. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (5), 2110–2115.

(194) Roberts, P. J.; Mitin, N.; Keller, P. J.; Chenette, E. J.; Madigan, J. P.; Currin, R. O.; Cox, A. D.; Wilson, O.; Kirschmeier, P.; Der, C. J. Rho Family GTPase Modification and Dependence on CAAX Motif-Signaled Posttranslational Modification. *J. Biol. Chem.* **2008**, *283* (37), 25150–25163.

(195) Zheng, G. X. Y.; Terry, J. M.; Belgrader, P.; Ryvkin, P.; Bent, Z. W.; Wilson, R.; Ziraldo, S. B.; Wheeler, T. D.; McDermott, G. P.; Zhu, J.; et al. Massively Parallel Digital Transcriptional Profiling of Single Cells. *Nat. Commun.* **2017**, *8*, 1–12.

(196) 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* (5), 411–420.

(197) McGinnis, C. S.; Murrow, L. M.; Gartner, Z. J. DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Syst.* **2019**, *8* (4), 329-337.e4.

(198) Robinson, M. D.; McCarthy, D. J.; Smyth, G. K. EdgeR: A Bioconductor Package for Differential Expression Analysis of Digital Gene Expression Data. *Bioinformatics* **2010**, *26* (1), 139–140.

(199) Ritchie, M. E.; Dai, Z.; Sheridan, J. M.; Gearing, L. J.; Moore, D. L.; Su, S.; Wormald, S.; Wilcox, S.; O'Connor, L.; Dickins, R. A.; et al. EdgeR: A Versatile Tool for the Analysis of ShRNA-Seq and CRISPR-Cas9 Genetic Screens. *F1000Research* **2014**, *3*.

(200) Love, M. I.; Huber, W.; Anders, S. Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2. *Genome Biol.* **2014**, *15* (12), 550.

(201) Ritchie, M. E.; Phipson, B.; Wu, D.; Hu, Y.; Law, C. W.; Shi, W.; Smyth, G. K. Limma Powers Differential Expression Analyses for RNA-Sequencing and Microarray Studies. *Nucleic Acids Res.* **2015**, *43* (7), e47.

(202) Langmead, B.; Trapnell, C.; Pop, M.; Salzberg, S. L. Ultrafast and Memory-Efficient Alignment of Short DNA Sequences to the Human Genome. *Genome Biol.* **2009**, *10* (3), R25.

(203) Langmead, B.; Salzberg, S. L. Fast Gapped-Read Alignment with Bowtie 2. *Nat. Methods* **2012**, *9* (4), 357–359.

(204) Fan, J.; Slowikowski, K.; Zhang, F. Single-Cell Transcriptomics in Cancer: Computational Challenges and Opportunities. *Experimental and Molecular Medicine*. Springer Nature

September 1, 2020, pp 1452–1465.

(205) Yuan, J.; Levitin, H. M.; Frattini, V.; Bush, E. C.; Boyett, D. M.; Samanamud, J.; Ceccarelli, M.; Dovas, A.; Zanazzi, G.; Canoll, P.; et al. Single-Cell Transcriptome Analysis of Lineage Diversity in High-Grade Glioma. *Genome Med.* **2018**, *10* (1), 57.

(206) Couturier, C. P.; Ayyadhury, S.; Le, P. U.; Nadaf, J.; Monlong, J.; Riva, G.; Allache, R.; Baig, S.; Yan, X.; Bourgey, M.; et al. Single-Cell RNA-Seq Reveals That Glioblastoma Recapitulates a Normal Neurodevelopmental Hierarchy. *Nat. Commun.* **2020**, *11* (1), 3406.

(207) Vladoiu, M. C.; El-Hamamy, I.; Donovan, L. K.; Farooq, H.; Holgado, B. L.; Sundaravadanam, Y.; Ramaswamy, V.; Hendrikse, L. D.; Kumar, S.; Mack, S. C.; et al. Childhood Cerebellar Tumours Mirror Conserved Fetal Transcriptional Programs. *Nature* **2019**, *572* (7767), 67–73.

(208) Hovestadt, V.; Smith, K. S.; Bihannic, L.; Filbin, M. G.; Shaw, M. K. L.; Baumgartner, A.; DeWitt, J. C.; Groves, A.; Mayr, L.; Weisman, H. R.; et al. Resolving Medulloblastoma Cellular Architecture by Single-Cell Genomics. *Nature* **2019**, *572* (7767), 74–79.

(209) Ho, B.; Johann, P. D.; Grabovska, Y.; De Dieu Andrianteranagna, M. J.; Yao, F.; Frühwald, M.; Hasselblatt, M.; Bourdeaut, F.; Williamson, D.; Huang, A.; et al. Molecular Subgrouping of Atypical Teratoid/Rhabdoid Tumors—a Reinvestigation and Current Consensus. *Neuro. Oncol.* **2020**, *22* (5), 613–624.

(210) Alimova, I.; Pierce, A.; Danis, E.; Donson, A.; Birks, D. K.; Griesinger, A.; Foreman, N. K.; Santi, M.; Soucek, L.; Venkataraman, S.; et al. Inhibition of *MYC* Attenuates Tumor Cell Self-renewal and Promotes Senescence in SMARCB1-deficient Group 2 Atypical Teratoid Rhabdoid Tumors to Suppress Tumor Growth *in Vivo*. *Int. J. Cancer* **2019**, *144* (8), 1983–1995.

(211) Gravina, G. L.; Festuccia, C.; Popov, V. M.; Di Rocco, A.; Colapietro, A.; Sanità, P.; Delle Monache, S.; Musio, D.; De Felice, F.; Di Cesare, E.; et al. C-Myc Sustains Transformed Phenotype and Promotes Radioresistance of Embryonal Rhabdomyosarcoma Cell Lines. *Radiat. Res.* **2016**, *185* (4), 411–422.

(212) Zhang, J.; Song, N.; Zang, D.; Yu, J.; Li, J.; Di, W.; Guo, R.; Zhao, W.; Wang, H. C-Myc Promotes Tumor Proliferation and Anti-Apoptosis by Repressing P21 in Rhabdomyosarcomas. *Mol. Med. Rep.* **2017**, *16* (4), 4089–4094.

(213) Kouraklis, G.; Triche, T. J.; Tsokos, M.; Wesley, R. D. Myc Oncogene Expression and Nude Mouse Tumorigenicity and Metastasis Formation Are Higher in Alveolar than Embryonal Rhabdomyosarcoma Cell Lines. *Pediatr. Res.* **1999**, *45* (4), 552–558.

(214) Bouchard, C.; Marquardt, J.; Brás, A.; Medema, R. H.; Eilers, M. Myc-Induced Proliferation and Transformation Require Akt-Mediated Phosphorylation of FoxO Proteins. *EMBO J.* **2004**, *23* (14), 2830–2840.

(215) Adameyko, I.; Lallemend, F.; Aquino, J. B.; Pereira, J. A.; Topilko, P.; Müller, T.; Fritz, N.; Beljajeva, A.; Mochii, M.; Liste, I.; et al. Schwann Cell Precursors from Nerve Innervation Are a Cellular Origin of Melanocytes in Skin. *Cell* **2009**, *139* (2), 366–379.

(216) Brunet, A.; Pagès, G.; Pouysségur, J. Constitutively Active Mutants of MAP Kinase Kinase (MEK1) Induce Growth Factor-Relaxation and Oncogenicity When Expressed in Fibroblasts. *Oncogene* **1994**, *9* (11), 3379–3387.

(217) Mansour, S. J.; Matten, W. T.; Hermann, A. S.; Candia, J. M.; Rong, S.; Fukasawa, K.; Vande Woude, G. F.; Ahn, N. G. Transformation of Mammalian Cells by Constitutively Active MAP Kinase Kinase. *Science (80-.).* **1994**, *265* (5174), 966–970.

(218) Cowley, S.; Paterson, H.; Kemp, P.; Marshall, C. J. Activation of MAP Kinase Kinase Is Necessary and Sufficient for PC12 Differentiation and for Transformation of NIH 3T3 Cells. *Cell* **1994**, *77* (6), 841–852.

(219) Zhang, J.; Tao, Z.; Wang, Y. Long Non-coding RNA DANCR Regulates the Proliferation and Osteogenic Differentiation of Human Bone-derived Marrow Mesenchymal Stem Cells via the P38 MAPK Pathway. *Int. J. Mol. Med.* **2018**, *41* (1), 213–219.

(220) Jin, S. J.; Jin, M. Z.; Xia, B. R.; Jin, W. L. Long Non-Coding RNA DANCR as an Emerging Therapeutic Target in Human Cancers. *Frontiers in Oncology*. Frontiers Media S.A. November 15, 2019, p 1225.

(221) Stine, Z. E.; Walton, Z. E.; Altman, B. J.; Hsieh, A. L.; Dang, C. V. MYC, Metabolism, and Cancer. *Cancer Discovery*. American Association for Cancer Research Inc. October 1, 2015, pp 1024–1039.

(222) Hoxhaj, G.; Manning, B. D. The PI3K–AKT Network at the Interface of Oncogenic Signalling and Cancer Metabolism. *Nature Reviews Cancer*. Nature Research February 1, 2020, pp 74–88.

(223) Stine, Z. E.; Walton, Z. E.; Altman, B. J.; Hsieh, A. L.; Dang, C. V. MYC, Metabolism, and Cancer. *Cancer Discovery*. American Association for Cancer Research Inc. October 1, 2015, pp 1024–1039.

(224) Liu, Y. C.; Li, F.; Handler, J.; Huang, C. R. L.; Xiang, Y.; Neretti, N.; Sedivy, J. M.; Zeller, K. I.; Dang, C. V. Global Regulation of Nucleotide Biosynthetic Genes by C-Myc. *PLoS One* **2008**, *3* (7).

(225) Wang, X.; Yang, K.; Xie, Q.; Wu, Q.; Mack, S. C.; Shi, Y.; Kim, L. J. Y.; Prager, B. C.; Flavahan, W. A.; Liu, X.; et al. Purine Synthesis Promotes Maintenance of Brain Tumor Initiating Cells in Glioma. *Nat. Neurosci.* **2017**, *20* (5), 661–673.

(226) Barfeld, S. J.; Fazli, L.; Persson, M.; Marjavaara, L.; Urbanucci, A.; Kaukoniemi, K. M.; Rennie, P. S.; Ceder, Y.; Chabes, A.; Visakorpi, T.; et al. Myc-Dependent Purine Biosynthesis Affects Nucleolar Stress and Therapy Response in Prostate Cancer. *Oncotarget* **2015**, *6* (14), 12587–12602.

(227) O'Flanagan, C. H.; Campbell, K. R.; Zhang, A. W.; Kabeer, F.; Lim, J. L. P.; Biele, J.; Eirew, P.; Lai, D.; McPherson, A.; Kong, E.; et al. Dissociation of Solid Tumor Tissues with Cold Active Protease for Single-Cell RNA-Seq Minimizes Conserved Collagenase-Associated Stress Responses. *Genome Biol.* **2019**, *20* (1), 210.

(228) Denisenko, E.; Guo, B. B.; Jones, M.; Hou, R.; De Kock, L.; Lassmann, T.; Poppe, D.;

Poppe, D.; Clément, O.; Simmons, R. K.; et al. Systematic Assessment of Tissue Dissociation and Storage Biases in Single-Cell and Single-Nucleus RNA-Seq Workflows. *Genome Biol.* **2020**, *21* (1), 130.

(229) Van Den Brink, S. C.; Sage, F.; Vértesy, Á.; Spanjaard, B.; Peterson-Maduro, J.; Baron, C. S.; Robin, C.; Van Oudenaarden, A. Single-Cell Sequencing Reveals Dissociation-Induced Gene Expression in Tissue Subpopulations. *Nature Methods*. Nature Publishing Group October 1, 2017, pp 935–936.

(230) Walsh, N. C.; Kenney, L. L.; Jangalwe, S.; Aryee, K. E.; Greiner, D. L.; Brehm, M. A.; Shultz, L. D. Humanized Mouse Models of Clinical Disease. *Annual Review of Pathology: Mechanisms of Disease*. Annual Reviews Inc. January 24, 2017, pp 187–215.

(231) Rodriques, S. G.; Stickels, R. R.; Goeva, A.; Martin, C. A.; Murray, E.; Vanderburg, C. R.; Welch, J.; Chen, L. M.; Chen, F.; Macosko, E. Z. Slide-Seq: A Scalable Technology for Measuring Genome-Wide Expression at High Spatial Resolution. *Science (80-.).* **2019**, *363* (6434), 1463–1467.

(232) Vickovic, S.; Eraslan, G.; Salmén, F.; Klughammer, J.; Stenbeck, L.; Schapiro, D.; Äijö, T.; Bonneau, R.; Bergenstråhle, L.; Navarro, J. F.; et al. High-Definition Spatial Transcriptomics for in Situ Tissue Profiling. *Nat. Methods* **2019**, *16* (10), 987–990.

(233) Ståhl, P. L.; Salmén, F.; Vickovic, S.; Lundmark, A.; Navarro, J. F.; Magnusson, J.; Giacomello, S.; Asp, M.; Westholm, J. O.; Huss, M.; et al. Visualization and Analysis of Gene Expression in Tissue Sections by Spatial Transcriptomics. *Science*. American Association for the Advancement of Science July 1, 2016, pp 78–82.

(234) Takebe, T.; Wells, J. M. Organoids by Design. *Science*. American Association for the Advancement of Science June 7, 2019, pp 956–959.

(235) Guye, P.; Ebrahimkhani, M. R.; Kipniss, N.; Velazquez, J. J.; Schoenfeld, E.; Kiani, S.; Griffith, L. G.; Weiss, R. Genetically Engineering Self-Organization of Human Pluripotent Stem Cells into a Liver Bud-like Tissue Using Gata6. *Nat. Commun.* **2016**, *7* (1), 1–12.

(236) Garreta, E.; Prado, P.; Tarantino, C.; Oria, R.; Fanlo, L.; Martí, E.; Zalvidea, D.; Trepat, X.; Roca-Cusachs, P.; Gavaldà-Navarro, A.; et al. Fine Tuning the Extracellular Environment Accelerates the Derivation of Kidney Organoids from Human Pluripotent Stem Cells. *Nat. Mater.* **2019**, *18* (4), 397–405.

(237) Homan, K. A.; Gupta, N.; Kroll, K. T.; Kolesky, D. B.; Skylar-Scott, M.; Miyoshi, T.; Mau, D.; Valerius, M. T.; Ferrante, T.; Bonventre, J. V.; et al. Flow-Enhanced Vascularization and Maturation of Kidney Organoids in Vitro. *Nat. Methods* **2019**, *16* (3), 255–262.

(238) Low, J. H.; Li, P.; Chew, E. G. Y.; Zhou, B.; Suzuki, K.; Zhang, T.; Lian, M. M.; Liu, M.; Aizawa, E.; Rodriguez Esteban, C.; et al. Generation of Human PSC-Derived Kidney Organoids with Patterned Nephron Segments and a De Novo Vascular Network. *Cell Stem Cell* **2019**, *25* (3), 373-387.e9.

(239) Phipson, B.; Er, P. X.; Combes, A. N.; Forbes, T. A.; Howden, S. E.; Zappia, L.; Yen, H. J.; Lawlor, K. T.; Hale, L. J.; Sun, J.; et al. Evaluation of Variability in Human Kidney Organoids. *Nat. Methods* **2019**, *16* (1), 79–87.

(240) Takasato, M.; Er, P. X.; Chiu, H. S.; Maier, B.; Baillie, G. J.; Ferguson, C.; Parton, R. G.; Wolvetang, E. J.; Roost, M. S.; De Sousa Lopes, S. M. C.; et al. Kidney Organoids from Human IPS Cells Contain Multiple Lineages and Model Human Nephrogenesis. *Nature* **2015**, *526* (7574), 564–568.

(241) Elcheva, I.; Brok-Volchanskaya, V.; Kumar, A.; Liu, P.; Lee, J. H.; Tong, L.; Vodyanik, M.; Swanson, S.; Stewart, R.; Kyba, M.; et al. Direct Induction of Haematoendothelial Programs in Human Pluripotent Stem Cells by Transcriptional Regulators. *Nat. Commun.* **2014**, *5* (1), 1–11.

(242) Tsang, K. M.; Hyun, J. S.; Cheng, K. T.; Vargas, M.; Mehta, D.; Ushio-Fukai, M.; Zou, L.; Pajcini, K. V.; Rehman, J.; Malik, A. B. Embryonic Stem Cell Differentiation to Functional Arterial Endothelial Cells through Sequential Activation of ETV2 and NOTCH1 Signaling by HIF1α. *Stem Cell Reports* **2017**, *9* (3), 796–806.

(243) Augustin, H. G.; Koh, G. Y. Organotypic Vasculature: From Descriptive Heterogeneity to Functional Pathophysiology. *Science*. American Association for the Advancement of Science August 25, 2017.

(244) Ferland-McCollough, D.; Slater, S.; Richard, J.; Reni, C.; Mangialardi, G. Pericytes, an Overlooked Player in Vascular Pathobiology. *Pharmacology and Therapeutics*. Elsevier Inc. March 1, 2017, pp 30–42.

(245) Wimmer, R. A.; Leopoldi, A.; Aichinger, M.; Wick, N.; Hantusch, B.; Novatchkova, M.; Taubenschmid, J.; Hämmerle, M.; Esk, C.; Bagley, J. A.; et al. Human Blood Vessel Organoids as a Model of Diabetic Vasculopathy. *Nature* **2019**, *565* (7740), 505–510.

(246) Liu, X. S.; Wu, H.; Ji, X.; Stelzer, Y.; Wu, X.; Czauderna, S.; Shu, J.; Dadon, D.; Young, R. A.; Jaenisch, R. Editing DNA Methylation in the Mammalian Genome. *Cell* **2016**, *167* (1), 233-247.e17.

(247) Victor, M. B.; Richner, M.; Hermanstyne, T. O.; Ransdell, J. L.; Sobieski, C.; Deng, P. Y.; Klyachko, V. A.; Nerbonne, J. M.; Yoo, A. S. Generation of Human Striatal Neurons by MicroRNA-Dependent Direct Conversion of Fibroblasts. *Neuron* **2014**, *84* (2), 311–323.

(248) Wang, W.; Xue, Y.; Zhou, S.; Kuo, A.; Cairns, B. R.; Crabtree, G. R. Diversity and Specialization of Mammalian SWI/SNF Complexes. *Genes Dev.* **1996**, *10* (17), 2117–2130.

(249) Yusa, K.; Zhou, L.; Li, M. A.; Bradley, A.; Craig, N. L. A Hyperactive PiggyBac Transposase for Mammalian Applications. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (4), 1531–1536.

(250) Albini, S.; Coutinho, P.; Malecova, B.; Giordani, L.; Savchenko, A.; Forcales, S. V.; Puri, P. L. Epigenetic Reprogramming of Human Embryonic Stem Cells into Skeletal Muscle Cells and Generation of Contractile Myospheres. *Cell Rep.* **2013**, *3* (3), 661–670.

(251) 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* (1), 1–12.

(252) Han, X.; Wang, R.; Zhou, Y.; Fei, L.; Sun, H.; Lai, S.; Saadatpour, A.; Zhou, Z.; Chen, H.;

Ye, F.; et al. Mapping the Mouse Cell Atlas by Microwell-Seq. *Cell* **2018**, *172* (5), 1091-1107.e17.

(253) Alexa, A.; Rahnenfuhrer, J.; Lengauer, T. Improved Scoring of Functional Groups from Gene Expression Data by Decorrelating GO Graph Structure. *Bioinformatics* **2006**, *22* (13), 1600–1607.

(254) Buchman, V. L.; Davies, A. M. *Different Neurotrophins Are Expressed and Act in a Developmental Sequence to Promote the Survival of Embryonic Sensory Neurons*; The Company of Biologists, 1993; Vol. 118.

(255) Kirschenbaum, B.; Goldman, S. A. Brain-Derived Neurotrophic Factor Promotes the Survival of Neurons Arising from the Adult Rat Forebrain Subependymal Zone. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92* (1), 210–214.

(256) Kumar, A.; D'Souza, S. S.; Moskvin, O. V.; Toh, H.; Wang, B.; Zhang, J.; Swanson, S.; Guo, L. W.; Thomson, J. A.; Slukvin, I. I. Specification and Diversification of Pericytes and Smooth Muscle Cells from Mesenchymoangioblasts. *Cell Rep.* **2017**, *19* (9), 1902–1916.

(257) Gorski, D. H.; Walsh, K. Control of Vascular Cell Differentiation by Homeobox Transcription Factors. *Circulation Research*. Lippincott Williams and Wilkins January 19, 2001, pp 7–8.

(258) Eklund, L. Lack of Type XV Collagen Causes a Skeletal Myopathy and Cardiovascular Defects in Mice. *Proc. Natl. Acad. Sci.* **2001**, *98* (3), 1194–1199.

(259) Ito, A.; Yamamoto, Y.; Sato, M.; Ikeda, K.; Yamamoto, M.; Fujita, H.; Nagamori, E.; Kawabe, Y.; Kamihira, M.; Levenberg, S.; et al. Induction of Functional Tissue-Engineered Skeletal Muscle Constructs by Defined Electrical Stimulation. *Sci. Rep.* **2014**, *4*, 879–884.

(260) Khodabukus, A.; Madden, L.; Prabhu, N. K.; Koves, T. R.; Jackman, C. P.; Muoio, D. M.; Bursac, N. Electrical Stimulation Increases Hypertrophy and Metabolic Flux in Tissue-Engineered Human Skeletal Muscle. *Biomaterials* **2019**, *198*, 259–269.

(261) Qian, H.; Kang, X.; Hu, J.; Zhang, D.; Liang, Z.; Meng, F.; Zhang, X.; Xue, Y.; Maimon, R.; Dowdy, S. F.; et al. Reversing a Model of Parkinson's Disease with in Situ Converted Nigral Neurons. *Nature* **2020**, *582* (7813), 550–556.

(262) 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* (10), 963–966.

(263) Polstein, L. R.; Gersbach, C. a. A Light-Inducible CRISPR-Cas9 System for Control of Endogenous Gene Activation. *Nat. Chem. Biol.* **2015**, *11* (3), 198–200.

(264) Toda, S.; Blauch, L. R.; Tang, S. K. Y.; Morsut, L.; Lim, W. A. Programming Self-Organizing Multicellular Structures with Synthetic Cell-Cell Signaling. *Science (80-.).* **2018**, *361* (6398), 156–162.

(265) Liao, L. De; Lin, C. T.; McDowell, K.; Wickenden, A. E.; Gramann, K.; Jung, T. P.; Ko, L.

W.; Chang, J. Y. Biosensor Technologies for Augmented Brain-Computer Interfaces in the next Decades. *Proc. IEEE* **2012**, *100* (SPL CONTENT), 1553–1566.

(266) He, B.; Cohen, R. J. Body Surface Laplacian ECG Mapping. *IEEE Trans. Biomed. Eng.* **1992**, *39* (11), 1179–1191.

(267) Fukuoka, Y.; Miyazawa, K.; Mori, H.; Miyagi, M.; Nishida, M.; Horiuchi, Y.; Ichikawa, A.; Hoshino, H.; Noshiro, M.; Ueno, A. Development of a Compact Wireless Laplacian Electrode Module for Electromyograms and Its Human Interface Applications. *Sensors* **2013**, *13* (2), 2368–2383.

(268) Besio, W.; Chen, T. Tripolar Laplacian Electrocardiogram and Moment of Activation Isochronal Mapping. *Physiol. Meas.* **2007**, *28* (5), 515–529.

(269) Prats-Boluda, G.; Gil-Sanchez, L.; Ye-Lin, Y.; Ibafiez, J.; Garcia-Casado, J.; Garcia-Breijo, E. Flexible Concentric Ring Electrode for Non Invasive Bioelectrical Surface Recordings. *Procedia Eng.* **2012**, *47*, 1223–1226.

(270) Street, R. A.; Ng, T. N.; Schwartz, D. E.; Whiting, G. L.; Lu, J. P.; Bringans, R. D.; Veres, J. From Printed Transistors to Printed Smart Systems. *Proc. IEEE* **2015**, *103* (4), 607–616.

(271) Swisher, S. L.; Lin, M. C.; Liao, A.; Leeflang, E. J.; Khan, Y.; Pavinatto, F. J.; Mann, K.; Naujokas, A.; Young, D.; Roy, S.; et al. Impedance Sensing Device Enables Early Detection of Pressure Ulcers in Vivo. *Nat. Commun.* **2015**, *6*, 6575.

(272) Kim, J.; Kim, W. S. A Paired Stretchable Printed Sensor System for Ambulatory Blood Pressure Monitoring. *Sensors Actuators, A Phys.* **2016**, *238*, 329–336.

(273) Ng, T. N.; Schwartz, D. E.; Mei, P.; Krusor, B.; Kor, S.; Veres, J.; Broms, P.; Eriksson, T.; Wang, Y.; Hagel, O.; et al. Printed Dose-Recording Tag Based on Organic Complementary Circuits and Ferroelectric Memories. *Sci. Rep.* **2015**, *5*, 13457.

(274) Ng, T. N.; Schwartz, D. E.; Mei, P.; Kor, S.; Veres, J.; Bröms, P.; Karlsson, C. Pulsed Voltage Multiplier Based on Printed Organic Devices. *Flex. Print. Electron.* **2016**, *1* (1), 015002.

(275) Hammock, M. L.; Knopfmacher, O.; Ng, T. N.; Tok, J. B.-H.; Bao, Z. Electronic Readout Enzyme-Linked Immunosorbent Assay with Organic Field-Effect Transistors as a Preeclampsia Prognostic. *Adv. Mater.* **2014**, *26* (35), 6138–6144.

(276) Leleux, P.; Badier, J. M.; Rivnay, J.; Binar, C.; Herve, T.; Chauvel, P.; Malliaras, G. G. Conducting Polymer Electrodes for Electroencephalography. *Adv. Healthc. Mater.* **2014**, *3* (4), 490–493.

(277) Abidian, M. R.; Martin, D. C. Experimental and Theoretical Characterization of Implantable Neural Microelectrodes Modified with Conducting Polymer Nanotubes. *Biomaterials* **2008**, *29* (9), 1273–1283.

(278) Martin, D. C. Molecular Design, Synthesis, and Characterization of Conjugated Polymers for Interfacing Electronic Biomedical Devices with Living Tissue. *MRS Commun.* **2015**, *5*

(02), 131–153.

(279) Bihar, E.; Roberts, T.; Saadaoui, M.; Herve, T.; De Graaf, J. B.; Malliaras, G. G. Inkjet-Printed PEDOT:PSS Electrodes on Paper for Electrocardiography. *Adv. Healthc. Mater.* **2017**, 1601167.

(280) Rivnay, J.; Inal, S.; Collins, B. A.; Sessolo, M.; Stavrinidou, E.; Strakosas, X.; Tassone, C.; Delongchamp, D. M.; Malliaras, G. G. Structural Control of Mixed Ionic and Electronic Transport in Conducting Polymers. *Nat. Commun.* **2016**, *7*, 11287.

(281) Ramuz, M.; Hama, A.; Huerta, M.; Rivnay, J.; Leleux, P.; Owens, R. M. Combined Optical and Electronic Sensing of Epithelial Cells Using Planar Organic Transistors. *Adv. Mater.* **2014**, *26* (41), 7083–7090.

(282) Jeong, S. H.; Zhang, S.; Hjort, K.; Hilborn, J.; Wu, Z. PDMS-Based Elastomer Tuned Soft, Stretchable, and Sticky for Epidermal Electronics. *Adv. Mater.* **2016**, *28* (28), 5830–5836.

(283) Yin, X.; Mead, B. E.; Safaee, H.; Langer, R.; Karp, J. M.; Levy, O.; Atala, A.; Yoo, J. J.; Barker, N.; Es, J. H. van; et al. Engineering Stem Cell Organoids. *Cell Stem Cell* **2016**, *18* (1), 25–38.

(284) Hofer, M.; Lutolf, M. P. Engineering Organoids. *Nature Reviews Materials*. Nature Research February 19, 2021, pp 402–420.

(285) Brassard, J. A.; Lutolf, M. P. Engineering Stem Cell Self-Organization to Build Better Organoids. *Cell Stem Cell*. Cell Press June 6, 2019, pp 860–876.

(286) Passaro, A. P.; Stice, S. L. Electrophysiological Analysis of Brain Organoids: Current Approaches and Advancements. *Frontiers in Neuroscience*. Frontiers Media S.A. January 12, 2021, p 1405.

(287) Kiskinis, E.; Kralj, J. M.; Zou, P.; Weinstein, E. N.; Zhang, H.; Tsioras, K.; Wiskow, O.; Ortega, J. A.; Eggan, K.; Cohen, A. E. All-Optical Electrophysiology for High-Throughput Functional Characterization of a Human IPSC-Derived Motor Neuron Model of ALS. *Stem Cell Reports* **2018**, *10* (6), 1991–2004.

(288) Hochbaum, D. R.; Zhao, Y.; Farhi, S. L.; Klapoetke, N.; Werley, C. A.; Kapoor, V.; Zou, P.; Kralj, J. M.; MacLaurin, D.; Smedemark-Margulies, N.; et al. All-Optical Electrophysiology in Mammalian Neurons Using Engineered Microbial Rhodopsins. *Nat. Methods* **2014**, *11* (8), 825–833.

(289) Li, Q.; Nan, K.; Le Floch, P.; Lin, Z.; Sheng, H.; Blum, T. S.; Liu, J. Cyborg Organoids: Implantation of Nanoelectronics via Organogenesis for Tissue-Wide Electrophysiology. *Nano Lett.* **2019**, *19* (8), 5781–5789.

(290) Shim, C.; Jo, Y.; Cha, H. K.; Kim, M. K.; Kim, H.; Kook, G.; Kim, K.; Son, G. H.; Lee, H. J. Highly Stretchable Microelectrode Array for Free-Form 3D Neuronal Tissue. In *Proceedings of the IEEE International Conference on Micro Electro Mechanical Systems (MEMS)*; Institute of Electrical and Electronics Engineers Inc., 2020; Vol. 2020-January, pp 380–383.

(291) Borda, E.; Ferlauto, L.; Schleuniger, J.; Mustaccio, A.; Lütolf, F.; Lücke, A.; Fricke, S.; Marjanović, N.; Ghezzi, D. All-Printed Electrocorticography Array for In Vivo Neural Recordings. *Adv. Eng. Mater.* **2020**, *22* (3), 1901403.

(292) Adly, N.; Weidlich, S.; Seyock, S.; Brings, F.; Yakushenko, A.; Offenhäusser, A.; Wolfrum, B. Printed Microelectrode Arrays on Soft Materials: From PDMS to Hydrogels. *npj Flex. Electron.* **2018**, *2* (1), 1–9.

(293) Adly, N.; Weidlich, S.; Seyock, S.; Brings, F.; Yakushenko, A.; Offenhäusser, A.; Wolfrum, B. Printed Microelectrode Arrays on Soft Materials: From PDMS to Hydrogels. *npj Flex. Electron.* **2018**, *2* (1), 15.

(294) Tseng, P.; Pushkarsky, I.; Di Carlo, D. Metallization and Biopatterning on Ultra-Flexible Substrates via Dextran Sacrificial Layers. *PLoS One* **2014**, *9* (8), e106091.

(295) Bandodkar, A. J.; Nuñez-Flores, R.; Jia, W.; Wang, J. All-Printed Stretchable Electrochemical Devices. *Adv. Mater.* **2015**, *27* (19), 3060–3065.

(296) Tseng, P.; Pushkarsky, I.; Di Carlo, D. Metallization and Biopatterning on Ultra-Flexible Substrates via Dextran Sacrificial Layers. *PLoS One* **2014**, *9* (8), e106091.

(297) Makeyev, O.; Besio, W. Improving the Accuracy of Laplacian Estimation with Novel Variable Inter-Ring Distances Concentric Ring Electrodes. *Sensors* **2016**, *16* (6), 858.

(298) Besio, W. G.; Koka, K.; Aakula, R.; Dai, W. Tri-Polar Concentric Ring Electrode Development for Laplacian Electroencephalography. *IEEE Trans. Biomed. Eng.* **2006**, *53* (5), 926–933.

(299) Prats-Boluda, G.; Ye-Lin, Y.; Bueno-Barrachina, J.; Rodriguez de Sanabria, R.; Garcia-Casado, J. Towards the Clinical Use of Concentric Electrodes in ECG Recordings: Influence of Ring Dimensions and Electrode Position. *Meas. Sci. Technol.* **2016**, *27* (2), 025705.

(300) Lee, S. M.; Kim, J. H.; Byeon, H. J.; Choi, Y. Y.; Park, K. S.; Lee, S. A Capacitive, Biocompatible and Adhesive Electrode for Long-Term and Cap-Free Monitoring of EEG Signals. *J. Neural Eng.* **2013**, *10* (3), 036006.

(301) Lin, C. T.; Liao, L. De; Liu, Y. H.; Wang, I. J.; Lin, B. S.; Chang, J. Y. Novel Dry Polymer Foam Electrodes for Long-Term EEG Measurement. *IEEE Trans. Biomed. Eng.* **2011**, *58* (5), 1200–1207.

(302) Norton, J. J. S.; Lee, D. S.; Lee, J. W.; Lee, W.; Kwon, O.; Won, P.; Jung, S.-Y.; Cheng, H.; Jeong, J.-W.; Akce, A.; et al. Soft, Curved Electrode Systems Capable of Integration on the Auricle as a Persistent Brain–Computer Interface. *Proc. Natl. Acad. Sci.* **2015**, *112* (13), 3920–3925.

(303) Huigen, E.; Peper, A.; Grimbergen, C. a. Investigation into the Origin of the Noise of Surface Electrodes. *Med. Biol. Eng. Comput.* **2002**, *40* (3), 332–338.

(304) Rocha, P. R. F.; Schlett, P.; Kintzel, U.; Mailänder, V.; Vandamme, L. K. J.; Zeck, G.; Gomes, H. L.; Biscarini, F.; Leeuw, D. M. de. Electrochemical Noise and Impedance of

Au Electrode / Electrolyte Interfaces Enabling Extracellular Detection of Glioma Cell Populations. *Sci. Rep.* **2016**, *6* (April), 34843.

(305) Rubinson, J.; Kayinamura, Y. Charge Transport in Conducting Polymers: Insights from Impedance Spectroscopy. *Chem. Soc. Rev.* **2009**, *38* (12), 3339–3347.

(306) Orazem, M. E.; Tribollet, B. Electrochemical Impedance Spectroscopy. *Annu. Rev. Anal. Chem.* **2010**, *3*, 207–229.

(307) Bisquert, J.; Garcia-Belmonte, G.; Bueno, P.; Longo, E.; Bulhões, L. O. . Impedance of Constant Phase Element (CPE)-Blocked Diffusion in Film Electrodes. *J. Electroanal. Chem.* **1998**, *452* (2), 229–234.

(308) Kim, D.-H.; Lu, N.; Ma, R.; Kim, Y.-S.; Kim, R.-H.; Wang, S.; Wu, J.; Won, S. M.; Tao, H.; Islam, A.; et al. Epidermal Electronics. *Science (80-.).* **2011**, *333* (September), 838–844.

(309) Tseng, P.; Pushkarsky, I.; Di Carlo, D. Metallization and Biopatterning on Ultra-Flexible Substrates via Dextran Sacrificial Layers. *PLoS One* **2014**, *9* (8), 106091.

(310) Linder, V.; Gates, B. D.; Ryan, D.; Parviz, B. A.; Whitesides, G. M. Water-Soluble Sacrificial Layers for Surface Micromachining. *Small* **2005**, *1* (7), 730–736.

(311) Wang, Z.; Volinsky, A. A.; Gallant, N. D. Crosslinking Effect on Polydimethylsiloxane Elastic Modulus Measured by Custom-Built Compression Instrument. *J. Appl. Polym. Sci.* **2014**, *131* (22).

(312) Nussbaum, E. L.; Houghton, P.; Anthony, J.; Rennie, S.; Shay, B. L.; Hoens, A. M. Neuromuscular Electrical Stimulation for Treatment of Muscle Impairment: Critical Review and Recommendations for Clinical Practice. *Physiotherapy Canada*. University of Toronto Press Inc. 2017, pp 1–76.

(313) Yen, H. C.; Chen, W. S.; Jeng, J. S.; Luh, J. J.; Lee, Y. Y.; Pan, G. S. Standard Early Rehabilitation and Lower Limb Transcutaneous Nerve or Neuromuscular Electrical Stimulation in Acute Stroke Patients: A Randomized Controlled Pilot Study. *Clin. Rehabil.* **2019**, *33* (8), 1344–1354.

(314) Howlett, O. A.; Lannin, N. A.; Ada, L.; Mckinstry, C. Functional Electrical Stimulation Improves Activity after Stroke: A Systematic Review with Meta-Analysis. In *Archives of Physical Medicine and Rehabilitation*; W.B. Saunders, 2015; Vol. 96, pp 934–943.

(315) Vafadar, A. K.; Côté, J. N.; Archambault, P. S. Effectiveness of Functional Electrical Stimulation in Improving Clinical Outcomes in the Upper Arm Following Stroke: A Systematic Review and Meta-Analysis. *BioMed Research International*. Hindawi Limited 2015.

(316) Cuesta-Gómez, A.; Molina-Rueda, F.; Carratala-Tejada, M.; Imatz-Ojanguren, E.; Torricelli, D.; Miangolarra-Page, J. C. The Use of Functional Electrical Stimulation on the Upper Limb and Interscapular Muscles of Patients with Stroke for the Improvement of Reaching Movements: A Feasibility Study. *Front. Neurol.* **2017**, *8* (MAY), 186.

(317) Straudi, S.; Baroni, A.; Mele, S.; Craighero, L.; Manfredini, F.; Lamberti, N.; Maietti, E.;

Basaglia, N. Effects of a Robot-Assisted Arm Training Plus Hand Functional Electrical Stimulation on Recovery After Stroke: A Randomized Clinical Trial. *Arch. Phys. Med. Rehabil.* **2020**, *101* (2), 309–316.

(318) Sharififar, S.; Shuster, J. J.; Bishop, M. D. Adding Electrical Stimulation during Standard Rehabilitation after Stroke to Improve Motor Function. A Systematic Review and Meta-Analysis. *Annals of Physical and Rehabilitation Medicine*. Elsevier Masson SAS September 1, 2018, pp 339–344.

(319) Kapadia, N.; Moineau, B.; Popovic, M. R. Functional Electrical Stimulation Therapy for Retraining Reaching and Grasping After Spinal Cord Injury and Stroke. *Front. Neurosci.* **2020**, *14*, 718.

(320) Lozano, A. M.; Lipsman, N.; Bergman, H.; Brown, P.; Chabardes, S.; Chang, J. W.; Matthews, K.; McIntyre, C. C.; Schlaepfer, T. E.; Schulder, M.; et al. Deep Brain Stimulation: Current Challenges and Future Directions. *Nature Reviews Neurology*. Nature Publishing Group March 1, 2019, pp 148–160.

(321) Ochoa, M.; Rahimi, R.; Ziaie, B. Flexible Sensors for Chronic Wound Management. *IEEE Rev. Biomed. Eng.* **2014**, *7*, 73–86.

(322) Dargaville, T. R.; Farrugia, B. L.; Broadbent, J. A.; Pace, S.; Upton, Z.; Voelcker, N. H. Sensors and Imaging for Wound Healing: A Review. *Biosens. Bioelectron.* **2013**, *41*, 30–42.

(323) Swisher, S. L.; Lin, M. C.; Liao, A.; Leeflang, E. J.; Khan, Y.; Pavinatto, F. J.; Mann, K.; Naujokas, A.; Young, D.; Roy, S.; et al. Impedance Sensing Device Enables Early Detection of Pressure Ulcers in Vivo. *Nat. Commun.* **2015**, *6* (1), 6575.

(324) Zhao, M. Electrical Fields in Wound Healing—An Overriding Signal That Directs Cell Migration. *Semin. Cell Dev. Biol.* **2009**, *20* (6), 674–682.

(325) Nuccitelli, R.; Nuccitelli, P.; Li, C.; Narsing, S.; Pariser, D. M.; Lui, K. The Electric Field near Human Skin Wounds Declines with Age and Provides a Non-Invasive Indicator of Wound Healing. *Wound Repair Regen.* **2012**, *19* (5), 645–655.

(326) Ud-Din, S.; Bayat, A. Electrical Stimulation and Cutaneous Wound Healing: A Review of Clinical Evidence. *Healthc. (Basel, Switzerland)* **2014**, *2* (4), 445–467.

(327) Ud-Din, S.; Sebastian, A.; Giddings, P.; Colthurst, J.; Whiteside, S.; Morris, J.; Nuccitelli, R.; Pullar, C.; Baguneid, M.; Bayat, A. Angiogenesis Is Induced and Wound Size Is Reduced by Electrical Stimulation in an Acute Wound Healing Model in Human Skin. *PLoS One* **2015**, *10* (4), e0124502.

(328) Houghton, P. E. Electrical Stimulation Therapy to Promote Healing of Chronic Wounds: A Review of Reviews. *Chronic Wound Care Manag. Res.* **2017**, *4*, 25–44.

(329) Hunckler, J.; de Mel, A. A Current Affair: Electrotherapy in Wound Healing. *J. Multidiscip. Healthc.* **2017**, *10*, 179–194.

(330) Ashrafi, M.; Alonso-Rasgado, T.; Baguneid, M.; Bayat, A. The Efficacy of Electrical

Stimulation in Experimentally Induced Cutaneous Wounds in Animals. *Vet. Dermatol.* **2016**, *27* (4), 235-e57.

(331) Thakral, G.; Lafontaine, J.; Najafi, B.; Talal, T. K.; Kim, P.; Lavery, L. A. Electrical Stimulation to Accelerate Wound Healing. *Diabet. Foot Ankle* **2013**, *4*, 22081.

(332) Sebastian, A.; Iqbal, S. A.; Colthurst, J.; Volk, S. W.; Bayat, A. Electrical Stimulation Enhances Epidermal Proliferation in Human Cutaneous Wounds by Modulating P53–SIVA1 Interaction. *J. Invest. Dermatol.* **2015**, *135* (4), 1166–1174.

(333) Zhao, M.; Song, B.; Pu, J.; Wada, T.; Reid, B.; Tai, G.; Wang, F.; Guo, A.; Walczysko, P.; Gu, Y.; et al. Electrical Signals Control Wound Healing through Phosphatidylinositol-3-OH Kinase-γ and PTEN. *Nature* **2006**, *442* (7101), 457–460.

(334) Cohen, D. J.; Nelson, W. J.; Maharbiz, M. M. M.; James Nelson, W.; Maharbiz, M. M. M. Galvanotactic Control of Collective Cell Migration in Epithelial Monolayers. *Nat. Mater.* **2014**, *13* (4), 409–417.

(335) Baird, A.; Deng, C.; Eliceiri, M. H.; Haghi, F.; Dang, X.; Coimbra, R.; Costantini, T. W.; Torbett, B. E.; Eliceiri, B. P. Mice Engrafted with Human Hematopoietic Stem Cells Support a Human Myeloid Cell Inflammatory Response in Vivo. *Wound Repair Regen.* **2016**, *24* (6), 1004–1014.

(336) Major, M. R.; Wong, V. W.; Nelson, E. R.; Longaker, M. T.; Gurtner, G. C. The Foreign Body Response: At the Interface of Surgery and Bioengineering. *Plast. Reconstr. Surg.* **2015**, *135* (5), 1489–1498.

(337) Daley, J. M.; Brancato, S. K.; Thomay, A. A.; Reichner, J. S.; Albina, J. E. The Phenotype of Murine Wound Macrophages. *J. Leukoc. Biol.* **2010**, *87* (1), 59–67.

(338) Qin, S.; Dorschner, R. A.; Masini, I.; Lavoie-Gagne, O.; Stahl, P. D.; Costantini, T. W.; Baird, A.; Eliceiri, B. P. TBC1D3 Regulates the Payload and Biological Activity of Extracellular Vesicles That Mediate Tissue Repair. *FASEB J.* **2019**, fj.201802388R.

(339) Diegelmann, R. F.; Lindblad, W. J.; Cohen, I. K. A Subcutaneous Implant for Wound Healing Studies in Humans. *J. Surg. Res.* **1986**, *40* (3), 229–237.

(340) Wang, K.; Parekh, U.; Pailla, T.; Garudadri, H.; Gilja, V.; Ng, T. N. Stretchable Dry Electrodes with Concentric Ring Geometry for Enhancing Spatial Resolution in Electrophysiology. *Adv. Healthc. Mater.* **2017**, *6* (19), 1700552.

(341) Khan, Y.; Pavinatto, F. J.; Lin, M. C.; Liao, A.; Swisher, S. L.; Mann, K.; Subramanian, V.; Maharbiz, M. M.; Arias, A. C. Inkjet-Printed Flexible Gold Electrode Arrays for Bioelectronic Interfaces. *Adv. Funct. Mater.* **2016**, *26* (7), 1004–1013.

(342) Venkatraman, S.; Hendricks, J.; King, Z. A.; Sereno, A. J.; Richardson-Burns, S.; Martin, D.; Carmena, J. M. In Vitro and In Vivo Evaluation of PEDOT Microelectrodes for Neural Stimulation and Recording. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2011**, *19* (3), 307–316.

(343) Ganji, M.; Tanaka, A.; Gilja, V.; Halgren, E.; Dayeh, S. A. Scaling Effects on the

Electrochemical Stimulation Performance of Au, Pt, and PEDOT:PSS Electrocorticography Arrays. *Adv. Funct. Mater.* **2017**, *27* (42), 1703019.

(344) Schulz, K. R.; Danna, E. A.; Krutzik, P. O.; Nolan, G. P. Single-Cell Phospho-Protein Analysis by Flow Cytometry. In *Current Protocols in Immunology*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2007; Vol. 78, pp 8.17.1-8.17.20.

(345) Olsen, J. V; Blagoev, B.; Gnad, F.; Macek, B.; Kumar, C.; Mortensen, P.; Mann, M. Global, in Vivo, and Site-Specific Phosphorylation Dynamics in Signaling Networks. *Cell* **2006**, *127* (3), 635–648.

(346) Rao, K. M. K. MAP Kinase Activation in Macrophages. *J. Leukoc. Biol.* **2001**, *69* (1), 3–10.

(347) Lu, S. X.; Alpdogan, O.; Lin, J.; Balderas, R.; Campos-Gonzalez, R.; Wang, X.; Gao, G.-J.; Suh, D.; King, C.; Chow, M.; et al. STAT-3 and ERK 1/2 Phosphorylation Are Critical for T-Cell Alloactivation and Graft-versus-Host Disease. *Blood* **2008**, *112* (13), 5254–5258.

(348) Shaul, Y. D.; Seger, R. The MEK/ERK Cascade: From Signaling Specificity to Diverse Functions. *Biochim. Biophys. Acta - Mol. Cell Res.* **2007**, *1773* (8), 1213–1226.

(349) Manning, B. D.; Toker, A. AKT/PKB Signaling: Navigating the Network. *Cell* **2017**, *169*, 381–405.

(350) Hawkins, P. T.; Stephens, L. R. PI3K Signalling in Inflammation. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **2015**, *1851* (6), 882–897.

(351) Mendoza, M. C.; Er, E. E.; Blenis, J. The Ras-ERK and PI3K-MTOR Pathways: Cross-Talk and Compensation. *Trends Biochem. Sci.* **2011**, *36* (6), 320–328.

(352) Deskins, D. L.; Ardestani, S.; Young, P. P. The Polyvinyl Alcohol Sponge Model Implantation. *J. Vis. Exp.* **2012**, No. 62.

(353) Merrill, D. R.; Bikson, M.; Jefferys, J. G. R. Electrical Stimulation of Excitable Tissue: Design of Efficacious and Safe Protocols. *J. Neurosci. Methods* **2005**, *141* (2), 171–198.

(354) Minhas, P.; Bansal, V.; Patel, J.; Ho, J. S.; Diaz, J.; Datta, A.; Bikson, M. Electrodes for High-Definition Transcutaneous DC Stimulation for Applications in Drug Delivery and Electrotherapy, Including TDCS. *J. Neurosci. Methods* **2010**, *190* (2), 188–197.

(355) Reid, B.; Zhao, M. The Electrical Response to Injury: Molecular Mechanisms and Wound Healing. *Adv. wound care* **2014**, *3* (2), 184–201.

(356) Schander, A.; Tesmann, T.; Strokov, S.; Stemmann, H.; Kreiter, A. K.; Lang, W. In-Vitro Evaluation of the Long-Term Stability of PEDOT:PSS Coated Microelectrodes for Chronic Recording and Electrical Stimulation of Neurons. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*; IEEE, 2016; pp 6174–6177.

(357) Tee, B. C. K.; Ouyang, J. Soft Electronically Functional Polymeric Composite Materials

for a Flexible and Stretchable Digital Future. *Adv. Mater.* **2018**, *30* (47), 1802560.

(358) Wu, Z.; Yao, W.; London, A. E.; Azoulay, J. D.; Ng, T. N. Temperature-Dependent Detectivity of Near-Infrared Organic Bulk Heterojunction Photodiodes. *ACS Appl. Mater. Interfaces* **2017**, *9*, 1654–1660.

(359) Thomas, M. L. *THE LEUKOCYTE COMMON ANTIGEN FAMILY*; 1989; Vol. 7.

(360) Fukui, T.; Dai, Y.; Iwata, K.; Kamo, H.; Yamanaka, H.; Obata, K.; Kobayashi, K.; Wang, S.; Cui, X.; Yoshiya, S.; et al. Frequency-Dependent ERK Phosphorylation in Spinal Neurons by Electric Stimulation of the Sciatic Nerve and the Role in Electrophysiological Activity. *Mol. Pain* **2007**, *3* (1), 1744-8069-3–18.

(361) Wang, Y.; Rouabhia, M.; Zhang, Z. Pulsed Electrical Stimulation Benefits Wound Healing by Activating Skin Fibroblasts through the TGFβ1/ERK/NF-KB Axis. *Biochim. Biophys. Acta - Gen. Subj.* **2016**, *1860* (7), 1551–1559.

(362) Ng, A. H. M.; Khoshakhlagh, P.; Rojo Arias, J. E.; Pasquini, G.; Wang, K.; Swiersy, A.; Shipman, S. L.; Appleton, E.; Kiaee, K.; Kohman, R. E.; et al. A Comprehensive Library of Human Transcription Factors for Cell Fate Engineering. *Nat. Biotechnol.* **2021**, *39* (4), 510–519.

(363) Rackham, O. J. L.; Firas, J.; Fang, H.; Oates, M. E.; Holmes, M. L.; Knaupp, A. S.; Suzuki, H.; Nefzger, C. M.; Daub, C. O.; Shin, J. W.; et al. A Predictive Computational Framework for Direct Reprogramming between Human Cell Types. *Nat. Genet.* **2016**, *48* (3), 331–335.

(364) Cahan, P.; Li, H.; Morris, S. A.; Lummertz Da Rocha, E.; Daley, G. Q.; Collins, J. J. CellNet: Network Biology Applied to Stem Cell Engineering. *Cell* **2014**, *158* (4), 903–915.

(365) D'Alessio, A. C.; Fan, Z. P.; Wert, K. J.; Baranov, P.; Cohen, M. A.; Saini, J. S.; Cohick, E.; Charniga, C.; Dadon, D.; Hannett, N. M.; et al. A Systematic Approach to Identify Candidate Transcription Factors That Control Cell Identity. *Stem Cell Reports* **2015**, *5* (5), 763–775.

(366) Toda, S.; McKeithan, W. L.; Hakkinen, T. J.; Lopez, P.; Klein, O. D.; Lim, W. A. Engineering Synthetic Morphogen Systems That Can Program Multicellular Patterning. *Science (80-.).* **2020**, *370* (6514), 327–331.

(367) Tybrandt, K.; Khodagholy, D.; Dielacher, B.; Stauffer, F.; Renz, A. F.; Buzsáki, G.; Vörös, J. High-Density Stretchable Electrode Grids for Chronic Neural Recording. *Adv. Mater.* **2018**, *30* (15), 1706520.

(368) Khodagholy, D.; Gelinas, J. N.; Thesen, T.; Doyle, W.; Devinsky, O.; Malliaras, G. G.; Buzsáki, G. NeuroGrid: Recording Action Potentials from the Surface of the Brain. *Nat. Neurosci.* **2015**, *18* (2), 310–315.