Arts & Sciences Electronic Theses and Dissertations

Arts & Sciences

Winter 12-15-2021

Regional Reprogramming and the Small Intestine: Analysis and Modeling of Adaptive Regeneration of the Epithelium

Sarah Elizabeth Waye
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

Division of Biology and Biomedical Sciences
Developmental, Regenerative & Stem Cell Biology

Dissertation Examination Committee:
Samantha A. Morris, Chair
Helen McNeill
Deborah C. Rubin
Laura Graves Schuettpelz
Thorold W. Theunissen

Regional Reprogramming and the Small Intestine: Analysis and Modeling of Adaptive
Regeneration of the Epithelium
by
Sarah Elizabeth Waye

A dissertation presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

December, 2021
St. Louis, Missouri

Table of Contents

List of Figures

Chapter 1:

Chapter 2:

Chapter 3:

Chapter 4:

List of Tables

Chapter 2:

Chapter 3:

Chapter 4:

List of Abbreviations

AU – arbitrary units
AP-1 – activator protein 1
CBC – crypt base columnar cell
cDNA – complementary DNA
cGMP – cyclic guanosine monophosphate
ESC – embryonic stem cell
FISH – fluorescent in-situ hybridization
GFP – green fluorescent protein
GC-C – guanylate cyclase-C
GRN – gene regulatory network
GI – gastrointestinal
GLP-2 – glucagon-like-peptide-2
HIIP – human induced intestinal progenitor
HIOs – human intestinal organoids
hPSC – human pluripotent stem cell
HUVEC – human umbilical vein endothelial cell
IBD – inflammatory bowel disease
iEP – induced endoderm progenitor
iHep – induced hepatocyte
iISC – induced intestinal stem cell
iPSC – induced pluripotent stem cell
ISC – intestinal stem cell
MEF – mouse embryonic fibroblast
MOI – multiplicity of infection
PBS – phosphate-buffered saline
PN – parenteral nutrition
PPAR – peroxisome proliferator-activated receptor
PPRE – PPAR response element
PSCs – pluripotent stem cells
QP – quadratic programming
qPCR – quantitative polymerase chain reaction
RA – retinoic acid
RAR – retinoic acid receptor
RARE – retinoic acid response element

RNA-FISH – RNA fluorescence in situ hybridization
RXR – retinoid X receptor
SBR – small bowel resection
SBS – Short Bowel Syndrome
SCNT – somatic cell nuclear transfer
scRNA-seq – single cell RNA sequencing
SI – small intestine
smFISH – single molecule fluorescent *in situ* hybridization
TA – transit amplifying cell
TESI – tissue engineered small intestine
tSNE – t-distributed stochastic neighbor embedding
UMAP – Uniform Manifold Approximation and Projection
UMI – unique molecular identifier

<u>Acknowledgments</u>

I first must acknowledge how grateful I am to my PI and mentor, Samantha Morris, who accepted me into her lab and made me into the scientist I am today. Sam's door was always open, allowing her to provide me with the support, guidance, and the tough love I required for my projects. From the beginning of my time in the lab, I cherish the training I received from Sam herself at the bench and by the end of my training, I cherish the training I received from Sam on writing, editing, and presenting data. Without Sam, I wouldn't be half the scientist I am today.

In addition, I would like to acknowledge how grateful I am to my committee members: Dr. Laura Schuettpelz, Dr. Deborah Rubin, Dr. Helen McNeill, and Dr. Thorold Theunissen. Their generous donation of time, support, and ideas helped to shape my project into what it has become today. Each time we met, my committee was supportive but tough, asking me real questions and making me think about outcomes. They always provided me with ideas and encouragement to follow the project where it was taking me.

I want to acknowledge two women who provided me with the push I needed to pursue science early in my undergraduate degree—Dr. Heidi Elmendorf and Dr. Maria Donoghue at Georgetown University. Dr. Elmendorf introduced me to research as a career path and instilled in me a love of teaching, while Dr. Donoghue supported my research through providing summer research opportunities and mentorship throughout my degree.

I can't forget to thank the entire Morris Lab, members past and present, who have been a constant source of energy and support in my project. The Morris Lab is full of amazing collaboration, ideas, and talented people. Thanks especially go to Wenjun Kong, Kunal Jindal, and Kenji Kamimoto for their support and assistance with all things computational. I am

In my personal life, I thank the friends and family who supported me in completion of this degree. To my friends, the years of Fantasy Football, Alphabet Dinners, and group Halloween costumes as well as our shared love of science and the lab has made this process so much more enjoyable. To my parents Wendell and Cindy, thank you for never pushing me to do something I didn't want to do but being unerringly supportive of all my decisions. Thanks for being proud of my work even when you didn't know how to describe it to your friends. To my brother, Ryan, thanks for being interested in my work and listening to me talk about it. To Andrew, thanks for always crossing your fingers (and toes) when I needed experiments to turn

out, but still saying you had faith in me. Finally, I have to say thank you to Darwin, my dog. I couldn't have done it without the unconditional love of Dogtor Darwin, Ph.Dog. I've been blessed with an unbelievable support system and can't express enough gratitude.

This work is dedicated in loving memory of David Martin, my grandfather.

Sarah Elizabeth Waye

Washington University in St. Louis

December, 2021

ABSTRACT OF THE DISSERTATION

Regional Reprogramming and the Small Intestine: Analysis and Modeling of Adaptive

Regeneration of the Epithelium

by

Sarah Elizabeth Waye

Doctor of Philosophy in Biology and Biomedical Sciences

Developmental, Regenerative & Stem Cell Biology

Washington University in St. Louis, 2021

Dr. Samantha A. Morris, Chair

The small intestine in homeostasis is capable of regular regeneration, but in cases of massive injury like Short Bowel Syndrome, the innate human response often fails to fully compensate for the loss of nutrient absorptive surface area that accompanies bowel resection. Murine models display an active compensatory reaction deemed "adaptation" in which the surface area of the bowel is increased to accommodate nutrient absorptive needs. This observation has highlighted several gaps in knowledge regarding bowel adaptation. Firstly, what occurs on a molecular level in murine models during adaptation? Secondly, how can the findings in mice be applied to humans in an accurate modeling system?

Using a murine model of bowel resection, single cell RNA sequencing was employed to analyze the molecular changes accompanying adaptation of the intestinal epithelium. We identify a process we deem "regional reprogramming" through which distal enterocytes upregulate genes associated with proximal intestinal nutrient transport during adaptation. *In silico* analysis indicates the importance of region-specific diffusible signals, notably Vitamin A, in driving this process through activation of transcription factors and signaling cascades. We further apply these

findings to build an *in vitro* model of regional reprogramming utilizing chimerization of proximal and distal intestinal organoids. The chimeric organoid model, when analyzed using single cell RNA-seq, indicates proximalization of distal organoid tissue upon exposure to the proximal organoid signaling environment—potentially related to changes in Guanylate Cyclase-C signaling. Finally, we utilize direct reprogramming to investigate the ability potential upstream transcription factor *Gata4* to induce regional reprogramming *in vitro* in organoids. We find that the use of *Gata4* alone is not sufficient to induce regional reprogramming *in vitro*, necessitating further study of the role of transcription factor signaling in regional reprogramming. This work concludes that regional reprogramming occurs during adaptation in mice, introduces a novel preliminary model of regional reprogramming due to diffusible signals—the chimera, and lays the groundwork for use of direct reprogramming to study regional reprogramming related to transcription factor activation *in vitro*.

Chapter 1: Introduction

       Much of translational biology today focuses on the repair and regeneration of tissues in order to address ailments from a cellular level. However, studies directly on human or patient tissue are difficult—not only are there restrictions on tissue procurement, but also an inability to perform *in vivo* studies. Because of this, most translational research takes place in animals and the findings are later applied to humans. Animal research allows for *in vivo* regeneration and repair studies and provides critical data, but comes at a high monetary and time cost in addition to obvious shortcomings regarding the differences in animal and human biology. Modeling diseases and potential treatments *in vitro* provides higher throughput experimentation at a lower cost, but many specialized cells fail to grow in generalized culture conditions *in vitro*. In addition, the complex nature of heterogeneous tissues, such as the small intestine, makes *in vitro* modeling and expanding such cells difficult.[1] In recent years, advances in cell culture techniques and cellular reprogramming have opened doors to *in vitro* modeling applications allowing researchers to better understand disease physiology *ex vivo* for applications in healthcare fields. This dissertation will focus on cellular reprogramming, disease and repair modeling, and applications of single cell RNA-seq (scRNA-seq) in a small intestine-specific manner. Specifically, we will use of scRNA-seq to uncover information in *in vivo* animal models and apply these findings to develop *in vitro* models to aid in discovery and future human studies.

1.1   Review of the Literature

       The following literature review, Subsections (1.1.1-1.1.4), will focus on the four topics relevant to this dissertation: 1) Intestinal damage and regeneration, with a focus on Short Bowel

Syndrome, 2) Intestinal organoids, 3) Cellular reprogramming, and 4) scRNA-seq technology. In these sections, a brief historical perspective of each topic and biological background relevant to the work will be discussed. In addition, important models, methods, and technology relating to each topic will be introduced.

1.1.1 Intestinal Damage and Regeneration: Short Bowel Syndrome, an Example of Disease and Adaptation

The small intestine is the organ responsible for the majority of nutrient absorption in the human body. The epithelium, or absorptive surface, of the small intestine is formed of a series of peaks and valleys called villi and crypts respectively (Figure 1.1). These folds in the intestinal epithelium help to increase the surface area available for nutrient absorption.[2] The epithelium consists of multiple, heterogeneous cell types with various functions. The stem cells (also known as crypt base columnar cells or CBCs), originally observed in 1974, reside in the base of the crypt. Transit amplifying cells (a progenitor-like cell located above the stem cell) differentiate into enterocytes, goblet cells, enteroendocrine cells, tuft cells, and Paneth cells, among a few other rarer types.[3] As most cell types of the intestine differentiate, they migrate up out of the crypt onto the villi. Enterocytes are the major absorptive cell of the intestine, forming the majority (~80%) of the epithelium.[4] Interspersed with enterocytes are the mucus secreting goblet cells and hormone producing enteroendocrine cells. Most distinct among the small intestinal cell types is the Paneth cell, a differentiated cell that migrates downward into crypt. These Paneth cells provide antibacterial products to protect the stem cells but have additionally been shown to help provide and maintain the stem cell niche by producing factors needed to maintain Wnt, EGF, and Notch signaling (Figure 1.1).[2,5–7]
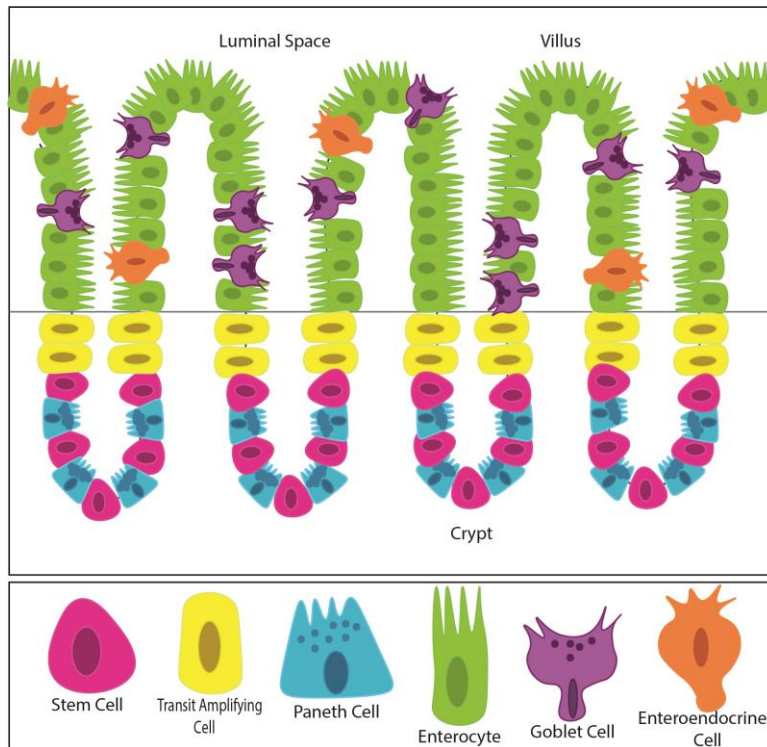
Figure 1.1: Diagram of the small intestinal epithelium. The small intestinal epithelium is lined by multiple cell types. The invaginated crypt area houses stem cells, Paneth cells, and transit amplifying cells, while the protruding villi house enterocytes, goblet cells, and enteroendocrine cells. The grey line delineates the approximate cutoff between villus and crypt.

The lifespan of most intestinal cells is relatively short. After differentiating from stem cells (a constant process), cells move up the villi for 3-5 days until they reach the peak of the villus and are sloughed off and replaced by younger cells.[2,8] Because of this continual movement, the intestinal epithelium is considered to regenerate every 5 days, making it one of the most regenerative tissues in a mammal.[9] The major exception to this rule, again, is the Paneth cell, which has a much longer lifespan—near 1-2 months.[2] Because of this short lifecycle of the epithelium, the intestine is considered to be a highly regenerative organ and minor insults to the epithelium often heal quickly. In addition, studies have shown that damage that eliminates or mutates the Lgr5+ crypt-base stem cells can still be repaired by "reserve" stem cells, cellular competition, and dedifferentiation.[9,10] However, many intestinal diseases and major insults or

damage, such as necrotizing enterocolitis, Crohn's disease, ischemia, volvulus, cancer treatment, and colitis can cause much deeper issues—even complete death of the tissue—that are not simply fixed by the turnover of the epithelium. Because of this, intestinal diseases and regenerative treatments for such diseases comprise a large portion of GI research today.

Of significant interest to physicians and researchers alike is Short Bowel Syndrome (SBS). SBS, while rare (24.5 children per 100,000 live births), has significant morbidity (5-year mortality >30%) and is most common amongst neonatal and pediatric populations.[11,12] SBS occurs when an individual exhibits massive disease or injury to the bowel, due to issues like necrotizing enterocolitis or gastroschisis, necessitating the removal of a large portion of bowel in what is referred to as "small bowel resection" (SBR). The remaining bowel is anastomosed back to a continuous tube. When a large portion of the bowel is removed, the patient suffers from malabsorption, the extent of which is directly related to the amount of bowel removed. Malabsorption occurs in these populations due to the loss of the surface area of the intestine. Less intestinal length means less absorptive epithelium leading to fewer nutrients absorbed. These patients often require total parenteral nutrition in order to survive. Unfortunately, parenteral nutrition incurs a health and financial cost, causing liver damage, sepsis, stunted growth, and costing up to $390,000/year.[13,14] Because of these significant issues, alternative treatments for SBS are highly sought after, but current treatments leave something to be desired. Current treatments are not curative, relying on surgical lengthening of the intestine, or on full intestinal transplantation, which can have dangerous complications such as leakage, bowel dilation, transplant rejection, or donor scarcity, leading to a 3- year survival rate between 32% and 60% [15–17].

Much current research into treatment of SBS focuses on tissue-engineered small intestine (TESI) to fully replace the lost tissue length. TESI encompasses the use of intestinal cells seeded on a scaffold (either decellularized biological matrices or synthetic) mimicking the intestinal architecture.[18–20] Much excitement surrounds the use of human intestinal organoids (HIOs) (derived from human embryonic stem cells, induced pluripotent stem cells, or biopsies) to recellularlize scaffolds for personalized regenerative therapies (organoids will be further discussed in Subsection 1.1.2).[21] While recellularized synthetic scaffolds have shown promise when implanted *in vivo,* seemingly faithfully recapitulating the growth of necessary intestinal cell types and architecture, the difficulties of inducing or providing neural components and vascularization to the implanted TESI cripple the growth of these tissues.[21–24] In addition, the starting cell type for reseeding of the intestine is often pluripotent stem cells (PSCs), however the use of PSCs for transplantation can be dangerous if there are populations of cells that have not fully differentiated, and use of such cells can require immunosuppression.[22] While the future of TESI is promising, fully functional TESI grafts seem in the distant future. For this reason, in this dissertation, we focus on potential methods for inducing and modeling endogenous regeneration capacities of tissue.

Interestingly, in many animal models (mouse, rat, pig, and zebrafish), an *in vivo* response to SBR called "adaptation" can produce enough absorptive epithelium to induce weight gain by normal feeding after resection.[25,26] In this adaptation process it is predominantly noted that the villi get taller and the crypts deeper, a process called "structural adaptation". It is also seen that there is functionally increased nutrient absorption, creating more functional absorptive epithelium to compensate for the lost intestinal sections, with the greatest adaptive response occurring in the ileum.[25,27] Unfortunately, however, this response is not as consistent or

significant in humans, highlighting the need for more knowledge about the molecular underpinnings of adaptation in animals, such that it could potentially be induced or increased in humans.[25,26,28] Because studies in human populations are low-throughput in the SBS field due to sample sizes, we utilize a murine model of resection and adaptation. The murine model is considered to be effective for translational studies due to extensive characterization of intestinal tissue and presence of genetic models.[29] Specifically, this murine model used in our study consists of control sham operations (bowel transection and reanastomosis) or SBR operations (50% resection of proximal small bowel and anastomosis) to induce adaptive responses.[30,31]

Current literature on adaptation after SBR is growing as more studies look into the pathways, hormones, and physiology associated with structural adaptation. Studies in zebrafish have shown that fibroblast growth factor-1 signaling is critical for intestinal adaptation in SBS, Igf1 signaling is important in progenitor cell replication, and EGFR signaling increases during adaptation.[26,32] In mammalian systems such as mouse and rat, many of these signals have also been shown to play a role.[14] IGF-1 treatment increased bowel length and crypt cell proliferation in mice and increased body weight in rats with SBS.[33,34] Rat models have showed that Vitamin A signaling, or retinoic acid (RA), appears to have trophic effects on cell proliferation during adaptation.[35] Adapting mice show a doubling of EGFR expression in crypts, while administration of exogenous EGF during adaptation in mice enhanced many aspects of adaptation.[36,37] Physiologically speaking, recent studies of murine ileostomy have shown the importance of luminal flow in the intestine in the stimulation of cell proliferation during the adaptation period.[29] Of critical importance to our study, it has been shown that HIOs produced from human pluripotent stem cells (hPSCs) and engrafted under the kidney capsule undergo an adaptation-like process (increased villus height, cell proliferation, etc.) when the host animal undergoes

ileocecal resection. These engrafted organoids even retain absorptive functions. Presumably, the engrafted organoids are responding to humoral circulating and diffusible factors in the host after resection, promoting adaptation.[1]

As it stands, current research shows "hyperadaptation" as a promising avenue for treatment of SBS through enhancing endogenous responses of the intestine. Currently, one approved method, administration of glucagon-like-peptide-2 (GLP-2) assists in induction of "hyperadaptation" in human adults, but the long term effects of treatment are unknown.[38] Therefore, many questions regarding enhancement of adaptation remain to be answered. Primarily, what is occurring in the animal models that is driving adaptation on a cell-population level? On a high-throughput molecular level, what changes are occurring that drive the structural and functional changes observed? To resolve these questions, we use scRNA-seq to profile the adapting intestine to gain high-resolution insight into molecular changes occurring in single cells to better understand what triggers effective adaptation in mice, the results of which are found in Chapter 2.

However, animal models are not optimal for many studies due to their long development times, potential inaccuracy in representing human data, and high cost of maintenance. Therefore, we ask, can adaptation be modeled in a way that allows for higher throughput studies and greater expandability at lower cost? Perhaps using *in vitro* organoids from adaptive model species from which results could be applied to human tissue in the future? To address this, in Chapters 3 and 4, we test avenues of reprogramming of mouse intestinal organoids to simulate aspects of adaptation *in vitro*, producing a preliminary model and results that lay the groundwork for future applications in human tissue.

1.1.2 Modeling the Intestine: Organoids Shed 3-Dimensional Light on Intestinal Development and Disease

After our findings in the in vivo SBR model (Chapter 2), we turned to organoids for use in in vitro modeling of SBR induced adaptation.

In the past 50 years, "organoids" have quickly become a popular model for development and disease. Organoids are defined as:

> A 3-D structure derived from either pluripotent stem cells (ESCs or iPSCs), neonatal or adult stem/progenitor cells, in which cells spontaneously self-organise [sic] into properly differentiated functional cell types, and which recapitulates at least some function of the organ…[39]

Organoids are being developed for a range of organs.[40] These models can be used to look at disease states and development with current applications ranging from brain development to the effects of Zika virus to the actual functionality of neuronal networks organoids.[39,41] Even more promising to some is the use of disease organoids, such as those derived from cancerous tissue or cystic fibrosis, in testing therapeutics on human tissue. The heterogeneous nature of the organoids—many expressing multiple cell types representative of the tissue they model—provides a more accurate representation of the *in vivo* environment than traditional 2-dimensional culture of single cell types.[40,42] Despite the development of organoids from a variety of complex tissues, small intestinal organoids are one of the best developed and studied. They generate a structure similar to that found in the small intestine and contain the major differentiated cell types of the small intestine.[7,43]

The development of the intestinal organoid began in the 1980's. Researchers were able to culture rat duodenal tissue into self-enclosed structures with lumens. Research in this vein continued, showing differentiation potential of these *ex vivo* tissues.[6] In 2009, Sato and colleagues created what they deemed "mini-gut organoids."[7] These "mini-guts" rely on the self-

organizational capabilities of intestinal stem cells as well as a multi-factored environment of growth signals to mimic those found in the intestinal epithelium. The "mini-guts," now referred to as small intestinal organoids (SIOs), constituted the first culture system to establish and maintain the budding crypt-villus structure of the intestinal epithelium (Figure 1.2). In order to generate the budding 3-dimensional structure, the SIOs are embedded in Matrigel to support epithelial growth because the high levels of laminin in Matrigel mimic the laminin found at the crypt base. Specific factors are necessary for the growth of the SIOs and a nutrient dense medium containing factors Wnt, R-spondin 1, EGF, and Noggin is used to promote the proliferation of stem cells.[5,44]
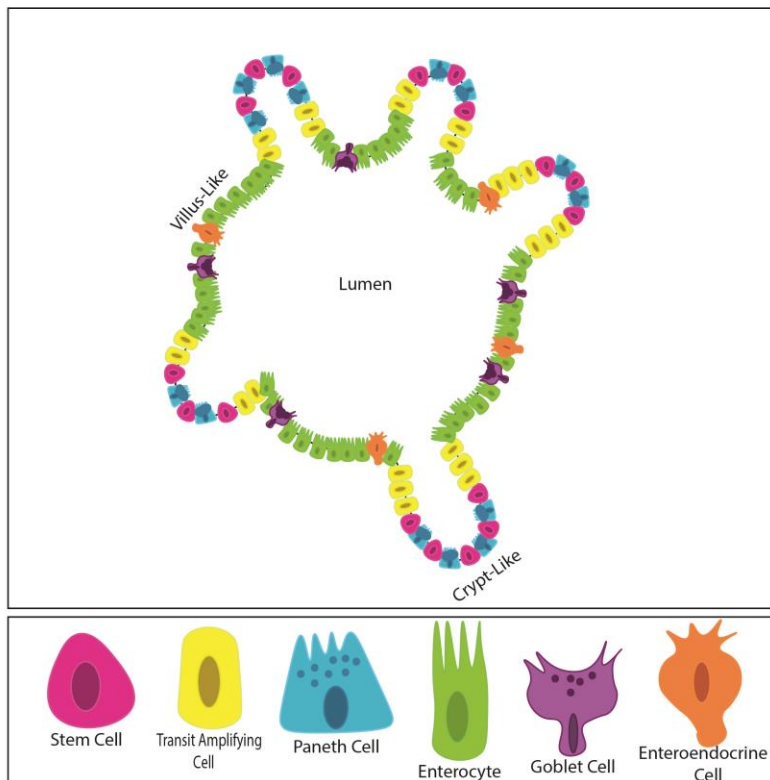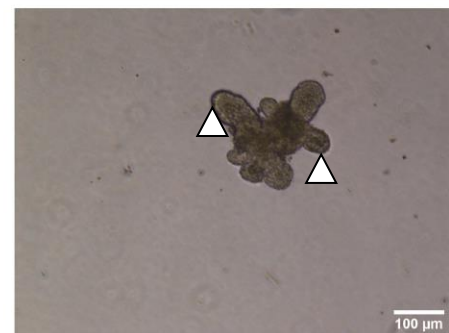
Figure 1.2: Visualization of small intestinal organoids. *(A-B)* Small intestinal organoids are characterized by a layer of epithelial cells surrounding a hollow lumen. Crypt-like protrusions (white arrowheads on *B*) contain stem and Paneth cells while villus-like regions surrounding the lumen house other differentiated cell types. Scale bar = 100 µm. Image taken on Nikon Eclipse Ts2 with Nikon DS-Fi2 camera at with a 10× objective.

SIO lines can be continually cultured with regular passaging for 8 months or more. They are easily generated in two ways: the culture of full crypts from the native intestine or from single Lgr5+ intestinal stem cells.[7,45] In either case, as the intestinal stem cells (ISCs) proliferate, the openings seal off, creating a self-enclosed lumen inside the monolayer of cells. Just as with the *in vivo* epithelium, the cells are continually sloughed off into the lumen. Organoids can also be derived from human pluripotent stem cells (hPSCs), and these additionally include a mesenchymal layer, however this model is immature and resembles fetal tissue unlike the adult tissue modeled by culture of ISCs.[46]

Many questions remain about just how close the organoid is to the tissue it mimics. Does an intestinal organoid express all the signals of the *in vivo* tissue? Does organoid physiology mimic intestinal physiology? In terms of physiology, evidence suggests that organoids reflect aspects of intestinal physiology such as cellular differentiation, region-specific patterning and gene expression, cellular transport abilities, and even interaction with pathogens.[47,48] Evidence from studies of growth of organoids from single stem cells without mesenchymal cells suggests that organoids, with the addition of stromal-derived factors such as Noggin and R-spondin 1 in the medium, recapitulate many aspects of the *in vivo* tissue, namely the stem cell niche.[7,49] Importantly, microarray analysis indicates that the gene expression of organoids remains very similar to that of endogenous crypts, indicating that the extraction from the intestinal epithelium and subsequent *ex vivo* growth change very little about the transcriptome.[7] Staining has shown that the main cell types of the intestine are present in SIOs in comparable locations to the endogenous intestine.[4] Notably, Lgr5+ stem cells touch Paneth cells in the organoid crypt domain, much as they are found endogenously. Paneth cells, therefore, can provide some of the necessary factors for the organoid stem cell niche.[45] The stem cell niche *in vivo* also involves

10

many non-epithelial factors, such as mesenchymal tissue and myofibroblasts.[50] The work of Sato has shown that the intestinal organoids, with the help of the Matrigel matrix and supplemented growth factors, appear to provide a physical stem cell niche without the necessity of mesenchyme.[7] This suggests that the niche can be created *in vitro* without the supporting stroma but with the addition of necessary stromal factors in the medium. In fact, a single Lgr5+ stem cell can differentiate to all the cell types of the organoid without true mesenchyme, only soluble factors.[50] Additionally, they can do it with a uniform solution of growth signals found in the growth medium. Interestingly, modulation of the factors and growth signals in the medium can quickly influence the cellular composition of the SIOs allowing for studies of rare cell types and lineage specification.[4] Ultimately, the 3-dimensional *in vitro* organoid culture system reportedly recapitulates much more of the *in vivo* niche *in vitro* than other *in vitro* culture systems.[46]

Promoting intestinal epithelial regeneration to help improve nutrient absorption and treat a number of intestinal issues through tissue engineering methods is an active area of research.[51] Studies have shown that intestinal organoids (mouse and human) can repopulate crypt or crypt-villus structures in experimentally damaged mouse colon [51–53]. In addition, the use of organoids to repopulate biological and synthetic intestinal scaffolds for transplantation has shown promise[16,54].

Many fascinating studies show the utility of organoids for modeling and studying disease. These cultures have proven valuable in modeling intestinal development and regeneration.[55–57] Because organoids can be derived from a patient, from diseased or healthy tissue, they are being used to model onset of diseases such as cancer of various organs, diarrheal disease, host-microbe interactions, and viral infections.[56,58] Organoids additionally provide the ability to expand patient cells for high-throughput screens for treatment purposes for cancer and other diseases like Cystic

11

Fibrosis.[46] As previously mentioned, in the frame of looking at adaptation after SBR, research has shown that human intestinal organoids implanted under the kidney capsule in mice who have undergone small bowel resection "adapt" with the structures of the organoid extending to increase surface area in response to circulating humoral factors.[1] This is great preliminary evidence for the ability to model response to intestinal damage in organoids, but a major roadblock exists: the *in vivo* environment must be exposing the organoids to unknown humoral cues that we must recapitulate in order to model the response accurately.[46] While organoids can currently be used to study genetics of congenital SBS, methods of regenerating lost cells in SBS, and can be used to elongate intestine using scaffolds, there is a gap of knowledge on how to use organoids to model SBS adaptive responses fully *in vitro*.[59]

Even with all the current research and results from using organoids as disease and development models, many questions still remain. Pertinent to this study, how can we use genetic manipulations, perhaps such as cellular reprogramming (see Subsection 1.1.3), to recapitulate adaptation, a response to damage, that can be studied *in vitro*? How can regional organoids inform us more about the processes occurring during adaptation?

1.1.3 Cellular Reprogramming Presents Promise for *in vitro* Modeling

In Chapter 2, the results of our in vivo scRNA-seq experiments in mice indicate that the distal intestine proximalizes or "regionally reprograms" during the adaptation period. These results directed us to look at utilizing cellular reprogramming methods to model this effect in vitro in organoids in Chapters 3 and 4.

Cellular differentiation is commonly imagined to follow what is called "Waddington's Landscape"—a trajectory through which each cell "rolls" downhill through the landscape during development from a stem cell state to a mature fate, where it is "trapped" in a valley and fate is

fully determined (Figure 1.3).[60] Once a cell reaches its valley, it cannot roll back up the hill and transition to a different valley (or fate) naturally, as its fate is fully determined. However, cellular reprogramming has changed this "black and white" fate determination to a more flexible model in which differentiated cells can be "rolled" back up the hill to less differentiated states or pushed to move to new valleys to different differentiated fates (Figure 1.3).[61]
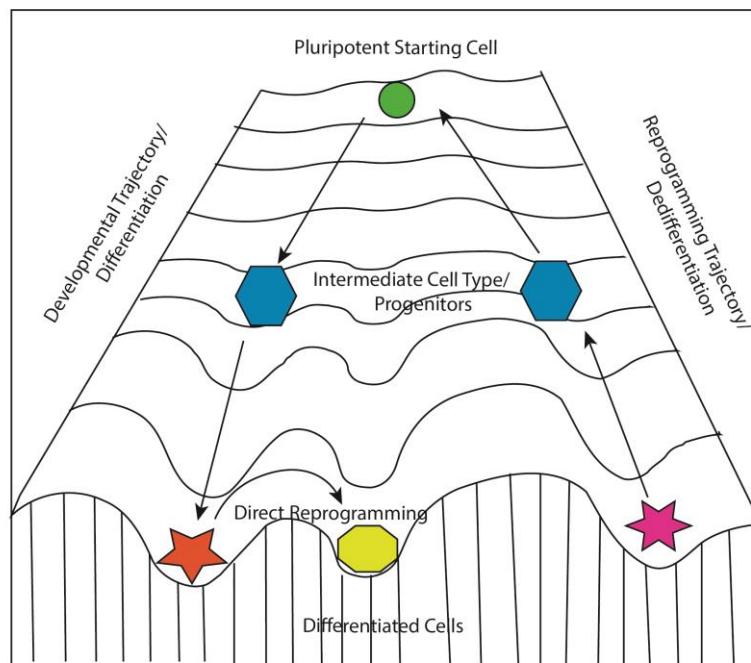
Figure 1.3: Waddington's Landscape. In this model, cells begin at the top of the hill during development and progress downward as they become more differentiated. Using reprogramming methods, cells can move from a fully differentiated state to another identity through direct reprogramming or back toward a less differentiated state through dedifferentiation (adapted from [60]).

Historically, at the forefront of the field of regenerative biology is John Gurdon. In 1958, he successfully cloned amphibians by taking the nucleus of one cell and implanting it into an enucleated egg in a process called Somatic Cell Nuclear Transfer (SCNT). These experiments revealed that a differentiated nucleus could be "reprogrammed" to a totipotent state through exposure to endogenous developmental factors.[62,63] In another historic landmark study, researchers fused muscle and non-muscle cells together to create "heterokaryons." They found

that the non-muscle cells began to activate muscle gene expression, indicating that the expression of genes in differentiated cells is plastic and that cells can be "reprogrammed" with diffusible factors altering cell-cell communication.[64,65] Diffusible factors were also shown to be capable of reprogramming in the "organizer" experiments carried out by Spemann and Mangold where dorsal signals induced ventral cells to take on different fates in an embryo.[66]

The discoveries surrounding diffusible factors, cellular communication, and reprogramming capabilities inspired the development of our novel model, the chimeric intestinal organoid (chimera). Chimeras themselves are a commonly used developmental model defined as:

> A composite organism in which the different cell populations are derived from more than one fertilized egg, thereby combining tissues with distinct genetic origins and identities. The distinct biological mechanisms underpinning chimera formation begin with the persistence of donor cells after transplantation…culminating in donor cell differentiation in a manner paralleling the tissue in which they reside.[67]

Because of the ability of chimeras to combine distinct tissues and promote cellular differentiation via diffusible factors and cellular communication, it is not a far leap to imagine using organoids, which have self-organizational properties and capabilities to recapitulate the niches in the tissue of origin, to create chimeras in an attempt to differentiate stem cells. This idea has been carried out in kidney organoids successfully with mouse embryonic kidney cells capable of forming organoids after aggregation. One such example used human amniotic fluid stem cells combined with murine embryonic kidney cells. Using centrifugation, the human stem cells were aggregated with mouse embryonic kidney cells to create chimeric organoids in which podocytes differentiated from the amniotic fluid stem cells when engrafted *in vivo*.[68,69] Using this inspiration, we developed the chimeric intestinal organoid model by chimerizing proximal and distal small intestinal organoid tissue. As previously mentioned, the ileum is more prone to adaptation, so we investigated the hypothesis that we could model adaptation through distal

14

tissue exposure to and communication with proximal diffusible factors. We asked the question: can we mimic the *in vivo* environment after resection *in vitro* with organoids to model regional reprogramming in a dish to help us learn more about the process of regional reprogramming?

More recently, more controlled methods of reprogramming have been developed. Notably, it was found that *Oct4*, *Sox2*, *Klf4*, and c-*Myc* are sufficient to move somatic cells from differentiated cells to induced pluripotent stem cells (iPSCs).[70] The popular process of directed differentiation induces differentiation from pluripotent cells. In this method, iPSCs or embryonic stem cells (ESCs) are directed to the desired cell type through addition of growth factors and small molecules, such as ESCs to cardiomyocytes with addition of Activin-A and *BMP4*.[71] Limitations of this process include length of time to reprogram and inefficient differentiation into mature cell types. This results in heterogeneous populations of developmentally immature cells.[65,72]

While directed differentiation shows that differentiated cells can be made pluripotent and then directed to differentiate into another cell type, this raised the question of whether differentiated cell types could be directly converted to other differentiated cell types, bypassing the intermediate pluripotent state to produce more mature cell types. Some terminally differentiated cells *in vivo* can change fate without going through a dedifferentiated state through transdifferentiation.[73] This "direct conversion" or "direct reprogramming" as we will refer to it, is possible *in vitro* through the forced expression of ectopic transcription factors or application of developmental signaling factors, which can direct lineage adoption.[66,73,74]

It is reported that use of transcription factors to convert fate is a more robust method and will be utilized in our study.[73] In a series of groundbreaking studies in the 1980's, it was found that ectopic expression of *MYOD1* in fibroblasts produced myoblasts, a direct lineage and fate

change without movement through an intermediate state.[75] These types of direct reprogramming are in use in many labs. A relevant example of the process is the production of iEPs (induced Endoderm Progenitors) utilized in the Morris Lab. This process begins using mouse embryonic fibroblasts (MEFs) and reprograms them through forced overexpression of the transcription factors *Hnf4α* and *Foxa1/2/3*. The resulting cells move from a fibroblast identity to a more open, endoderm progenitor identity—capable of functionally engrafting in the liver and colon.[76–78] Current work on direct reprogramming from the Morris Lab aims to lineage trace reprogramming cells in order to determine which cells will reprogram and also the states of the cells as they move along the reprogramming trajectory.[79] Building on the ability of iEPs to engraft colon, direct reprogramming to intestinal organoids has also been achieved. Using a cocktail of *Hnf4α, Foxa3, Gata6,* and *Cdx2*, it has been shown that mouse fibroblasts can be reprogrammed to intestinal progenitors capable of forming organoids, broadening potential applications of direct reprogramming approaches for use in modeling development and organs.[80] In addition, there are direct reprogramming methods to produce pancreatic β-cells, cardiomyocytes, neurons, and many more cell types.[73]

Neuronal reprogramming is a very active area due to the applications and the many types of neurons that can be and need to be produced. While our study focuses on intestine, neuronal reprogramming can inform on aspects of our goals—primarily, information regarding positional identity. It has been shown that in reprogramming from fibroblast to neuron, the fibroblasts maintain their positional identity markers and *Hox* gene expression but also that *Hox* gene expression can be induced and influenced by signaling molecules, indicating that positional identity is part of cellular memory but can be influenced by signals. [81,82] Both of these aspects are

critical to our goals to reprogram distal tissue to be more proximal to model regional reprogramming.

New avenues of research in direct reprogramming branch out from reprogramming cells in a dish for use in studying reprogramming processes. Now, researchers are looking toward using direct reprogramming of cells to model disease and developmental processes and even for cellular replacement therapies such as in the injured brain.[73,83] A current example is that of direct conversion of somatic cells to neurons (using a variety of transcription factors but typically including *Ascl1, Brn2,* and/or *Myt1l*) to be used to model patient specific diseases which animal models failed to recapitulate.[84] In one study, fibroblasts from patients with amyotrophic lateral sclerosis (ALS) were directly reprogrammed to induced neurons. These neurons expressed mislocalized FUS protein to the cytoplasm and recapitulated known mutant-FUS pathologies, while *in vivo* rat models failed to recapitulate the pathology accurately.[85] In another study, induced neurons modeling pantothenate kinase-associated neurodegeneration (PKAN) were reprogrammed from fibroblasts. Again, patient derived induced neurons showed mitochondrial dysfunction and aberrant oxidation status representative of the disease, unlike animal models, making a "suitable" reprogrammed model for therapeutic testing for the disease.[86]

As previously mentioned, directly reprogrammed cells maintain "memory" of identity, and importantly, they can maintain genetic and epigenetic information from the patient donor.[87] Therefore, we ask the question, can we utilize transcription factor direct reprogramming to reprogram organoids to model regional reprogramming and dissect the roles and effects of specific transcription factors during this process?

1.1.4 scRNA-seq Technology Advances Provide Much Needed Resolution to Transcriptomic Analysis of Modeling and Reprogramming

To analyze subtle molecular changes during adaptation in our in vivo model (Chapter 2) and to classify cell types and identities in our reprogramming efforts (Chapters 3 and 4) we turned to single cell RNA sequencing.

Advances in technology are the foundation of scientific advances, and advances in the resolution of transcriptomics—driven by progress in the field of single cell RNA sequencing (scRNA-seq)—are allowing researchers to truly understand what molecular changes occur on a cellular level like never before.

Historically, the study of transcriptomics on a single cell level began with single-cell qPCR and single-molecule fluorescence *in situ* hybridization (smFISH) in the 1990's.[88,89] Soon after, microarrays were used for the first whole-transcriptome single cell analyses.[90] In a final push, bulk RNA-seq allowed the measurement of population level transcriptomes, pushing forward the knowledge of expression patterns and identity in cells and tissues.[91] However, RNA-seq is performed on bulk populations in which the expression data for many cells in a "homogeneous" population is averaged. These populations of cells can be truly heterogeneous at the RNA expression level as well as in identity, microenvironment, and function. Because of this intrinsic heterogeneity, it is easy to lose rare cell types in the large population, causing the truly distinct transcriptomes of some cells to be blurred and lost.[92] scRNA-seq follows directly on the heels of RNA sequencing.[93] scRNA-seq avoids the losses of RNA seq due to averaging of heterogeneity. It allows us to see the expression pattern of each individual cell, instead of an averaged, bulk reading. This allows for very rare cells to be detected in the population, helping to further define the heterogeneity and identity of each cell within a population.[92,94,95] Additional

biology accessible through scRNA-seq includes high resolution tissue composition, transcriptional dynamics, and even gene regulatory relationships.[96,97]

While there are multiple methods of scRNA-seq following a similar basic framework, we now use the 10x Genomics Chromium 3' platform (10x) for the analysis of thousands of single cells. 10x is a method using aqueous droplets to capture the transcripts from many individual cells for parallel sequencing and analysis. A microfluidic device is utilized to combine an aqueous single-cell suspension, lysis buffer, microparticle beads, and oil together to encapsulate a single cell with a single microparticle bead. When a droplet is formed that contains a cell and bead, the lysis buffer lyses the cell, and the mRNAs attach to oligonucleotides on the bead due to presence of poly-A tails. mRNAs are then reverse-transcribed, PCR amplified, and sequenced. The resulting dataset allows for tracing of each transcript to a specific cell due to the presence of a unique barcode on each bead and additionally gives gene expression levels for individual transcripts in each cell due to presence of unique molecular identifiers (UMIs) (process overview in Figure 1.4).[94,97–99]

The datasets generated through 10x can be analyzed in many ways, our work predominantly uses the R package Seurat and additional in-house tools.[41,100–102] Using Seurat, data can be put through quality control, removing cells with aberrant barcodes and low-quality reads and regressing out factors like mitochondrial genes. Gene expression is quantified and normalized, and then analyzed by dimensional reduction and clustering analysis using differential gene expression data.[94] Such dimensional analysis, now visualized using UMAP, allows users to visualize the transcriptional heterogeneity of populations in two dimensions where cells of similar transcriptional expression cluster together quickly and with high
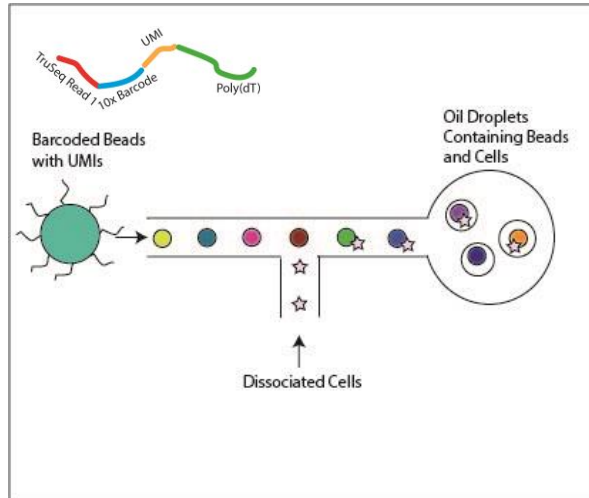
Figure 1.4: Overview of scRNA-seq microfluidic pipeline. A microfluidic device is utilized to combine an aqueous single-cell suspension (bottom, stars), lysis buffer, microparticle beads including poly(dT) primers, barcodes, and UMIs (left), and oil (right) together to encapsulate a single cell with a single microparticle bead. The design of the oligonucleotide on the beads is pictured above the microparticle bead (adapted from [98]).

reproducibility.[103] These clusters can then be used to identify cell types based on expression of canonical or novel markers or comparison to reference atlases.[96,97,100,104]

As previously mentioned, scRNA-seq is superior to bulk sequencing in terms of the resolution it allows for determining heterogeneity in populations—allowing the tracing of lineages, reprogramming processes, development, and even identification of heterogeneity in models like organoids and *in vivo* tissues. For example, our lab utilized lentiviral barcodes called CellTags to map lineage and identity of iEPs during the direct reprogramming process. The resolution provided by scRNA-seq allowed for the CellTags to be utilized to construct lineage trees, leading to the discovery of two trajectories in iEP reprogramming: successful reprogramming and the 'dead-end' state. Using the scRNA-seq analysis, *Mettl7a1* was identified as a novel factor promoting efficient successful reprogramming. This study successfully uses scRNA-seq and novel technologies to look into direct reprogramming dynamics as never seen in the iEP system and provides additional information on how to improve the efficiency of

reprogramming.[79] Additionally, scRNA-seq has been utilized to investigate the process of gastrulation in mice, looking deep into developmental stages in which the molecular mechanisms driving the process were once considered an unknown "black box". The use of scRNA-seq on 1,205 cells from various stages of gastrulation allowed the researchers to resolve the transcriptional programs driving gastrulation, finding new markers and regulators of gastrulation such as *Slc35d3,* as well as a role of *Tal1* in mesodermal blood formation.[105] Heterogeneity of the mouse small intestine was explored using scRNA-seq, even identifying new diversity in enteroendocrine cells and previously unknown gene signatures.[106] Even intestinal organoids have been successfully studied using scRNA-seq. Organoids undergo a symmetry breaking event early in development when a stochastic event produces a Paneth cell, creating the first budding stem cell niche in the organoid. To study the gene variability causing this symmetry breaking event, scRNA-seq in a time course was utilized to investigate variability in *Yap1* expression in single cells during symmetry breaking. They found that variability in expression of *Yap1* induced Paneth cell differentiation at high levels, inducing symmetry breaking critical to the formation of traditional budding organoids.[57]

Based on these studies and others, we ask several questions. Can we use scRNA-seq to learn more about the drivers of the regenerative adaptation state that occurs after SBR? How can we combine scRNA-seq and novel analysis techniques built in our lab to predict reprogramming factors capable of achieving organoid regional reprogramming? Can scRNA-seq help us to accurately analyze the degree of reprogramming achieved by reprogramming factors in organoids?[100,101] Finally, can scRNA-seq data elucidate important changes in cell-cell communication occurring during regional reprogramming?[107]

1.2 Key Outstanding Questions in the Field and Hypotheses

This dissertation will address the following major questions: What are molecular changes associated with adaptation after SBR on a single cell level? How can we apply scRNA-seq technology to determine potential upstream drivers of adaptation? And finally, how can we model and study adaptative responses using combinations of organoids, reprogramming, and scRNA-seq analysis?

Based on these questions and expert knowledge of applicable technologies, we form the following hypotheses:

1. scRNA-seq technology can provide high resolution data elucidating molecular changes occurring during the adaptation process in mouse models of adaptation.

2. Applications of analysis of scRNA-seq data will identify potential upstream regulators of adaptation.

3. Reprogramming of organoid models using potential regulators of adaptation will build a collection of models of regional reprogramming that can be used to model and test adaptation in human populations.

To answer these questions and study these hypotheses, a number of models and technologies will be utilized. Major models will be the *in vivo* mouse model of SBR discussed in Subsection 1.1.1 and mouse small intestinal organoids as discussed in Subsection 1.1.2. Major methods and technologies will include: chimerization of proximal-distal organoids and transcription factor-driven direct reprogramming of organoids as discussed in Subsection 1.1.3 and scRNA-seq of various experimental models as discussed in Subsection 1.1.4.

1.3   Contributions of Dissertation to the Field

The contributions of this dissertation to the field are both on a translational and basic science level. Firstly, the major dataset utilized in Chapter 2 is available to the public, allowing other researchers to analyze and mine data regarding SBR and adaptation without necessitating involvement of additional experiments. In addition, this dissertation provides SBS researchers with data about a previously undescribed process during adaptation after SBR, regional reprogramming, or the proximalization of enterocytes following resection during adaptation. It suggests the high importance of exposure to region-specific diffusible signals, such as Vitamin A, and changes in signaling cascades post-resection. This dissertation provides valuable information regarding the use of organoids to model and study adaptation for applications to personalized medicine in the future, specifically focusing on the role of proximal diffusible factors in regional reprogramming through building of the chimeric organoid model. Finally, this dissertation lays the groundwork to pursue further modeling of regional reprogramming through direct reprogramming of organoids to dissect the effects of upstream transcription factor signaling.

Chapter 2: In Vivo Regional Reprogramming of the Mouse Intestine after Small Bowel Resection

Adapted from:

Single-Cell Analysis Reveals Regional Reprogramming During Adaptation to Massive

Small Bowel Resection in Mice

Seiler, K.M.[*], Waye, S.E.[*], Kong, W., Kamimoto, K., Bajinting, A., Goo, W.H., Onufer, E.J., Courtney, C., Guo, J., Warner, B.W.[#], Morris, S.A.[#] (2019). Single-Cell Analysis Reveals Regional Reprogramming During Adaptation to Massive Small Bowel Resection in Mice. Cellular and Molecular Gastroenterology and Hepatology. *8*(3): 407-426.

[*]Authors Contributed Equally; [#]Co-Corresponding Authors

Sarah Elizabeth Waye contributed to the original manuscript in the following manner: co-wrote manuscript, data acquisition, data analysis, data interpretation, statistical analysis, figure preparation, critical review of manuscript.

2.1 Abstract

The small intestine (SI) displays known regionality in nutrient absorptive and immunological functions. After loss of intestinal tissue, such as in SBS, the remaining SI compensates, or "adapts" to the shortened intestine and resultant loss of surface area. However, while the morphological results of adaptation are well described, the molecular capacity of SI epithelium to reprogram its regional identity in response to resection has not been described. Here, we apply single-cell resolution analyses to characterize molecular changes underpinning adaptation to SBS. scRNA-seq was performed on epithelial cells isolated from distal SI of mice following 50% proximal small bowel resection (SBR) surgery vs. sham surgery. Single-cell profiles were clustered based on transcriptional similarity. An unsupervised computational approach to score cell identity was used to quantify changes in regional (proximal vs. distal) SI identity and findings were validated using molecular biology approaches. Cell identity scoring demonstrated segregation of enterocytes by regional SI identity: a population of SBR enterocytes assumed more mature proximal identities despite being sampled from the distal intestine. This was associated with significant upregulation of lipid metabolism, which was validated via orthogonal analyses. Observed upstream transcriptional changes suggest retinoid metabolism through exposure to proximal levels of diffusible factors Vitamin A/Retinoic Acid and proximal transcription factor *Creb3l3* drive proximalization of cell identity in response to SBR. Adaptation to proximal SBR involves regional reprogramming of ileal enterocytes toward a proximal identity. Interventions bolstering the endogenous reprogramming capacity of SI enterocytes—conceivably by engaging proximal diffusible factors and upstream transcription factors—merit further investigation, as they may increase enteral feeding tolerance, and obviate intestinal failure, in SBS.

2.2 Introduction

The small intestine absorbs nutrients necessary to sustain life, and displays regional specialization for absorption of specific nutrients from duodenum (proximal) to jejunum to ileum (distal). The majority of nutrient absorption occurs in the duodenum and jejunum, while the ileum is primarily responsible for absorbing bile, vitamin B12, and fat-soluble vitamins.

A variety of diseases require surgical resection of significant lengths of SI. These range from congenital anomalies and necrotizing enterocolitis in children to trauma, embolism, and malignancy in adults. The resulting loss of SI can cause short bowel syndrome—the inability of the SI to completely support the metabolic demands of a patient. Due to loss of SI length, there is a reduction in absorptive surface area for the patient, resulting in an inability to properly absorb the nutrients necessary for survival. Management options for SBS are limited, including parenteral nutrition (PN), intestinal lengthening procedures, and small bowel transplant, all incurring significant morbidity and mortality.[108,109]

The murine model of SBS is based on small bowel resection (SBR), in which 50% of the proximal SI is surgically removed.[31] Sham surgery consists of transection and anastomosis, without removal of any SI, and acts as a control for exposure to anesthesia, laparotomy, and intestinal transection. This model elicits known adaptive responses such as villus lengthening in the remnant ileum of SBR but not sham mice, increasing mucosal absorptive surface area to compensate for lost tissue. Importantly, the degree of this "structural adaptation" response correlates with "functional adaptation," as evidenced by increased oral tolerance and weight gain in mice. This model is relevant to clinical SBS, as structural adaptation correlates with oral tolerance and weaning from PN observed in human patients.[110] At the same time, structural adaptation does not intrinsically predict functional adaptation.[111] This leads us to conclude that

additional factors beyond simple tissue hyperplasia are at play, likely underscored by molecular changes detectable at the single-cell level.

While structural and, to a lesser extent, functional adaptation following SBR, has been characterized, relatively little is understood about the molecular changes that accompany the adaptation process. In this respect, messenger RNA (mRNA) expression analyses offer critical insight, as clinically appreciable adaptation may require that cells assume molecular identities mimicking those of the resected region as compensatory reactions. Studies of adaptation at the structural and functional level lack the resolution to explore this possibility. Here, in the case of proximal SBR, we hypothesize that remnant ileum (distal SI) upregulates gene expression patterns characteristic of the jejunum (proximal SI) at the single-cell level, a process we term "regional reprogramming."

Clinical therapies to induce regional reprogramming could enhance an SBS patient's ability to tolerate oral intake and wean from PN by augmenting the innate regional functionality of the lost epithelial cells. It is possible this approach may actually be more effective than the previous "holy grail" of SBS research, which has primarily focused on inducing structural adaptation. Enhanced structural adaptation through induction of tissue growth is intrinsically more metabolically demanding and may or may not affect the key absorptive, metabolic, and immunological pathways specifically deficient in an SBS patient.

To test our hypothesis and address gaps in our understanding of adaptation to SBS, we employed high-throughput single-cell RNA sequencing to characterize gene expression changes of distal SI epithelium during adaptation following massive proximal SBR. This allowed us to dissect population heterogeneity within the epithelium and characterize the regionalization pathways critical in the adaptive response. Here, we show that following SBR, the enterocytes of

the distal SI epithelium regionally reprogram toward mature proximal enterocyte identity, accompanied by increased proximal SI nutrient processing gene expression. These changes are punctuated by the increased expression of the proximal SI transcription factor, *Creb3l3*, a novel candidate for reprogramming distal SI gene regulatory networks to a more proximal identity. Analysis of upstream pathways suggests a role for Retinoic Acid (RA) signaling or retinoid metabolism in driving the adaptation response. Together, our single-cell analyses have enabled high-resolution characterization of the molecular changes and regional reprogramming that underlie adaptation and provide basis for the studies performed in Chapters 3 and 4.

2.3 Cellular Heterogeneity of the Small Intestinal Epithelium Is Captured by scRNA-seq

Structural adaptation—or villus growth— reliably occurs by day 7 after SBR, making it a commonly utilized experimental endpoint in SBR studies. Therefore, we utilized day 7 after sham or SBR surgery as our experimental endpoint. We confirmed typical SBR structural adaptation, with villi height increasing by 86.19 ± 19.14 μm ($P < 0.01$) relative to sham (Figure 2.1*A*). Epithelial cells from animals demonstrating structural adaptation or from sham surgery animals were harvested from SI in an area distally equidistant from the anastomosis, dissociated into single cells, and processed via high-throughput droplet-based scRNA-seq using the 10x Genomics platform.[99] In total, we sequenced 19,245 cells from 9 independent biological replicates (sham: 8209 cells, n = 5 replicates; SBR: 11036 cells, n = 4 replicates). A mean of 1767 and 1763 genes per cell in sham and SBR, respectively, and 6754 and 6111 transcripts per cell in sham and SBR, respectively, were detected (Figure 2.1*B*).

To reduce dimensionality of the data and visualize the clustering of cells, we used the R package Seurat along with Uniform Manifold Approximation and Projection (UMAP)

visualization.[102,103,112] UMAP analysis and plotting revealed 16 clusters of transcriptionally distinct cell types/states (Figure 2.1*C*). Scoring and projection of cell cycle state onto the UMAP plot revealed clustering of cells in S and G1 phases, corresponding to stem cells and transit-amplifying (TA) cells (Figure 2.1*D*) as identified through classical marker expression (not shown) and in Figure 2.2*A*. Gene expression between equivalent biological replicates was highly correlated, demonstrating a high degree of consistency between the independent biological replicates and reducing the possibility of batch effects between replicates driving results (Figure 2.1*E*). Furthermore, cells from every cluster were represented in each biological replicate, demonstrating consistency of cell capture and thereby our collection methods (Figure 2.1*F*).

To assign cell identity to each cluster in an unsupervised manner, we used a computational method based on quadratic programming (QP) to score individual cell identity against an existing single-cell atlas of well-annotated SI cells.[106,113] This reference atlas contains stem cells, TA cells, early and late enterocyte progenitors, immature proximal and distal enterocytes, mature proximal and distal enterocytes, goblet cells, Paneth cells, enteroendocrine cells, and tuft cells, all annotated based on a list of previously identified and novel markers.[114] We have previously demonstrated the efficacy of QP in placing cells into an identity continuum during lineage reprogramming.[79] The QP algorithm has since been enhanced and named Capybara, it is used in Chapters 3 and 4.[100]

Cell identity scores generated by QP were projected onto the UMAP plot, enabling cell cluster identity to be annotated (Figure 2.2*A*). Projection of all identity scores onto the UMAP demonstrated this clustering and visualization method does indeed retain both local and global information, capturing the differentiation trajectory from stem cells to mature enterocytes (Figure 2.2*B*). Further examination of these clusters revealed that expression of the proliferation
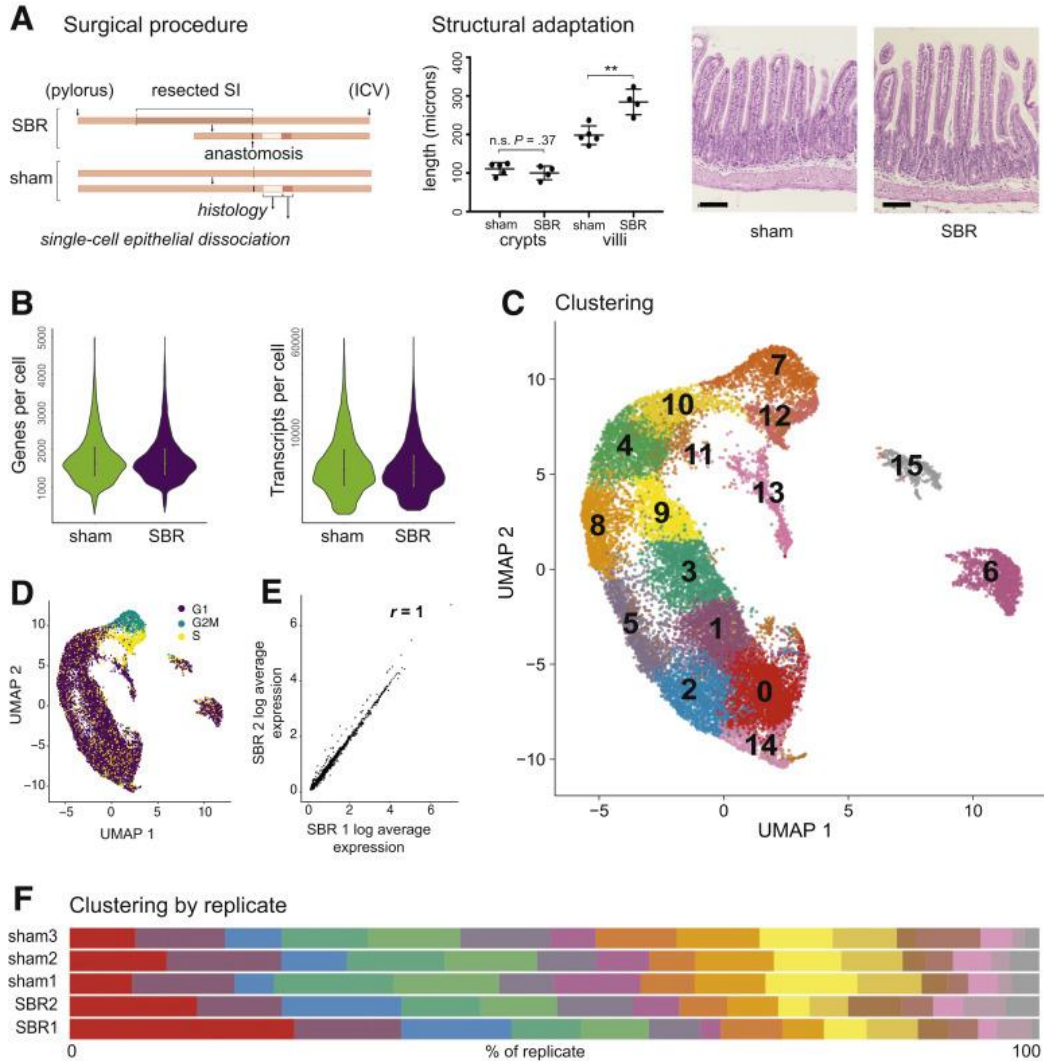
Figure 2.1: Experimental design, quality control, and single-cell analysis. (*A*) A 50% proximal SBR and sham operation were performed on mice. Seven days after surgery, the intestine distal to the anastomosis (ileum) was harvested and equal amounts of tissue equidistant from the anastomosis were used to generate single-cell epithelial suspensions. An area immediately adjacent to this was prepared for histological examination. Typical structural adaptation of SBR mice (lengthened villi) was confirmed ($P = 0.003$), with a representative hematoxylin and eosin image of SI tissue from a sham vs SBR mouse shown (20× image acquired using Nikon Eclipse 80i). Scale bar = 100 μm. Epithelium from mice demonstrating structural adaptation was prepared for scRNA-seq analysis. (*B*) A mean of 1767 and 1763 genes per cell in sham and SBR, respectively, and 6754 and 6111 transcripts per cell in sham and SBR, respectively, were detected. (*C*) UMAP of integrated biological replicates identified 16 unique cell clusters. (*D*) Cell cycle states projected onto the UMAP. (*E*) Representative plot of SBR experimental replicates demonstrated similar gene expression profiles. Correlation coefficients (R) of average gene expression are as shown between these biological replicates. Total biological replicates were 5 sham and 4 SBR (n = 3 "sham1," n = 1 "sham2," n = 1 "sham3," n = 3 "SBR1," n = 1 "SBR2"). (*F*) The same 16 clusters were identified in both sham and SBR, in all replicates, as described in panel *E*. Distribution of all cells across clusters 0–15 (from left to right), by replicate, is shown.

30

marker *Mki67* is enriched in stem, TA, and progenitor cells, and is downregulated as cells begin to differentiate. Conversely, expression of mature enterocyte marker alkaline phosphatase (*Alpi*) increases with differentiation or maturation, with highest expression colocalizing in areas identified by QP as mature enterocytes (Figure 2.2*C*), thereby supporting the QP analysis by identification of cell populations with known marker genes.

In summary, our single-cell analyses identified all major SI epithelial cell populations, demonstrating the strength of our unsupervised QP cell classification approach in concert with single-cell analysis, showing the effectiveness of our collection and analytical methods. Our analysis utilizing UMAP visualization to identify cells differentiating along a defined trajectory reflects the normal cellular differentiation of the rapidly cycling intestinal epithelium and identifies key cell types and transitional states. Together, this provides a comprehensive picture of complex SI epithelial heterogeneity, under both sham and SBR conditions in our model.

2.4 Regional Reprogramming Toward Mature Proximal Enterocyte Identity Occurs After SBR

Changes in cell type composition of the SI epithelium accompany adaptation. However, which cell types expand during adaptation is highly debated in the literature. Some reports describe expansion of enterocytes while others cite expansion of secretory cells.[115–119] It is possible that the differing reports exist due to use of different models, timepoints, and experimental conditions as well as the historic use of limited resolution techniques for cellular identification.

Here, we use our validated, unbiased single-cell resolution classification of cell identity (QP) to precisely quantify changes in epithelial composition following SBR specific to our model. We assessed the distribution of sham- vs SBR-derived cells by projecting the densities of
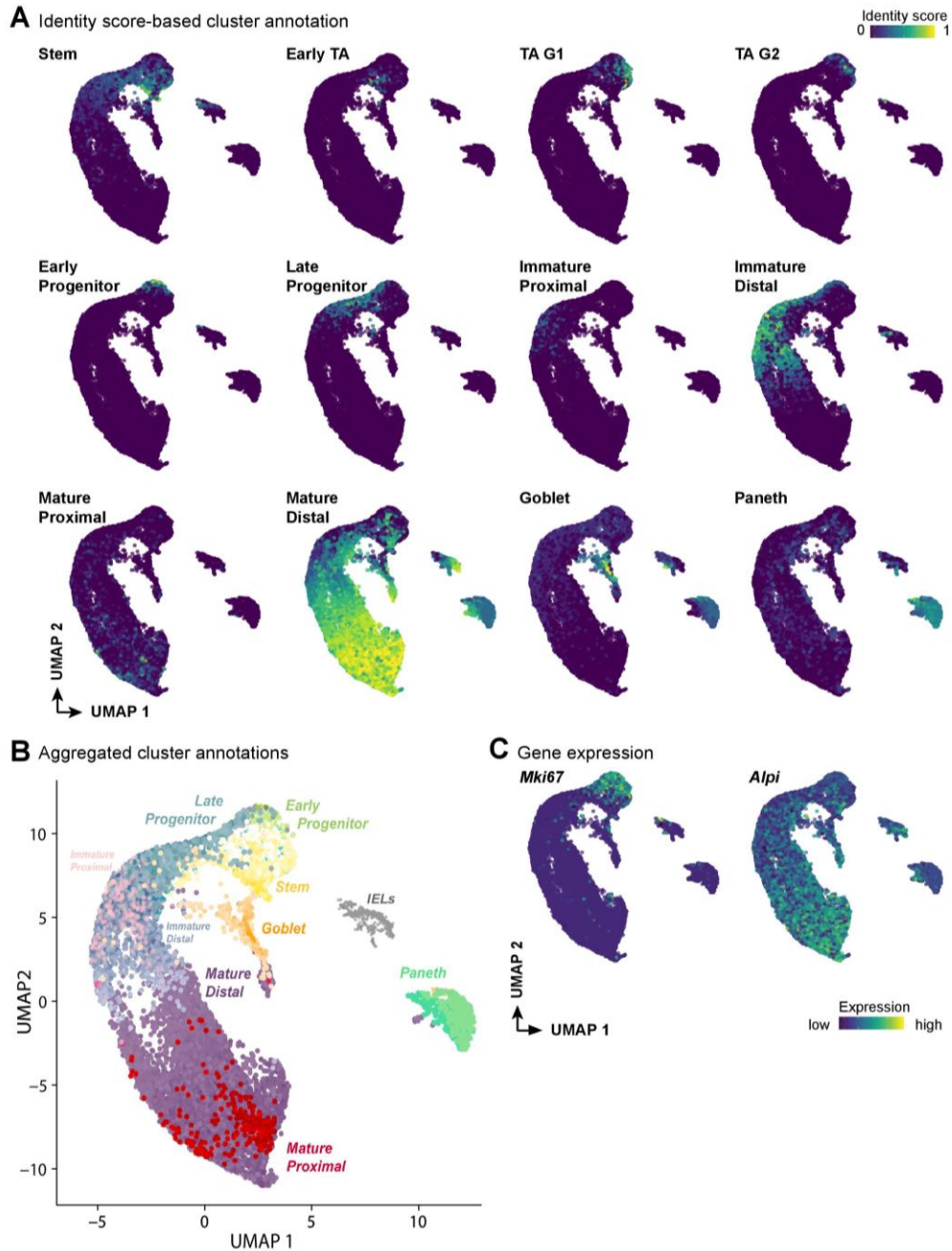
Figure 2.2: Annotation of cell identities using QP. (*A*) QP-based identity scores of intestinal epithelial populations, projected onto UMAP. Cell populations (from left to right) include stem, early TA, TA G1, TA G2, early enterocyte progenitors, late enterocyte progenitors, immature proximal enterocytes, immature distal enterocytes, mature proximal enterocytes, mature distal enterocytes, goblet cells, and Paneth cells. Low percentages of tuft and enteroendocrine cells (0.3% and 0.003%, respectively) were identified and are not shown. (*B*) Aggregated QP scores provide a summary of cell identities within the UMAP, demonstrating a maturation trajectory from stem cells to mature enterocytes. (*C*) Projection of transcript enrichment for selected QP validation markers, from left: proliferation marker antigen KI-67 (*Mki67*) is appropriately enriched in the stem, TA, and progenitor regions of UMAP; alkaline phosphatase (*Alpi*) expression increases as enterocyte maturation occurs. Color scale bar indicates relative intensity of cell identity scoring (*A*) or gene expression (*C*) across the UMAP.

32

each population onto the UMAP (Figure 2.3*A*). Looking at the enterocyte population (comprised of immature and mature enterocytes), we found a significant increase following SBR as a total of all cells surveyed (68.9% ± 3.1% of sham events sampled, vs. 76.8% ± 0.1% of SBR ($P < 0.05$)) (Figure 2.3*B*).
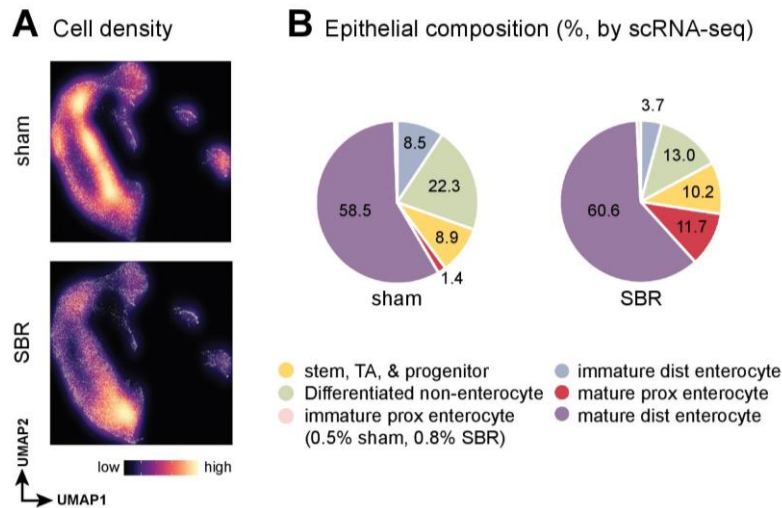
Figure 2.3: Relative expansion of mature proximal enterocytes occurs in SBR mice. (*A*) Density of sampled cells from sham and SBR epithelium, projected onto the UMAP, demonstrates a population shift toward mature proximal enterocytes in SBR. Color scale bar indicates relative density. (*B*) Graphical representation of how epithelial lineages (as identified by QP), contribute to the total composition of events sampled by scRNA-seq in sham vs SBR.

This finding supports reports in the literature of increases in enterocyte proliferation and total number post resection.[116,118,120] It has been reported that after resection, developing enterocytes have accelerated differentiation and maturation in expressing nutrient transporters.[121] Our use of a high-resolution, unbiased classification system based on gene expression may have allowed us to detect and identify earlier stage and immature enterocytes based on their early-stage transcriptional profiles than traditional methods, allowing for the detection of the expansion of full (mature and immature) enterocyte populations while previous studies could not detect such expansion.

Considering our hypothesis that cells adopt a different regional identity to aid adaptation, we quantified the balance between proximal and distal enterocyte identities in sham vs. SBR samples. Our hypothesis was initially supported by the density projections in Figure 2.3*A*, as the highest density of SBR cells fell in a region identified to contain mature proximal enterocytes in Figure 2.2*A*. As expected, considering the tissue was harvested from ileum, a large percentage of cells (58.5% ± 7.5%) from sham samples received high mature distal enterocyte scores with very few cells (1.4% ± 0.6%) scoring as mature proximal enterocytes (Figure 2.3*B*).

In contrast, in SBR samples we found a significant increase in the percentage of cells receiving high mature proximal enterocyte scores (11.7% ± 4.1%, $P < 0.05$) when quantified from QP scores (Figure 2.3*B*). This shift toward mature proximal enterocyte identity in SBR suggests a transcriptional "proximalization," we have termed "regional reprogramming", of ileum in response to proximal SI resection. The term "regional reprogramming" aims to describe the complex array of changes on a molecular, single-cell level related to regional identity detected in our study. While previous studies have looked at changes individual genes associated with intestinal regions, by detecting arrays of genes associated with different regional identities using QP instead of single genes, we gain higher resolution insight into the changing ileal enterocyte identity than previous studies and gain more insight into the regulation of the known transcriptional changes and subsequent regional reprogramming.

2.5 Regional Reprogramming After SBR Increases Proximal SI Nutrient Processing Gene Expression

From our previous analyses, changes in SI epithelial composition following SBR focus primarily on a shift toward mature proximal enterocyte identities, suggesting that these changes

are a key driver of the adaptive response. Thus, we next focused on characterizing the transcriptional changes underlying the shift toward mature proximal enterocyte identity in SBR. It has been reported that the ileum is the site of greatest adaptational changes in the small intestine and previous studies have cited increases in major proximal absorptive genes such as *Fabp1* and *Apoa4* in the ileum post-resection.[27,120,122] Our scRNA-seq analysis supports these previous findings, but additionally, use of scRNA-seq allows us to unbiasedly detect differential gene expression after resection—finding potentially previously novel markers and molecular regulators of proximalization through our expanded dataset.

To identify significant transcriptional changes after SBR, we performed differential gene expression analysis, identifying 174 differentially expressed genes between all sham and SBR cells. The 10 most significantly highly expressed genes in SBR, relative to sham, are shown in Table 2.1. This list of SBR-associated transcripts is enriched for signature genes of proximal SI nutrient processing function, including apolipoprotein A-IV (*Apoa4*), fatty acid binding protein 1 (*Fabp1*), apolipoprotein C-III (*Apoc3*), lactase (*Lct*), and epoxide hydrolase 2 (*Ephx2*) (Figure 2.4).[114,123] In contrast, distal SI transcript fatty acid binding protein 6 (*Fabp6*) was depleted in SBR (0.62 average log fold change depleted, $P < 0.001$).[114]

Table 2.1: Top 10 Genes Upregulated in SBR Relative to Sham Epithelium, With Average logFC and Adjusted *P* Values (logFC: log fold change; SBR: small bowel resection)

Gene	logFC	P
Apoa4	2.05	<0.001
Fabp1	1.76	<0.001
Apoc3	1.46	<0.001
Lct	1.11	<0.001
Rbp2	1.10	<0.001
Fapb2	0.96	<0.001
Sepp1	0.92	<0.001
Leap2	0.79	<0.001
Ephx2	0.71	<0.001
Apob	0.69	<0.001

Figure 2.4: Identification of signature proximal small intestine transcripts that increase after SBR. Violin plots (left) show relative expression of transcripts in sham vs SBR populations and UMAP plots (right) show relative transcript expression levels within cell populations. ****$P < 0.0001$.

Increased *Apoa4* and *Rbp2* expression following SBR has been previously described and is closely related to the transport of Vitamin A—an essential nutrient typically absorbed in the proximal bowel.[35,124,125] Contrary to our findings, one of these studies reported no significant changes in ileal *Fabp1* mRNA expression. However, this study examined whole tissue preparations, rather than epithelium at single-cell resolution which is likely to cause loss of potentially subtle changes.[124] Thus, to validate the transcriptional changes revealed by our scRNA-seq analysis, we surveyed *Fabp1* expression via RNA fluorescence in situ hybridization (RNA-FISH) on histological sections of sham and SBR animals at 7 and 70 days post-surgery, with the latter analysis designed to investigate whether the observed changes are stable over time, a property rarely investigated in the context of SBR. At day 7, *Fabp1* showed a 1.5 average

log2-fold AU increased expression in SBR, relative to sham ($P < 0.001$), and this response was maintained through day 70 post-surgery (1.4 average log2-fold AU increased expression, $P < 0.01$) (Figure 2.5*A*). Such stability over time indicates that the changes in lipid transport via *Fabp1* are sustained long-term after resection and adaptation to contribute to the functionality of the shortened intestine after adaptation.

To further validate our findings, we performed immunofluorescence and quantitative polymerase chain reaction (qPCR) for selected proteins and transcripts, including FABP1, FABP6, *Rbp2*, and *Ephx2*. These specific genes and proteins were chosen due to their relationship to lipid and nutrient metabolism commonly associated with the proximal or distal intestine in order to represent the "proximalization" of molecular expression we associate with regional reprogramming of the distal intestine. Immunofluorescence demonstrated significant increases in FABP1 (1.2 average log2-fold AU, $P < 0.001$) in SBR relative to sham (Figure 2.5*B*). In contrast, FABP6 was significantly decreased in SBR compared with sham (–1.2-fold relative fluorescent intensity, $P < 0.05$) (Figure 2.5*B*). qPCR confirmed upregulation of *Rbp2* and *Ephx2* (Figure 2.5*C*).

Together, these orthogonal validations confirm our scRNA-seq results, demonstrating that gene expression programs to support proximal SI nutrient processing are engaged following SBR. These mRNA-level changes underlie the regional reprogramming to mature proximal identity we observe in SBR. Such experiments served dual purpose in validating our study with benchmark studies in the literature reporting changes in gene expression and also introduces new potential candidates representing proximalization after resection in genes such as *Ephx2*.

37
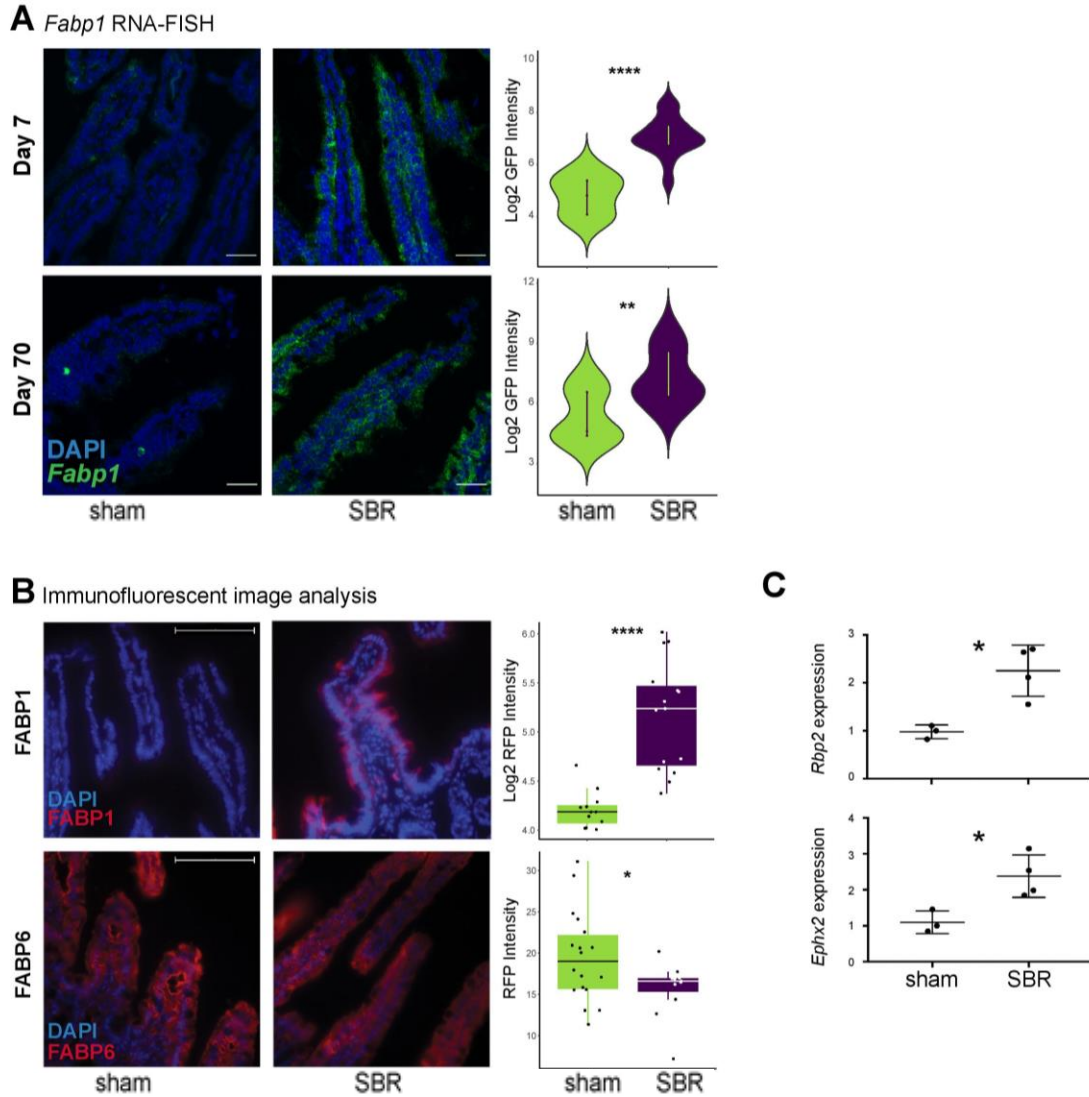
Figure 2.5: Validation of proximal small intestine markers that increase after SBR. (*A*) RNA FISH for *Fabp1* shows significant upregulation of transcripts (fluorescein signal) per nucleated cell (DAPI) at days 7 and 70 after surgery. Day 7: n = 15 sham images, n = 15 SBR images (3 biological replicates). Day 70: n = 12 sham images, n =15 SBR images (3 biological replicates) (images acquired using Olympus FV1200 Confocal Microscope). (*B*) Immunofluorescence staining images for FABP1 (n = 12 sham images, n = 15 SBR images, 3 biological replicates) and FABP6 (n = 18 sham images (4 biological replicates), n =11 SBR images (3 biological replicates)) (images acquired using Nikon Eclipse 80i with Ds-Ri2 camera). (*C*) Top to bottom: qPCR validation of upregulated SBR genes *Rbp2* and *Ephx2* from SI tissue (n = 3 sham and n = 4 SBR mice). RNA-FISH images are at magnification 60×, scale bar = 30 μm. IF are at 40×, scale bar = 100 μm. All graphs are presented as mean ± SD. *$P < 0.05$, **$P < 0.01$, ****$P < 0.0001$.

2.6 *Creb3l3* Shows Stable Upregulation in SBR Mice

We next aimed to identify candidate upstream transcription factors responsible for driving the observed regional reprogramming and shift toward a mature proximal nutrient transport and processing profile following SBR. While previous studies have seen changes in upregulation of proximal-related transporters and metabolic factors, it is unknown how these may be regulated. Of 44 transcription factors previously shown to be differentially expressed between proximal vs. distal enterocytes in the literature, only two associated with proximal identity, only two were upregulated in SBR. cAMP responsive element binding protein 3 like 3 (*Creb3l3*), was upregulated in SBR (0.46 average log-fold increase in SBR, *P* < 0.0001) (Figure 2.6*B-C*).[106] *Creb3l3* is a master regulator of lipid and cholesterol metabolism, a role fitting with the increased lipid metabolism necessary in the ileum after SBR.[126–128] Metabolically speaking, *Creb3l3* has been shown to bind to and directly regulate metabolic genes such as those in gluconeogenesis and fatty acid synthesis and elongation; interestingly, it regulates *Apoa4,* a significant player in our differential gene expression in SBR.[126] Another proximally-related transcription factor known to play a role in differentiation of enterocytes, Krüppel-like factor 4 (*Klf4),* was also upregulated after SBR (0.26 average log-fold increase in SBR, *P* < 0.001*)* (Figure 2.6*B-C*) (note: classified as distal in Haber et al., 2017).[129] It is worth noting the possibility that a wider variety of proximal transcription factors were upregulated immediately after surgery and stabilized to baseline by day 7, when structural adaptation was complete and we collected our samples. Transcription factors can be difficult to detect using scRNA-seq due to low levels and transient expression, so it is likely there are more, undetected transcription factors at play. However, we focused our analysis on *Creb3l3* as a novel putative driver of stable regional reprogramming due to its more significant increase post-SBR, its interplay with lipid

metabolism including *Apoa4,* and its ability to be regulated by other well-known intestinal players such as *Hnf4a* and *PPARa*.[126] Overall, the literature regarding *Creb3l3* indicates great ability of this factor to respond to signals like stress and nutrients, like those produced by SBR, and to regulate genes involved in nutrition and metabolism—making it a novel upstream candidate for response during adaptation. We further explore these factors in Chapter 4.

We sought to determine whether proximal SI transcription factor *Creb3l3* expression was transiently upregulated in SBR, or whether it was critical to maintaining a long-term adaptive response. As *Creb3l3* regulates lipid metabolism, and the increased demand for ileal lipid absorption should persist indefinitely after SBR, we expected its expression to remain elevated. Indeed, RNA-FISH for *Creb3l3* at days 7 and 70 demonstrated its significant and long-term upregulation following SBR (day 7: 1.3 average log2-fold AU increase in SBR, $P < 0.001$; day 70: 1.2 average log2-fold AU increase in SBR, $P < 0.01$) (Figure 2.6*A-C*). This long-term upregulation indicates that *Creb3l3* has long-term effects on lipid metabolism regulation, indicating a regulatory effect of *Creb3l3* upstream of the known molecular changes (i.e.: lipid absorption and metabolism) accompanying adaptation.

2.7 Interactome Analysis Indicates Regional Reprogramming is Driven by Retinoid Metabolism and Signaling

Although *Creb3l3* was elevated at day 7 and sustained through day 70 after SBR—suggesting a continued role in maintaining adaptation—no significant differences in *Creb3l3* expression were observed at day 3 post-surgery (Figure 2.6*D*). This suggested that inductive upstream signaling at earlier stages of adaptation may be critical to driving proximalization, which is subsequently mediated and sustained by *Creb3l3*. This indicated that

an early/immediate upstream signal triggered downstream effects such as increased *Creb3l3*. As *Creb3l3* expression is known to be induced by a number of events including inflammatory cytokines, stress, and nutrient level alterations, we sought to determine what events in the intestine during adaptation may affect *Creb3l3* activity.[128]

To investigate this, we generated an *in silico* interactome from single-cell gene expression profiles of all analyzed sham and SBR cells to determine which differentially expressed genes were most strongly co-expressed, thereby inferring gene-gene relationships and pathways (Figure 2.6*E*). This approach identified a network of highly interacting genes induced by SBR, including *Apoa1*, *Fabp2*, and *Rbp2*. We used a total of 59 genes from the SBR interactome to perform gene list functional enrichment analysis (5 genes from the interactome were excluded from analysis as they were absent from the database).[130] This analysis generated a list of pathways, some which were well-known and expected, including "lipid digestion, mobilization, and transport" ($P = 2.290 \times 10^{-11}$) and "digestion of dietary carbohydrate" ($P = 2.857 \times 10^{-8}$), and some that were more thought-provoking, such as "PPAR signaling pathway" ($P = 3.933 \times 10^{-7}$) and "retinoid metabolism and transport" ($P = 1.935 \times 10^{-8}$).

In the context of this study, retinoid metabolism was of particular interest because it has been shown to play a key role in structural adaptation.[35,124,125,131,132] Retinoids are derived from Vitamin A, which must be obtained from the diet, and are mostly absorbed in proximal SI. In fact, a retinoid/Vitamin A gradient exists in the intestine from proximal to distal, meaning that this gradient could be disrupted by resection and anastomosis.[133–135] Retinoic Acid (RA) is the intracellularly bioactive hydrolysis derivative of Vitamin A, the active product of retinoid metabolism, and it interacts with retinoid X receptor (RXR) and RA receptor (RAR) heterodimers, which bind to RA response elements (RAREs) within the nucleus, to drive effects,
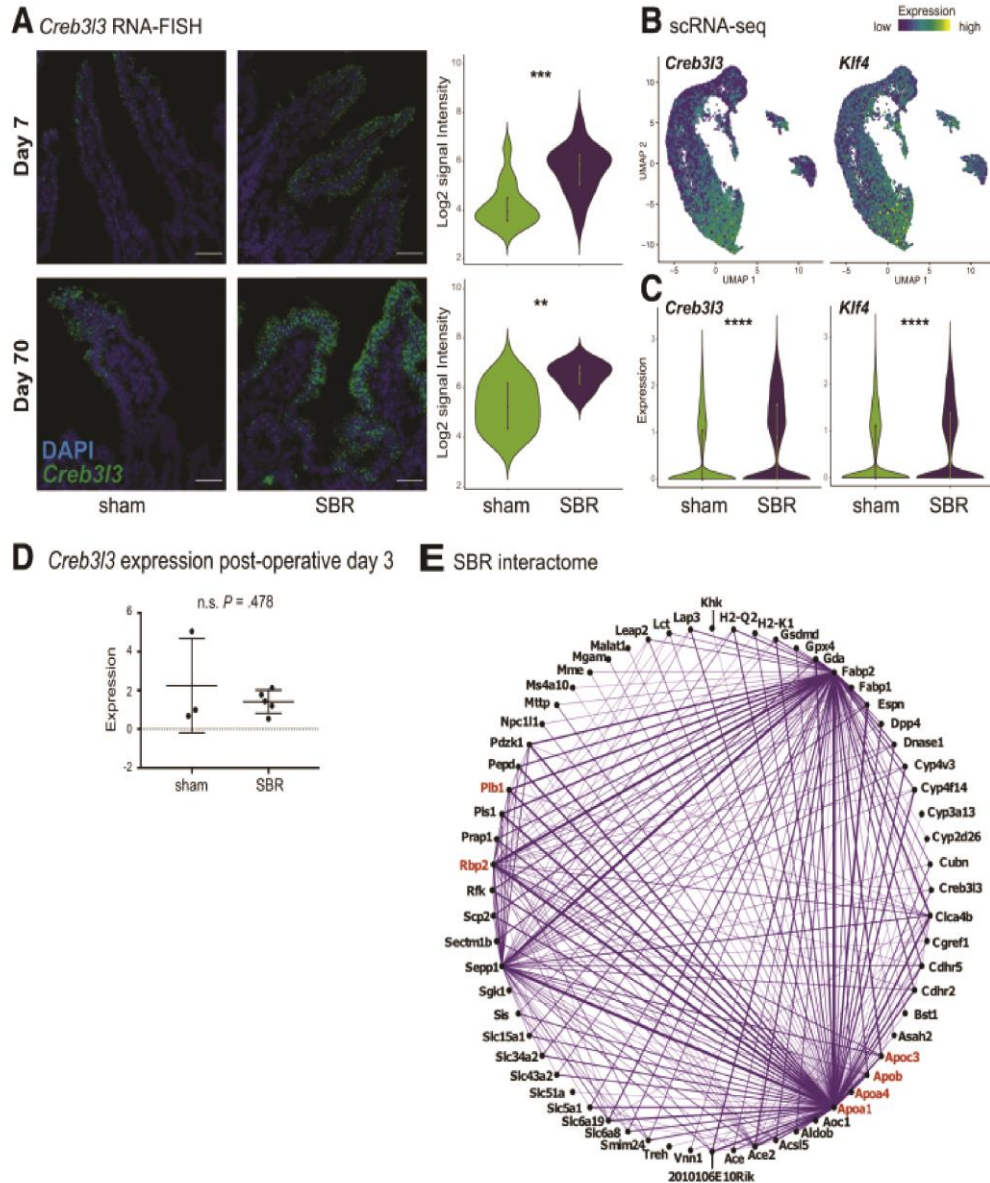
Figure 2.6: Dissecting genetic underpinnings of epithelial proximalization following SBR. (*A*) RNA-FISH for *Creb3l3* shows significant upregulation of transcripts (fluorescein signal) per nucleated cell (DAPI) at days 7 and 70 after surgery. Images are at 60×, scale bar = 30μm, acquired using Olympus FV1200 Confocal Microscope. Day 7: n = 15 sham images, n = 14 SBR images (3 biological replicates). Day 70: n = 9 sham images, n =10 SBR images (2 biological replicates). (*B*) Projection of *Creb3l3* and *Klf4* transcript enrichment onto the UMAP plot shows increased expression with enterocyte maturation. Color scale bar indicates relative intensity of gene expression. (*C*) Violin plots showing differential expression of *Creb3l3* and *Klf4* between sham and SBR. (*D*) Relative *Creb3l3* expression in SI from day 3 postoperative sham and SBR mice (n = 3 sham and n = 5 SBR mice) was measured using qPCR. (*E*) Interactome of genes upregulated in SBR epithelium. Genes in red are involved in RA signaling. All graphs are presented as mean ± SD. ns, not significant. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001.

hereby referred to as "RA signaling".[136] Notably, mice deficient in dietary Vitamin A do not adapt after SBR, and it was observed that administration of RA drives adaptation in part via regulation of enterocyte proliferation, migration, and apoptosis.[131,132] However, relatively little detail on the molecular changes induced by RA signaling has been revealed so far.

To investigate how retinoid metabolism leading to RA signaling induces the transcriptional changes accompanying SBR, we mined our dataset for genes differentially expressed between sham and SBR treatments which were either putative targets of RA signaling based on the literature, or contain a predicted RARE as determined by FIMO Motif Search (Table 2.2).[137] As a result, we found 45 genes differentially expressed between sham and SBR that are predicted to likely respond to RA signaling (Table 2.2). Several of these genes, including *Plb1* (phospholipase B1), *Rbp2* (retinol binding protein 2), *Apoa1* (apolipoprotein A1), *Apoa4* (apolipoprotein A4), *Apob* (apolipoprotein B), and *Apoc3* (apolipoprotein C3) (highlighted in red in Fig 2.6*E*), are known to be active in retinoid metabolism and transport, 4 of which were are also found in our list of the top 10 most differentially expressed genes in SBR (Table 2.1). Importantly, *Rbp2* (1.1 log fold change enriched, $P < 0.001$) is preferentially induced by RA in differentiated cells, such as the expanding enterocyte population we observe after SBR.[138] Furthermore, motif analysis indicated that *Creb3l3*, our novel proximal transcription factor, contains a RARE and can be activated by RXRα based on ENCODE transcription factor targets, making RA signaling a likely upstream regulator of the cascade of events attributed to *Creb3l3* activation (Table 2.3). Together, these findings confirm that RA signaling and metabolism is induced in cells responding to SBR, placing these signals upstream of the key transcriptional changes we observe, supporting a crucial role for RA signaling in adaptation both on a structural level, inducing enterocyte proliferation as shown in the literature,

43

and a molecular level inducing downstream changes in lipid transport and other nutrient

metabolism, as shown by our high-resolution scRNA-seq.[125]

Table 2.2: Genes Increased in Small Bowel Resection vs Sham that are Putative Responders to RA Signaling (rows in bold indicate prediction to contain retinoic acid response elements from Find Individual Motif Occurrences.[137] Rows in italics indicate putative direct and indirect retinoic acid signaling targets from literature.)

Apoa4	*Apob*	*Gda*	*Mgam*	*Creb3l3*
Slc43a2	*Treh*	*Gpx4*	*Gk*	*Cdhr5*
Fabp1[139]	*Ace2*[140]	*Slc2a2*[141]	*Mme*[142]	*Pdzk1*[143]
Nudt4[144]	*Plb1*[145]	*Dio1*[141]	*App*[146]	*H2-Q2*[141]
Klf4[147]	*Lap3*[144]	*Ace*[146]	*Clca4b*[148]	*Slc5q1*[144]
Mical1[144]	*Fos*[149]	*Chka*[150]	*Rnf128*[151]	*Fbln1*[152]
Prap1[144]	*H2-K1*[141]	*Vnn1*[153]	*Gls*[154]	*Malat1*[155]
Aqp1[156]	*Clec2e*[157]	*Slc6a19*[144]	*Apol7a*[144]	*Rbp2*[158]
Neat1[159]	*Ogdh*[160]	*Egr1*[161]	*Apoa1*[162]	*Apoc3*[163]

Table 2.3: Genes Identified in Small Bowel Resection vs Sham that are Target Genes of Retinoid X Receptor Alpha from ENCODE Transcription Factor Targets Dataset (Human) (Source: http://amp.pharm.mssm.edu/Harmonizome/gene_set/RXRA-ENCODE+Transcription+Factor+Targets)

Fabp1	*Apoc3*	*Sepp1*	*Leap2*	*Ephx2*	*Apob*	*Apoa1*	*Ace2*	*Slc2a2*	*Dnase1*
Pdzk1	*Creb3l3*	*Slc43a2*	*Acsl5*	*Pls1*	*Chka*	*Ano6*	*Rnf128*	*Fbln1*	*Prap1*
Rfk	*Khk*	*Malat1*	*Pepd*	*Gsdmd*	*Neat1*	*Ogdh*	*Dhrs1*	*Gpx4*	*Egr1*
Gk	*Acox1*	*Cdhr5*	*Nudt4*						

2.8 Discussion

Here, we report epithelial single-cell analysis of ileum following proximal SBR, showing

expansion of cells identified as mature proximal enterocytes in SBR vs. sham mice. SBR

enterocytes differentiated significantly on the basis of solute and nutrient transporters typically

associated with proximal SI, especially with regard to lipid metabolism. Because the duodenum

and jejunum, proximal portions of the small intestine, absorb the majority of nutrients under normal conditions, we propose that this "regional reprogramming," driven by transcriptional proximalization, is a critical component of the adaptation response to SBR in the remnant ileum. This is a novel principle, as the typically studied structural adaptation, which increases absorptive surface area, does not necessarily facilitate functional adaptation.[111] Rather than simple tissue hyperplasia, as others have suggested, we demonstrate that enterocyte-level alterations in transcriptional profiles of distal cells occurs after SBR, in order to mimic proximal SI function.[164] These changes mimic proximal SI function through the production of critical nutrient and lipid transporters typically found on proximal enterocytes in the enterocytes of the distal intestinal epithelium.

These findings highlight the significant contribution that single-cell analysis makes toward our understanding of organ pathophysiology. While previous studies could identify a small pool of "proximal" genes increasing in the ileum after resection, our scRNA-seq expands the pool of "proximalization factors" and allows us to look at what cell types and identities are changing during this process. Of note, we saw that the SBR epithelium also retained a mature distal enterocyte population (Figure 2.2*B* and 2.3*B*), and differential gene expression results suggest largely preserved ileal bile acid metabolism and cobalamin absorptive function. For example, bile acid metabolism genes *Slc10a2*, *Nr5a2*, *Slc51b*, *Abcc3*, and *Nr1h4*, and also cobalamin metabolism genes *Lrp2* and *Tcn2*, showed no significant differential expression after SBR. Enterocyte basolateral bile acid transporter *Slc51a* was depleted in SBR, (0.29 average log fold change, $P < 0.0001$), while cobalamin receptor *Cubn* was enriched (0.54 average log fold change, $P < 0.0001$). These results suggest a hybrid proximal-distal identity of ileal epithelium

following SBR, gaining functions lost from resection but not losing important distal functions as well, a true compensatory effect.

While investigating causative signaling mechanisms driving these proximalization changes, we identified upregulation of a transcription factor associated in the literature with proximal SI: *Creb3l3*. Upstream analysis reiterated the importance of retinoid metabolism to the adaptation response, and the depth of our analysis allowed identification of previously undescribed transcriptional changes likely mediated by RA signaling and Vitamin A gradients after SBR, including changes in *Creb3l3* expression levels. As *Creb3l3* is known to respond to nutrient availability among other things, the prediction of an RXR binding site (Table 2.3) for Vitamin A signaling provides a viable hypothesis for Vitamin A-induced activation of *Creb3l3*. We did not identify strong evidence for perturbed intestinal stem cell regional identity based on previously identified markers, consistent with previous observations that RA exerts its effects on differentiated cells. In addition, *Creb3l3* is localized to villus enterocytes rather than crypt-based cells (Figure 2.6*A*), again providing reasonable evidence that *Creb3l3* likely responds to Vitamin A and RA signaling through differentiated enterocytes and increased exposure of villus enterocytes to gradients in the lumen compared to crypt cells.[106,138] We therefore conclude that a critical component of the adaptation response to massive proximal SBR is transcriptomic "proximalization" of distal SI enterocytes, and that this is driven at least in part by an RA signaling gradient, which is upstream of "proximalization" transcription factors such as *Creb3l3* and signaling cascades. The appearance of *Creb3l3* as a potential driver of regional reprogramming directed our further interest in the potential of using known (from the literature) and novel (from our scRNA-seq data) transcription factors for direct reprogramming to investigate regional reprogramming in organoids *in vitro* (Chapter 4).

As it is a product of Vitamin A to retinoid to active RA metabolism, RA is derived entirely from the diet. Following proximal SBR, the ileum becomes exposed to nutrients and molecules in the luminal content which would otherwise have been largely absorbed more proximally, most notably lipids and other essential nutrients. Essential nutrient Vitamin A is among the list of nutrients commonly absorbed in the proximal intestine that we propose provides new stimulus to the ileum after resection as Vitamin A is found in a proximal-distal gradient in the small intestine.[134,165] When the ileum bypasses the previous jejunal and duodenal length once responsible for Vitamin A absorption post-resection, this new exposure to RA/Vitamin A provides a growth stimulus that is a likely driver of both structural and functional adaptation through resultant increases in RA signaling in ileal enterocytes.[121,133] Similar to the diffusible reprogramming factors described in the early days of reprogramming research, this alteration in the endogenous amount of Vitamin A available to the distal cells post resection provides us with a new hypothesis: diffusion of factors endogenous to the proximal intestine extends into the distal intestine post resection and we can harness these endogenous factors *in vitro* to model regional reprogramming by exposing distal organoids to proximal organoids as presented in Chapter 3.[64,66]

A series of studies utilizing an ileostomy model of SBS in mice and zebrafish have shown the critical effects of mechanoluminal flow on the structural adaptation process after SBR, including loss of structural adaptation and cellular proliferation in the distal bowel when isolated from the flow of luminal contents.[29,166] The importance of luminal nutrition or enteral feeding has also been highlighted by other groups.[167,168] As such, exposure to increased levels of dietary Vitamin A is a likely mechanism driving structural adaptation, consistent with prior reports.[35,125,131,132] We hypothesize that additional proximal factors likely play a role in this

mechanism as well. In addition to Vitamin A, the ileum also becomes exposed to higher levels of dietary fatty acids after SBR, and this likely constitutes an additional driving force for adaptation through "proximal" endogenous factor gradients.[121]

The PPAR signaling pathway was implicated by our interactome analysis. We investigated this further since PPARs form heterodimers with RXRs to activate PPAR response elements (PPREs) in the induction regions of many genes involved in lipid metabolism, including *Creb3l3,* which is well known to interact with PPAR.[128,169–171] Interestingly, treatment of mice with a PPARα agonist induced villus growth by facilitating cell differentiation, similar to the adaptation phenotype observed after SBR in which there are elongated villi with an expansion of mature enterocytes.[172,173] Ultimately, we observed neither changes in *Ppara* expression (1.5% decrease 3 days after SBR via qPCR, $P = 0.96$, not shown), nor significant changes in *Pparδ* expression (55.2% decrease 3 days after SBR, $P = 0.16$, not shown). Less is known about PPARδ, though it is thought to interact with corepressors and function as an inhibitor of PPARα.[147] Given these findings, we suspect that either (1) PPAR signaling is important to SBR adaptation, but was not captured in our analysis temporally, or (2) minimal to no transcriptional change in these specific genes is needed to drive a significant biological effect. In support of option 1, as we found that *Creb3l3* expression changes were not detectable by 3 days but were by 7, and *Creb3l3* and *Ppara* are interacting partners, it is possible that changes in PPARs occur later, post 3 days after surgery, and were therefore not detected. Further studies looking at the potential role of PPARs in early adaptation would be interesting.

In summary, our analysis has revealed a significant shift in metabolic machinery and regional identity at the enterocyte level following SBR. Moving forward, additional studies are warranted to better delineate causal factors driving changes between sham and SBR enterocytes,

in conjunction with or independent of RA signaling. This is especially true considering prior studies that demonstrated proliferative and morphometric effects of circulating factors on structural "jejunalization" of ileum following SBR, which implicates additional, non-luminal stimuli in driving structural adaptation.[1,174] Similar studies exploring molecular changes in response to circulating factors would provide further insight. Discerning the stimuli for functional proximalization of ileum following SBR will prove critical, as it will provide insight toward targeted therapeutic approaches, via the enteral or parenteral route, for patients suffering from SBS. Targeted therapeutic approaches combining aspects of adaptation from both structural and molecular avenues could induce heightened adaptive responses, yielding better patient outcomes. For example, as RA signaling seems to play a role in both structural adaptation and regional reprogramming, it is possible that therapeutic approaches involving increased Vitamin A in enteral nutrition or administration of Retinoic Acid to the small bowel could improve patient outcomes. Further studies regarding the exact role of RA and the pathways it activates, perhaps through transcription factor activation, are warranted.

However, further studies on the influence of regional luminal factors such as Vitamin A/RA (and others) and the effects of transcription factors are complicated in the multi-variate *in vivo* environment—as many confounding variables are hard to control within the mouse. Therefore, we propose that future studies should focus on moving from the mouse model of SBR to an *in vitro* model, using organoids. Using organoids allows for a controlled environment and ease of scaling up for experimentation. We will focus on the use of organoids to model regional reprogramming via diffusible proximal signaling and transcription factor direct reprogramming in Chapters 3 and 4. These studies aim to develop models of regional reprogramming that can be used for future study of SBS and SBR responses in patients.

2.9   Conclusions

Here, we have characterized the transcriptome of adapted intestinal epithelium at the single-cell level following massive SBR, a laboratory model for SBS, using scRNA-seq. Our analysis revealed the emergence of unique enterocyte gene expression patterns between sham and SBR mice, which distinguished themselves on the basis of proximal vs. distal SI patterning and cell identity, including critical absorptive features. Pathways driving these changes, such as proximal signaling and transcription factor activation, deserve further investigation, as they underlie the functional aspects of adaptation to SBS, which allow progressive tolerance of enteral feeding and weaning from PN for patients.

2.10 Materials and Methods

2.10.1 Mice

A total of 50% proximal SBR was performed on male C57/B6 mice at 8–12 weeks of age, according to standard protocol.[31] Briefly, the SI was extruded via a midline laparotomy and the ileocecal valve identified. The SI was transected 12 cm proximal to the ileocecal valve, and ~2 cm distal to the ligament of Treitz. The intervening SI was removed, the mesentery ligated with 3-0 silk suture, and the proximal and distal ends approximated with interrupted 9-0 nylon stitches. Sham surgery, consisting of distal transection and anastomosis only, was performed as control. Peritoneum and skin were approximated in separate layers, animals were resuscitated with a subcutaneous bolus of normal saline (repeated on postoperative day 1) and co-housed in a 33°C incubator until the end of the 7 day experiment. For longer studies (70 days), mice were moved to room temperature at day 7. Liquid diet (TestDiet PMI Micro-Stabilized Rodent Liquid Diet LD 101) was initiated 24 hours before surgery, withheld the morning of surgery, and subsequently provided on postoperative day 1 until the end of the experiment. Food and water

were available ad libitum, and animals were housed under 12-hour light/dark cycles with corn cob bedding and nestlet enrichment. All surgical and animal care procedures were approved by the Washington University Institutional Animal Care and Use Committee, and meet Animal Research: Reporting of In Vivo Experiments standards.

2.10.2 Tissue Isolation and Processing

At day 7 after surgery, epithelium was isolated from a 1-cm segment of SI 3 cm distal to the anastomosis, similar to previously published protocols.[175,176] Briefly, the SI was flushed with ice cold sterile saline, filleted lengthwise, and placed in a conical tube containing ice cold 30 mM EDTA in phosphate-buffered saline (PBS). After 15 minutes on ice without agitation, the SI segment was transferred to a fresh conical tube containing 30-mM EDTA in PBS, briefly shaken, and placed in a 37°C water bath for 15 minutes. Subsequently, the tube was shaken aggressively by hand for 2 minutes. Subepithelial tissue floated to the top and was removed, and epithelium was pelleted by centrifugation. Epithelium was then re-suspended in a 0.3-U/mL dispase solution (StemCell Technologies #07923) and incubated at 37°C for 15 minutes, shaking every 2 minutes. The solution was then quenched with media containing fetal bovine serum to a final concentration of 5%, pipetted several times, and sequentially passed through 100-, 70-, and 40-μm filters. Single-cell suspensions were confirmed by microscopy, pelleted by centrifugation, and resuspended in 200-μL ice cold PBS. Then, 800-μL ice cold 100% methanol was added, dropwise, with gentle mixing between drops. Samples were immediately stored in 80% methanol in PBS at –80°C, according to Alles et al., 2017, for later processing.[177] For Western blotting analysis, whole epithelium was isolated using 30-mM EDTA, pelleted by centrifugation, and lysed in sodium dodecyl sulfate sample buffer (50-mmol/L Tris-HCL, pH 6.8, 2% sodium

dodecyl sulfate, 10% glycerol, and 5% 2-mercaptoethanol). Lysate was heated to 100°C and stored at –20°C prior to processing. Protein concentration was measured using the RC DC (reducing agent and detergent compatible) Protein Assay Kit II (Bio-Rad 5000122). For qPCR experiments, RNA was isolated from homogenized whole SI using the standard Trizol method. RNA concentration was measured using a NanoDrop Spectrophotometer (NanoDrop Technologies ND-1000). A total of 1-μg RNA was converted to complementary DNA (cDNA) using qScript cDNA Synthesis Kit (Quanta Bio 95047), according to the manufacturer's instructions, and stored at –20°C until use.

2.10.3 scRNA-seq Library Preparation

For single-cell library preparation on the 10x Genomics Chromium platform, we used the Chromium Single 3′ Library & Gel Bead Kit v2 (PN-120237), Chromium Single Cell 3′ Chip kit v2 (PN-120236), and Chromium i7 Multiplex Kit (PN-120262), according to the manufacturer's instructions in the Chromium Single Cell 3′ Reagents Kits V2 User Guide. Methanol-fixed cells from sham (n = 3) and SBR (n = 3) animals were pooled for the first batch. Methanol-fixed cells from sham (n = 2) vs SBR (n = 1) were processed individually in a separate experiment for a final sample size of sham (n = 5) and SBR (n = 4). Just before cell capture, methanol-fixed cells were placed on ice, then spun at 3000 rpm for 5 minutes at 4°C, followed by resuspension and rehydration in PBS, as previously described.[177] Resulting cDNA libraries were quantified on an Agilent Tapestation and sequenced on an Illumina HiSeq 2500.

2.10.4 scRNA-seq Analysis

The Cell Ranger v2.1.0 pipeline was used to align reads to the mm10 genome build, and generate a digital gene expression (DGE) matrix: (https://.support.10xgenomics.com/single-cell-gene-expression/software/downloads/latest). For initial filtering of these DGE matrices, we first

excluded cells with a low number (<200) of unique detected genes. We then excluded cells for which the total number of unique molecules (UMIs) (after $\log_{10}$ transformation) was not within 3 standard deviations of the mean. This was followed by the exclusion of outlying cells with an unusually high or low number of UMIs/genes given their number of reads by fitting a Loess curve (span = 0.5, degree = 2) to the number of UMIs/genes with number of reads as predictor (after $\log_{10}$ transformation), removing cells with a residual more than 3 standard deviations the mean. Finally, we excluded cells in which the proportion of the UMI count attributable to mitochondrial genes was >25%. Raw and processed data files are available via GEO: accession number GSE130113. After filtering and normalization of the DGE, the R package Seurat (Version 3) was used to cluster and analyze the single-cell transcriptomes.[102] Independent biological replicates from the sham and SBR surgeries were integrated by Canonical Correlation Analysis, identifying common sources of variation to align the datasets, reducing batch effects.[112] Highly variable genes were identified and used as input for dimensionality reduction via canonical correlation analysis. The resulting Canonical Correlation Vectors and the correlated genes were examined to determine the number of components to include in downstream analysis, followed by clustering and visualization via UMAP.[103]

2.10.5 Quadratic Programming Analysis to Assess Cell Identity and State

QP, previously described in Treutlein et al., 2016 and successfully modified and used by our group in Biddy et al., 2018, was used to score cell identity. Here, we created a reference of SI epithelial cell types, collected previously.[79,106,113] The R Package QuadProg was used for QP to generate cell identity scores.

2.10.6 Immunohistochemistry

Ileal tissue adjacent to the region collected for single-cell preparation was fixed in 10% neutral buffered formalin, paraffin embedded, and sectioned at a thickness of 5 μm. Deparaffinization and immunolocalization were performed as previously.[178] Briefly, slides were deparaffinized in xylene, rehydrated in sequential ethanol baths, and prepared in 3% hydrogen peroxide in methanol. Antigen retrieval was performed using 1× Diva Decloaking Solution (Biocare Medical DV2004), and blocking was performed using the Avidin-Biotin kit (Biocare Medical AB972L). Primary antibodies were diluted in Da Vinci green (Biocare Medical PD900L) and incubated overnight at 4°C. Slides were rinsed in phosphate-buffered saline + TWEEN 20 (PBST), incubated in biotin-labeled secondary IgG diluted in PBST, rinsed in PBST, incubated in streptavidin-horseradish peroxidase diluted in PBST, developed in DAB (Sigma-Aldrich D9015), counterstained with hematoxylin and bluing agent, run in successive dilutions of ethanol, and xylene, and cover-slipped using MM 24 mounting medium (Surgipath 100109). Of note, samples used for confirmatory staining were from a different litter of mice than those used for scRNA-seq analysis. This was done to validate consistency of results across cage and littermates, which has been previously reported as a confounding variable in murine gastrointestinal research.[179] At least 3 sham and 3 SBR samples were analyzed after surgery; 20× images representative of the sample were obtained by a blinded investigator using a Nikon Eclipse 80i with Ds-Ri2 camera and NIS Elements V4.3 software (Nikon Instruments, Inc).

2.10.7 Quantitative PCR

cDNA was amplified using TaqMan Gene Expression Master Mix (Applied Biosystems 4369016) and the specified primer probe on the Applied Biosystems 7500 Fast Real-Time PCR

system. Primer probes were *Creb3l3* (Mm00520279_m1), *Ephx2* (Mm01313813_m1), and *Rbp2* (Mm00436300_m1), all from Applied Biosystems.

2.10.8 Immunofluorescence

Ileal tissue adjacent to the region collected for single cell-preparation was fixed overnight in 4% paraformaldehyde and then overnight in 30% sucrose before being embedded in O.C.T. Compound (Fischer Healthcare 23-730-571), sectioned at 5 μm, and stored at –80°C until use. Slides were washed in PBS and blocked in 5% goat serum with 0.3% Triton X-100 (Sigma-Aldrich T8787) prior to incubation with primary antibody in antibody staining solution (1% goat serum and 0.3% Triton X-100) at 4°C overnight. Slides were rinsed in PBS for 5 minutes 3 times, and secondary antibody was applied in antibody staining solution for 1 hour at room temperature. A total of 300-nM DAPI (Invitrogen D1306) was applied for 1 minute, and slides were rinsed in PBS for 5 minutes 3 times before a cover slip was applied using ProLong Gold Antifade Mountant (Invitrogen P10144). The 40× images representative of the sample were obtained by a blinded investigator using a Nikon Eclipse 80i with Ds-Ri2 camera and NIS Elements V4.3 software (Nikon Instruments, Inc). Images were subsequently analyzed using the FIJI Distribution of ImageJ for FABP6, or unbiased computational quantification (FABP1; see "Quantitative analysis of RNA-FISH images"). Representative images were chosen.

2.10.9 ImageJ Quantitative Analysis

For FABP6, the full area of each villus was chosen as a region of interest. The mean intensity of that region was calculated to represent the fluorescence of the diffuse FABP6 staining. For percent villus occupancy, ImageJ was used to measure the areas of *Creb3l3* expression, and this was divided by the total villus length. For *Creb3l3*, there were 12 sham villi (3 biological replicates) and 11 SBR villi (3 biological replicates).[180]

2.10.10 RNA Fluorescent in Situ Hybridization

Tissues were fixed and sectioned as previously described in methods for immunohistochemistry and immunofluorescence. RNA-FISH was performed using the RNAscope Multiplex Fluorescent v2 kit (Advanced Cell Diagnostics 323100), following the protocol for Fixed Frozen Tissue. Briefly, tissue was pretreated with RNAscope Hydrogen Peroxide (Advanced Cell Diagnostics 322335), target retrieval was performed for 5 minutes, and tissue was treated with RNAscope Protease III (Advanced Cell Diagnostics 322337). Then, specified probes were hybridized using the RNAscope HybEZ II Oven (Advanced Cell Diagnostics 321710/321720). Probes were then amplified and the HRP signal was developed using TSA Fluorescein Plus Evaluation Kit (Perkin Elmer NEL741E001KT). Finally, slides were counterstained with DAPI (Advanced Cell Diagnostics 323108) and mounted with ProLong Gold antifade reagent (Life Technologies Corporation P10144). Imaging was performed on an Olympus FV1200 Confocal Microscope, in which multiple 60× images were taken per sample, of which 2–3 samples were used for each timepoint (day 7 vs. day 70) in each treatment (sham vs. SBR). Images were subsequently analyzed using unbiased computational quantification and representative images were chosen.

2.10.11 Quantitative Analysis of RNA-FISH Images

RNA-FISH images were processed with a custom python script to quantify gene expression level at single-cell resolution. Individual cell segmentation was achieved based on the nuclear signal. First, DAPI images were transformed into binary images by thresholding the fluorescent signal. The threshold values were determined by the Otsu method.[181] Binarized nuclei images were processed by the watershed segmentation method to completely separate individual objects. The images were subjected to 2-step quality check: filtering of objects and

filtering of images. First, inappropriately sized objects were filtered to remove noise and cell multiplets. Then, images with a large number of inappropriate objects were removed. The intensity of the fluorescent signal per individual cell area was then quantified. Fluorescent signals per image were averaged to obtain mean signals per sample and treatment.

2.10.12 Gene Coexpression and Interactome Analysis

To further reveal relationships among differentially expressed genes following SBR, we constructed gene coexpression networks using weighted gene correlation network analysis, adapted for single-cell analysis as in https://hms-dbmi.github.io/scw/WGCNA.html. The analysis was performed using the R package WGCNA. In brief, differentially expressed genes between SBR and sham samples were identified via Seurat analysis, as previously (using the function *FindMarkers*). Genes that were upregulated in SBR treatment were used for network construction. Correlations of each gene pair among all significantly differentially expressed genes were then calculated. A gene-gene correlation matrix was used to construct an adjacency matrix by raising the correlations to a soft-threshold power, from which Topological Overlap Matrix (TOM) was further computed to remove spurious correlations. With the TOM matrix, the algorithm identifies modules/clusters of genes via clustering using Ward's method and Dynamic Branch Cut methods. Here, to identify the most significant connections, we select the top 5% of the most differentially expressed genes for visualization. The network was visualized using CytoScape based on the TOM matrix.

2.10.13 Statistical Analysis

Statistical comparison of the QP-generated identity scores between the 2 groups, sham and SBR, was performed using an unpaired Student's *t* test. For qPCR studies, statistical analyses were performed using Prism 7.0 (GraphPad Software). Differences in protein

expression and mRNA expression between sham and SBR were assessed using unpaired Student's t tests. Graphs with error bars represent mean $\pm$ SD. $P < 0.05$ was considered significant. For immunofluorescence quantitative analysis using FIJI, FABP6 replicate values were averaged between sham and SBR treatment identities. A Wilcoxon rank sum test was performed to determine statistical significance. For quantitative analysis of RNA-FISH images, and FABP1 immunofluorescence, statistics were calculated using the Wilcoxon rank sum test. These processes were run with Python 3.6.7 and its libraries: scikit-image 0.13.1, numpy 1.14.3, pandas 0.24.1, oiffile 2019.1.1, matplotlib 3.0.3, seaborn 0.8.1, and jupyter 1.0.0.

2.11 Acknowledgements

2.12 Authors' Contributions

Chapter 3: Novel Chimeric Intestinal Organoids as a Preliminary Model of Regional Reprogramming

3.1 Abstract

Moving studies from the *in vivo* environment to an *in vitro* environment is critical to the future applicability of translational medicine to patient-specific purposes. Issues like SBS and the accompanying process of regional reprogramming are difficult to model and study as they currently require the utilization of an animal model, traditionally mouse, that does not exactly recapitulate the human response. In addition, the *in vivo* models are low throughput and expensive, leading researchers to look *in vitro* for the future of disease and treatment modeling. In this chapter, we utilize findings from our *in vivo* study, specifically the presence of region-specific diffusible factors influencing regional reprogramming, to approach *in vitro* modeling of regional reprogramming utilizing intestinal organoids. We develop a novel, fully *in vitro*, chimeric intestinal organoid model ("chimera") to recapitulate murine regional reprogramming in a dish and investigate novel signaling in the process. We find that our chimera model likely increases proximal identity of distal organoid enterocytes based on scRNA-seq data and computational analyses. Additionally, we note that exploration of cell-cell signaling pathways through scRNA-seq data analysis in the chimera identifies Guanylate Cyclase-C signaling as playing a role in the transcriptional changes seen within the chimeric model. We conclude that our preliminary chimeric organoid model provides an *in vitro* environment capable of

60

recapitulating aspects of regional reprogramming, thereby producing a preliminary model for continued development and study of murine regional reprogramming after SBR.

3.2 Introduction

The field of stem cell biology, especially the focus on reprogramming and regeneration, is a relatively young but exponentially fast-growing field.[63] One focus of the field is on personalized medicine, or taking patient cells and reprogramming or differentiating them for therapeutic or disease modeling purposes *in vitro*. The focus on *in vitro* medicine is important in order for the field to move away from dependence on animal models that fail to completely recapitulate the human system and to increase experimental throughput for drug and treatment testing. The first step toward achieving this goal involves accurate *in vitro* models for disease processes that can be applied to humans.

The development of small intestinal organoids from mice and humans provides a unique opportunity to study intestinal disease and injury *in vitro.* Briefly, intestinal organoids are a 3-dimensional *in vitro* culture system in which intestinal stem cells self-organize into an enclosed structure containing the cell types and mimicking the functionality of the small intestinal epithelium.[7,43] Important to this study are two aspects of intestinal organoid biology. Firstly, when taken from the *in vivo* system and cultured into organoids *in vitro*, the intestinal cells retain their genetic memory of regional identity and even of disease states.[47,48] Organoids are currently being developed as a tool to model and test treatments for inflammatory bowel disease (IBD) and Crohn's Disease *in vitro* when derived from IBD or Crohn's patients.[182] Secondly, intestinal organoids have been shown to react to bowel resection through undergoing structural adaptation when engrafted *in vivo*.[1] Together, this evidence supports the use of intestinal organoids to model adaptation or regional reprogramming after SBR, as organoids maintain their *in vivo* genetic

61

backgrounds in *in vitro* culture and can undergo "adaptation" if exposed to the correct factors (for more detailed discussion of organoids, please refer to Subsection 1.1.2). The problem remains, though, of how to expose organoids to the "correct" factors to mimic adaptation *in vitro*.

Short Bowel Syndrome is a perfect system in which to utilize organoids as a model for treatment response as adaptation in mice is detectable via altered expression of nutrient and lipid transporters and other known genes, as discussed in Chapter 2. During our *in vivo* study in Chapter 2, we utilized scRNA-seq to identify a process known as regional reprogramming which occurs in the distal remnant intestine after resection. It is detectable based on the increased expression of proximal-related tissue markers in the distal intestine post-adaptation. Notably, our analysis identified Vitamin A metabolism (Retinoic Acid signaling) as an upstream trigger for regional reprogramming in the mouse. We hypothesized that upon removing the proximal and medial intestine, the distal intestine was exposed to much higher levels of Vitamin A signaling due to an endogenous gradient of Vitamin A in the small intestine from proximal to distal.[183]

Based on this finding, we generated a more general hypothesis: region-specific diffusible gradients of signals exist in the intestine and are interrupted by bowel resection such that changes in signal exposure play a role in proximalization (regional reprogramming) of the distal intestine. We further hypothesized that regional reprogramming could be modeled *in vitro* by reproducing the exposure of distal intestinal tissue to the proximal signaling environment using a chimeric organoid system and that we could further dissect the pathways at play using such an environment.

Our development of a chimeric organoid system to model regional reprogramming has great implications for the future study of SBS treatments for patients. Through building a model

using mouse organoids, we enable the further study of regional reprogramming as it occurs successfully in mice. In addition, we provide a framework for adaptation of this model to humans. Chimerization of human patient organoids using this model in the clinic could provide clinicians with valuable information regarding the extent of endogenous adaptation for individual patients and will guide treatment options through providing a high-throughput system to administer treatments *in vitro* and test effectiveness prior to treating the patient—lowering the administration rate of ineffective treatments for SBS in humans.[182]

In our model, we were inspired by the developmental model of the chimeric organism (see Subsection 1.1.3 for more details).[67] Groups working in kidney organoids have previously developed "chimeric" organoid systems in which they aggregate human stem cells with mouse kidney cells and allow them to develop into organoids then engraft them *in vivo* to achieve maturity of the cell types.[68,69] While these studies provided us with support that chimerization could influence cellular identity, the involvement of an *in vivo* maturation step needed improvement. In moving to the intestinal organoid system instead of kidney, we found an opportunity to utilize the unique attributes of the intestinal organoid—notably the presence of a stem cell niche mimicking the *in vivo* environment capable of influencing cell identity outside of an animal model.

We began by chimerizing iEP cells with mouse organoid cells to test the ability of the mouse organoid to influence endoderm progenitors via diffusible signals. We found that the iEPs chimerized with organoids took on more intestine-like identities. Based on this success, we refined the chimera model using the chimerization of proximal intestinal organoids with distal intestinal organoids. The resulting chimeras were analyzed using scRNA-seq for cellular identity changes in the distal enterocyte populations. We found preliminarily that distal organoid cells,

when exposed to the proximal organoid tissue in the chimera, acquire significantly higher proximal enterocyte identity scores than controls and that the chimeric environment changes Guanylate Cyclase-C signaling. However, these studies were underpowered and require further replication to increase confidence in results. We conclude that the chimeric organoid model recapitulates aspects of regional reprogramming *in vitro* and can be further developed to study mouse and human adaptation after SBR.

3.3   Pilot of Chimeric Mouse Organoids Induce Intestine-Like Transcriptomic Changes in iEPs

Taking a cue from chimeric kidney studies and the unique and valuable characteristics of intestinal organoids, we began development of the chimeric intestinal organoid system.[49,68,69] The Morris Lab works actively with the reprogrammed cell type known as induced endoderm progenitors (iEPs) (detailed discussion found in Section 1.1.3).[78,79] While these cells are directly reprogrammed progenitors themselves, studies have shown that *in vivo* transplantation of these cells in the mouse colon further reprograms them into functional intestinal cells.[77,184] Because of the known ability of these cells to reprogram *in vivo*, we chose them as our source cell type for testing our new *in vitro* methods mimicking the *in vivo* environment using the intestinal organoid stem cell niche.

In order to "hijack" the niche within the intestinal organoids, we developed our chimeric intestinal organoid ("chimera" or "chimeric organoid"). In the small intestine, it has been proposed that Paneth cells help to provide the niche for the Lgr5+ stem cells, encouraging the maintenance and differentiation of the stem cells.[49] It has been shown that dissociation of Lgr5+ stem cells and Paneth cells down to a single cell level, followed by centrifugation for

reaggregation, results in the re-forming of small intestinal organoids, just as dissociation of the kidney and reaggregation results in the formation of renal organoids.[49,68]

Based on this evidence, we generated the pilot chimera model through dissociation of iEPs and mouse small intestinal organoids and then centrifugation to aggregate them to form chimeras (Figure 3.1*A*). In brief, iEPs, identifiable by GFP expression, were cultured to maturity and mouse organoids were cultured in 3-dimensions using Nicotinamide to enhance the stem cell populations and enrich for the crypt-niche environment.[49] The iEPs were then dissociated to a single cell level while organoids were broken to small (2-5 cell) chunks to allow the presence of Paneth cells next to stem cells to encourage organoid re-growth. The populations were combined in a 1:2 (iEP to organoid) ratio in a low-bind U-bottom plate and centrifuged briefly to aggregate cells to the bottom of the plate. The 1:2 combination ratio reflects the competitive growth of iEPs vs. organoids to prevent competitive domination of a single cell type in culture. The aggregates were suspended in Matrigel and allowed to grow in IntestiCult Medium. After a growth period *in vitro*, we investigated the cells for chimerization via confocal microscopy (Figure 3.1*B).* We observed an infiltration of the GFP positive iEPs into the organoid structure, a preliminary indication of successful chimerization of the cell populations.

After the qualitative visual success of the chimerization process, we set out to determine if the chimeric iEPs were being influenced toward intestinal fates by their exposure to the stem cell niche of the *in vitro* organoids. We performed a pilot scRNA-seq study using Drop-seq in order to compare control iEPs, control intestinal organoid cells, and chimeric iEPs grown in our model (Figure 3.2).[185]
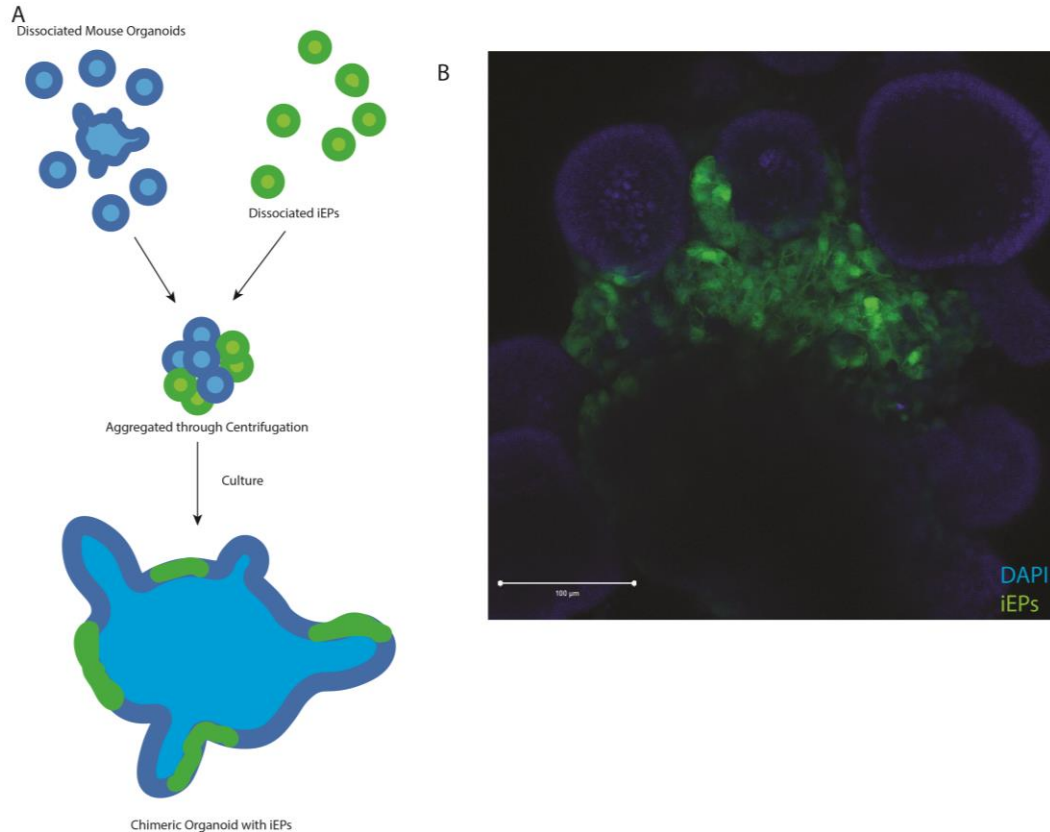
Figure 3.1: iEP and Organoid Chimerization. *(A)* Schematic of chimerization process using dissociated iEPs and mouse small intestinal organoids. The cells are aggregated via centrifugation and allowed to re-form into chimeric organoids. *(B)* Confocal image of an iEP-organoid chimera. Nuclei are stained with DAPI (blue), allowing visualization of the organoid budding structure. Integrated iEPs are visualized in GFP (green). White arrowhead denotes integrated iEPs. Scale bar = 100 µm. Image taken on the ZEISS LSM 880 II Airyscan FAST Confocal Microscope with a 20× objective.

Via Seurat-based tSNE analysis of the cell populations to dimensionally reduce the data, we found that 3 clusters were formed from our cell types: an intestinal cluster identified by the high proportion of control organoid cells (77.1% of cells in the cluster), a progenitor cluster identified by a high proportion of control iEPs (68.8% of cells in the cluster), and a "transition" cluster located between the two, built up of mixture of cells (15.1% organoid control, 42.5% chimeric iEP, and 28.3% iEP controls) (Figure 3.2).[102] Interestingly, we found that iEPs from the chimeras were present in all three clusters, indicating that some are similar to normal iEPs (25.7% of the progenitor cluster) while others are becoming more like organoid cell types

66
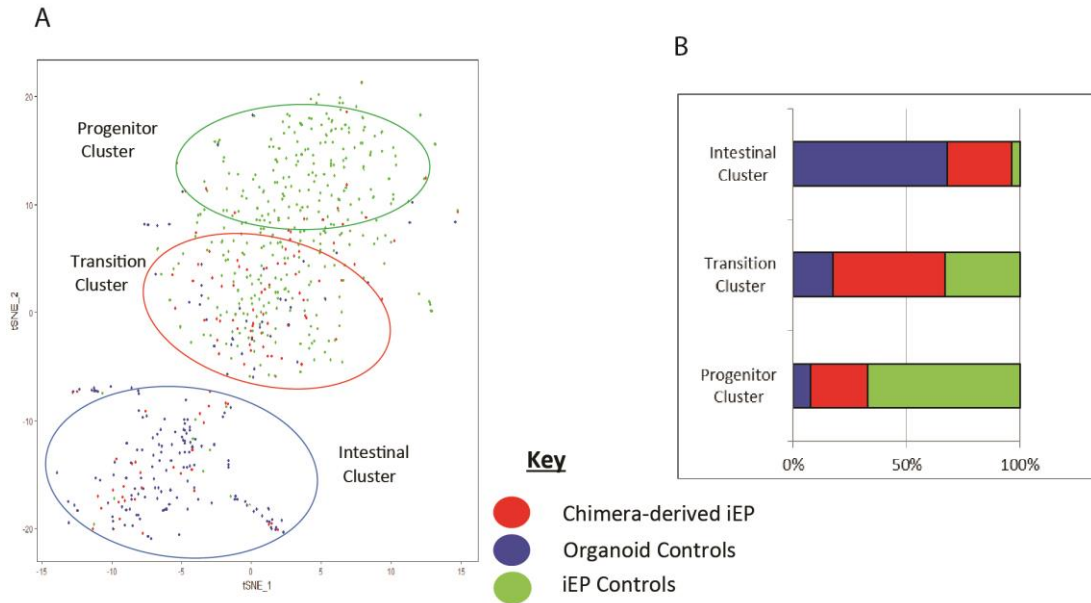
Figure 3.2: scRNA-seq analysis results of iEP chimeras. *(A)* tSNE plot representing the transcriptomic analysis of the iEP chimera dataset. Red dots represent iEPs derived from chimeras, blue dots represent mouse intestinal organoid cells, green dots represent control iEPs. Blue circle outlines intestinal cluster, red circle outlines transition cluster, green circle outlines progenitor cluster. *(B)* Breakdown of proportion of cell types in each cluster of *(A)*. Key of colors for *A* and *B*.

(31.9% of the intestinal cluster), and some are in a more intermediate, transitional state (42.5% of the transitional cluster) (Figure 3.2). Within the chimera-derived iEPs found in the intestinal organoid cluster, we found that some cells were expressing intestinal stem cell markers such as *Lgr5* and *Prom1*, indicating a potential transcriptomic change to a more intestinal stem cell-like identity in chimerized iEPs (data not shown).

Overall, this analysis provided preliminary evidence that iEPs can be influenced by the stem cell niche in intestinal organoids to take on more intestine-like identities. While the cells remained predominantly intestinal stem cell-like in this analysis, we hypothesized that extended culture time and alterations in culture conditions would lead to increased maturation. Even with the immaturity of the data, we could see that the use of the organoid environment to influence cells can induce reprogramming without the forced expression of exogenous factors, giving

credence to our use of *in vivo*-like environments to induce cellular identity changes with diffusible signals. Based on these results, we continued to refine and optimize the chimeric model for use in modeling of regional reprogramming.

3.4 Proximal-Distal Chimeric Organoids Produce Structurally Chimeric Organoids

Due to the preliminary success of the pilot study, we refined the chimera model to suit the needs of our proposed regional reprogramming model. This required the derivation of new organoid lines from distinct regions of the intestine (proximal/duodenum and distal/ileum) that could be differentiated from one another via scRNA-seq. To achieve this differentiation of starting regional identities via scRNA-seq, we derived proximal organoids from the duodenum of normal mice and distal organoids from the ileum of mT/mG mice, which express membrane-bound tdTomato in all cells, making the distal organoids fluoresce red and hopefully allowing their differentiation from the proximal organoid cells when utilized for scRNA-seq.[186,187] Proximal organoids were derived from the duodenum instead of the more medial jejunum to ensure proximal identity of the cells, as regions of the intestine during dissection have little delineation between jejunum and other regions.

The chimera aggregation process was adapted slightly from the pilot to encourage better association of the cells. We began by breaking the proximal and distal organoids down to small (2-5) cell clusters as before—to encourage the regrowth of organoids from stem cells.[49] The control (proximal or distal organoids broken down and individually reaggregated via centrifugation to control for effects of aggregation on organoids, creating single-identity controls) and experimental (proximal and distal organoid cells combined and chimerized through aggregation) populations were then created totaling 2000 cells/well of a 96-well U-bottom low-

bind plate supplemented with custom growth media to promote maturation and differentiation of the organoid cells. The movement from IntestiCult to a custom media blend allowed for modulation of differentiation factors such as Noggin and R-spondin 1 as opposed to the proprietary formulation of commercial IntestiCult Media (StemCell Technologies) which we found maintains highly stem-like populations. The cells were allowed to settle with the plates on ice for 10 minutes before centrifugation to aggregate. Finally, Matrigel was added to the wells for extracellular matrix support and the chimeras and controls allowed to grow for 5 days in an incubator prior to collection for downstream analysis (Figure 3.3).
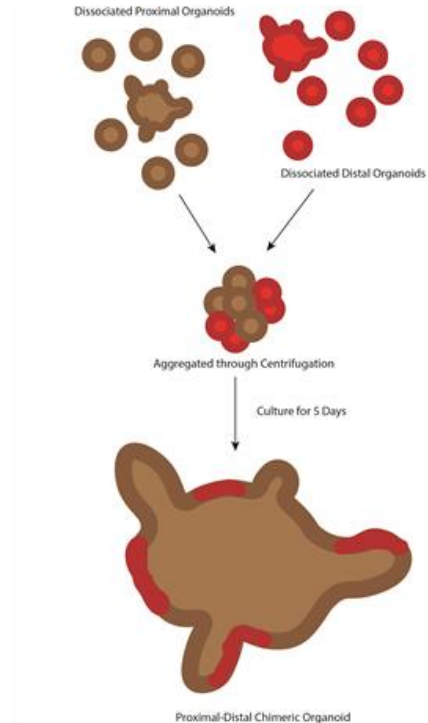
Figure 3.3: Overview of proximal-distal chimera process. Schematic of chimerization process using normal proximal mouse organoids and mT/mG distal mouse organoids. The cells are aggregated via centrifugation and allowed to re-form into chimeric organoids.

Initial analysis began with visual inspection of the preliminary chimeras and controls (organoids grown in IntestiCult). Brightfield imaging of the control and chimeric populations

qualitatively revealed similar growth patterns among chimeras and controls. While some variation existed between wells of each 96-well plate, all aggregates had re-formed visually normal organoids of relatively equivalent size and exhibiting many distinct crypt structures (Figure 3.4). This primary visual inspection indicated that both the aggregation and chimerization processes had limited impact on organoid growth and cellular proliferation, perhaps having no effect on structural adaptation characterized by expansion of cells (discussed in Chapter 2).
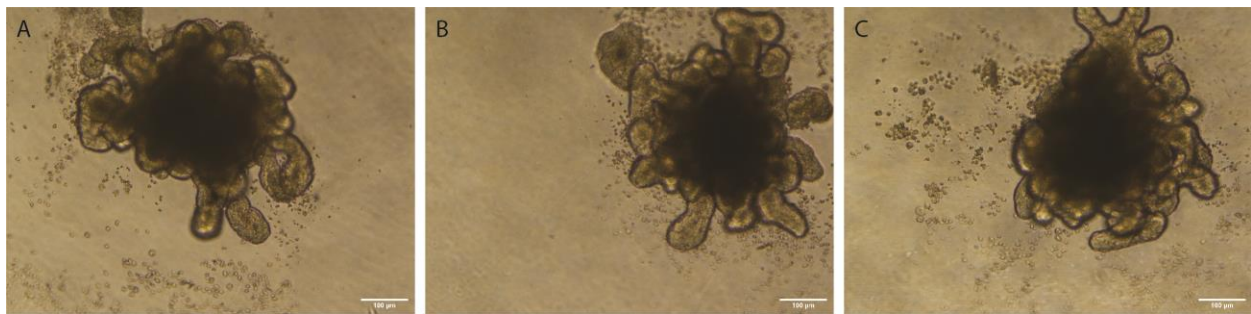
Figure 3.4: Bright field imaging of controls and proximal-distal chimeras. *(A)* Ileal-Ileal (distal) control. *(B)* Duodenal-Duodenal (proximal) control. *(C)* Chimeric (proximal-distal) experimental organoid. Images presented are representative of IntestiCult growth conditions and taken on a Nikon Eclipse Ts2 with Nikon DS-Fi2 camera with a 10× objective. Scale bar = 100 µm.

To gain better understanding of the structure and compositional patterning of chimeric organoids and to visualize the success of the chimerization process, confocal imaging was utilized to view the combination of normal proximal and mT/mG distal cells in chimeras grown in IntestiCult (Figure 3.5). The images display expected chimerization in which some regions (crypts) have been grown from an mT/mG positive stem cell derived from the distal organoids and are red and some regions (crypts) have been grown from an mT/mG negative stem cell derived from the proximal organoids. This interspersed nature of the regions is expected based on the growth patterns of organoids in which a single crypt is produced from an individual stem cell as cells differentiate and migrate outward. It would not be expected to have alternating

individual cells of differing colors due to this typical growth pattern. Instead, the regions of a single color indicate normal growth and differentiation patterns of the intestinal epithelium. These data indicate that chimerization results in regions of a single chimeric organoid being derived from different regions of the intestine, a reflection of the anastomosis of remnant distal and proximal intestine post resection. Imaging took place in IntestiCult prior to switch to custom growth media but reflects growth patterns of chimeras vs. organoids in identical conditions.

Based on our qualitative imaging investigation, we conclude that chimeric organoids composed of proximal and distal organoids display little change in growth or cellular proliferation due to the chimerization process. Structural imaging also indicates success of the chimerization process in combining regional organoids into a single structure, the chimeric organoid. With promising images, we moved forward to quantitative analysis of the chimeras using scRNA-seq.
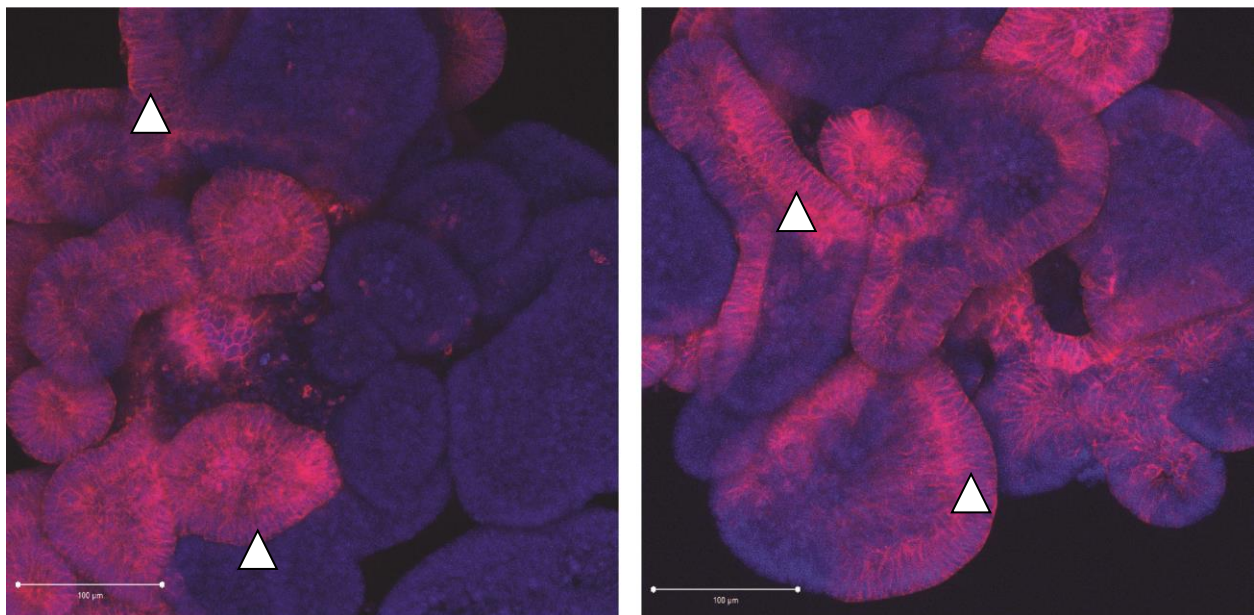
Figure 3.5: Confocal imaging displays regions of chimerization in organoids. Confocal imaging of mT/mG (distal) cells in red and nuclei in DAPI blue exhibits combined regions of red+ and red- cells in the chimeras, indicating successful chimerization of proximal and distal organoids. White arrow heads distinguish distally-derived crypt regions. Images presented are representative of IntestiCult growth conditions and taken on the ZEISS LSM 880 II Airyscan FAST Confocal Microscope with a 20× objective. Scale bar = 100 µm.

3.5 Distal Tissue in Chimeric Organoids Displays Significant Proximalization Scores via scRNA-seq

In two independent biological replicates, we used scRNA-seq to analyze the chimeric organoids and the control organoids in our custom growth media. The data was aligned and run through quality control measures before utilizing Seurat to perform dimensional reduction and investigate differential gene expression between populations.[102] Seurat analysis resulted in UMAP visualizations of Replicate 1 (R1) and Replicate 2 (R2) in which R1 consisted of 12 clusters and R2 consisted of 10 clusters (data not shown). R1 contained 2416 duodenal control cells, 1601 ileal control cells, and 343 chimeric experimental cells. R2 contained 3722 duodenal control cells, 3485 ileal control cells, and 2114 chimeric experimental cells. Because of difficulties in cell collection, demultiplexing, and subsetting, the following analysis was affected by low cell numbers and results are considered preliminary—future replications are needed to validate and increase confidence in conclusions.

Visualization of the UMAPs by original sample identity of the cells showed clear delineation between the control duodenal and control ileal populations (Figure 3.6*A*). The duodenal controls (green) are predominantly distinct from the ileal controls (blue), which indicates clear transcriptional differences between control populations. These differences validate the literature-reported maintenance of regional identity of organoids once derived.[3,7] The clear distinction also lends strength to the ability of this analysis to identify proximal vs. distal regional identities transcriptionally. In both replicates, the chimeric organoids (red) are distributed throughout the UMAP and fall into both control regions. This distribution is as expected due to the derivation of 50% of the chimeric cells from the ileal organoids and 50% of the chimeric cells from the duodenal organoids (Figure 3.6*A*). We expected the chimeras to fall

in both populations as the chimeric cells derived from the ileum are not likely to completely lose ileal identity but instead begin to take on markers of proximal or duodenal identity as well, ultimately resulting in a "hybrid" identity as suggested in Chapter 2.[183]
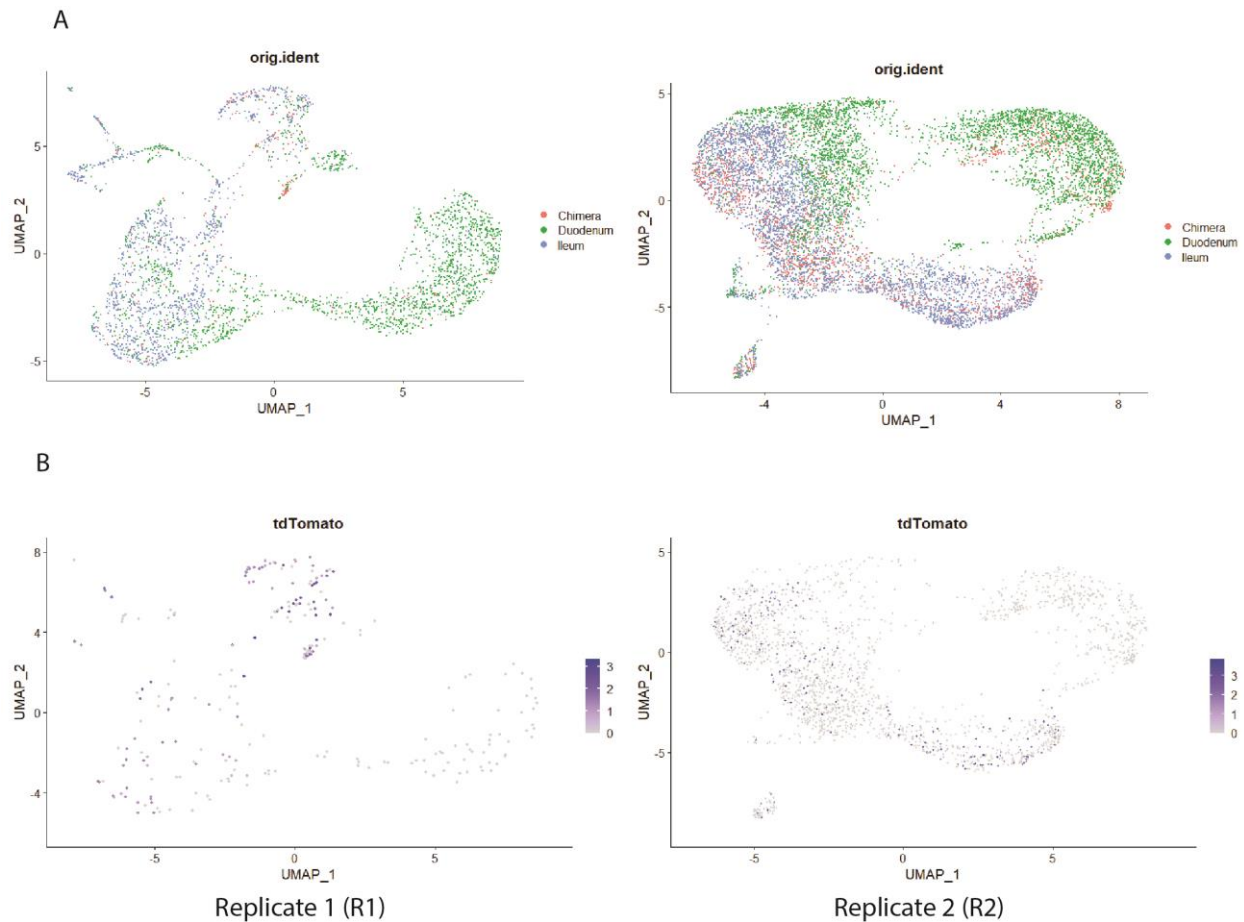
Figure 3.6: Single cell RNA-seq analysis visualizations of chimera experiments. *(A)* UMAP visualizations of Replicate 1 (left) and Replicate 2 (right) separated by original sample identity with key in center. *(B)* UMAP visualizations of tdTomato expression (purple) as detected by scRNA-seq in the chimeric sample subset in Replicate 1 (left) and Replicate 2 (right) denoting original identity of chimeric cells as from distal organoids (mT/mG positive).

To determine which chimeric cells came from which source (distal mT/mG cells or proximal normal cells) we subsetted out the chimeric sample computationally and utilized gene expression data to differentiate between the cells expressing tdTomato (distally derived) and

73

others (Figure 3.6*B*). By doing so, we were able to directly compare distally derived chimeric cells to the distal controls to look for differences in proximal enterocyte identity acquisition. Using Haber et al., 2017, we compiled a list of markers of proximal enterocyte identity (Table 3.1).[106] Seurat was utilized to generate a proximal identity score for each cell in the distal chimeric and distal control populations based on their respective expression of the genes in the list. Permutation testing was utilized to compare the mean proximal identity score of the control distal and chimeric distal populations to determine if the difference between proximal enterocyte identity scores was statistically significant. We found that in R1, distally derived chimeric cells had significantly greater proximal enterocyte identity scores than distal controls ($P = 0.01099$). Similarly, in R2, distally derived chimeric cells had significantly greater proximal enterocyte identity scores than distal controls ($P = 0.0091$). The significance of these statistics indicates that the chimera model induces acquisition of proximal enterocyte identity in distal organoid tissue based on scRNA-seq analysis.

Table 3.1: Proximal Enterocyte Gene List for Identity Scores

Casp6	Gata4	Apoa4	Fabp1	Apoc3	Rbp2	Apoc2
Leap2	Cyp2b10	Cyp3a11	Lct	Gsta1	Gstm1	Gstm3
Ephx2	Ms4a10	Fam213a	Cbr1	Adh6a	Cyb5r3	Dhrs1
Ifi27l2b	Cyb5a	Cyp3a25	Ckb	Prap1	Cgref1	Dnase1
Aldh1a1	Khk	Lpgat1	Treh	Reg3g	Acsl5	Ace
Aldob	H2-Q2	Rdh7	Ckmt1	Cyp3a13	P4hb	Mdh1
Ppap2a	Slc2a2	Cox7a1	Sec14l2	Gsta4	Mme	Retsat
Mttp	Creb3l3	Slc5a1	Sult1b1	Hsd17b6	Scp2	Cyb5b
Cyp2c65	Gpx4	Xdh	Cyp2d26	Ugdh	Gstm6	Ndufa1
Gpd1	Cyp2c66					

In seeking to validate our preliminary findings, we utilized Capybara, a quadratic programming approach, to measure the identity classifications of individual cells in the distally-derived chimera dataset and the distal controls of R2 (the larger dataset) using the Haber et al.,

2017 dataset as reference.[100,106] Based on this analysis, we find that a higher percentage of cells in the distally derived chimera dataset classify as proximal enterocytes (10.8%) compared to distal controls (5.9%) (Figure 3.7). This result provides support to our finding that proximal enterocyte identity scores increase in the chimeric organoid distal cells. Additionally, we noted that immature and mature enterocyte populations (distal and proximal combined) were expanded in the chimeric samples (27.1% vs. 13.3%). This finding reflects our finding of expanded enterocyte identity within the *in vivo* model after resection presented in Figure 2.3*B*. Together with the proximalization scores, these findings support the use of the chimeric organoid system to reflect cell identity and population changes due to regional reprogramming.
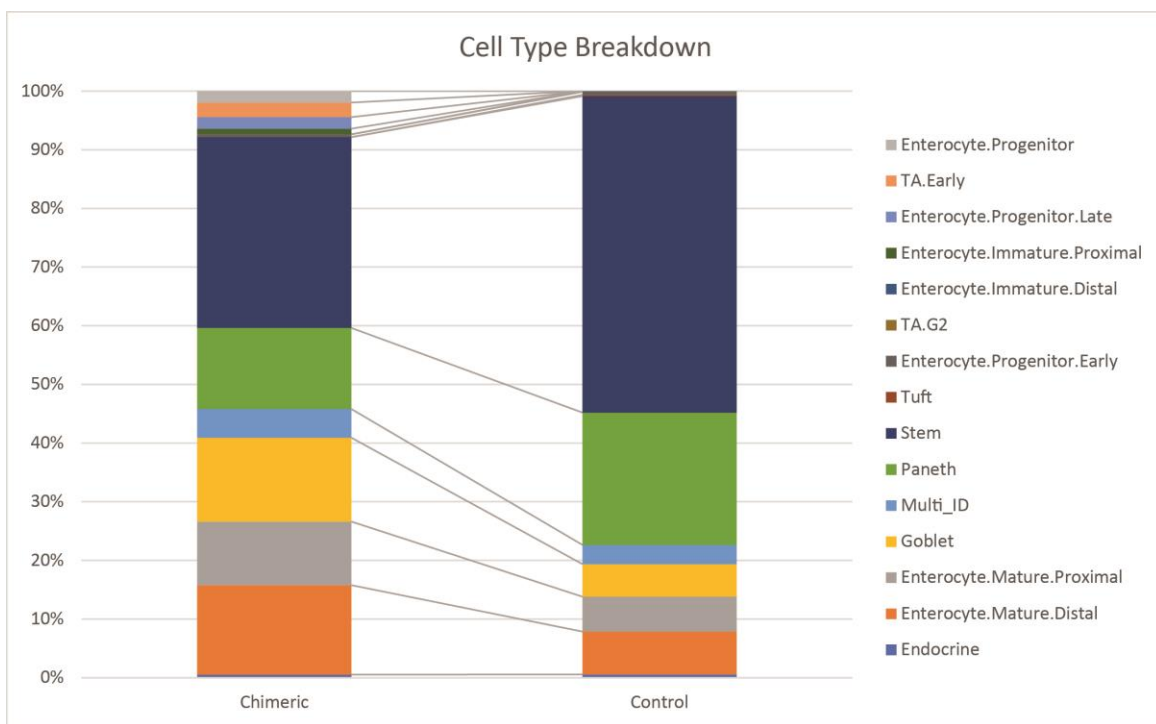
Figure 3.7: Capybara analysis of chimera experiment. Capybara classifications of cells in each sample as percent total cells. (No statistics available due to n = 1 and small sample sizes).

However, we must reiterate that these analyses were underpowered. While the distal control dataset contained 3485 cells, the distal chimera data only contained 203 cells due to issues with dropout resulting in distally-derived chimera cells going undetected (see Section 3.7).

Ultimately, sampling of 203 cells for analysis leaves this data preliminary and underpowered. However, we find that the increased proximalization scores and proximal enterocyte classifications are a promising sign that proximal identity is being acquired.

We conclude, based on these analyses, that the chimeric organoid model as developed in these studies provides a preliminary *in vitro* model of regional reprogramming using mouse tissue. scRNA-seq analysis indicates that exposure of the distal organoid tissue to proximal organoid tissue via chimerization significantly increases proximal enterocyte identity score expression in distal cells when compared to distal controls. We hypothesize that the acquisition of proximal identity is due to different exposure to diffusible factors or signals in the new chimeric environment, mimicking the exposure to signaling gradients disrupted by resection as discussed in Chapter 2.[183] Replication of this experiment is necessary to increase confidence in this analysis.

3.6 Computational Analysis Indicates Alterations in Guanylate Cyclase-C Signaling in Chimeras

To further dissect the mechanisms surrounding the potential regional reprogramming occurring in our chimera model, we turned to *in silico* analysis of the Replicate 2 chimeric scRNA-seq dataset. Replicate 2 was chosen for this purpose due to greater sample size for increased analytical confidence (though still underpowered). We explored potential changes in signaling pathways induced during chimerization compared to control organoids utilizing a computational tool called CellPhoneDB.[107] Briefly, CellPhoneDB generates "interaction scores" between populations in scRNA-seq datasets based on permutation testing of expression of ligands and receptors from a curated database. In the case of our dataset, we were able to measure interactions between the chimeric cells (proximally derived ←→ distally derived) and

compare them to baseline interaction levels in control organoids. A heatmap of overall interaction counts gave the first indication that signaling pathways had changed activity levels in distally derived chimera cells as the population displayed high levels of interaction with the other sample cell types (Figure 3.8*A*). When dissected, different combinations of sample cells showed signaling pathway interactions, of most interest, due to high significance and expression levels in chimeras, were the GUCY2C – GUCA2A and GUCY2C – GUCA2B interactions (Figure 3.8*B*).

GUCY2C is the receptor for the guanylate cyclase-C (GC-C) signaling pathway. GC-C activation leads to a cyclic guanosine monophosphate (cGMP) cascade in the intestine and plays roles in intestinal homeostasis, intestinal barrier integrity, and epithelial renewal among other functions, making it a logical and fascinating potential player in regional reprogramming.[188] GUCY2C is expressed along the length of the intestinal tract as a membrane bound receptor facing the intestinal lumen.[189] Of interest to regional reprogramming, *Gucy2c* is a downstream target gene of transcription factor *Cdx2*, a major player in the development of regional identity and patterning in the intestine.[188,190]

GUCY2C is known to have two endogenous peptide ligands: guanylin (*Guca2a*) and uroguanylin (*Guca2b*). Importantly, GUCA2A is predominantly expressed in a distal→proximal gradient in the intestine while GUCA2B is expressed in a proximal→distal gradient.[188,189,191] These gradients reportedly are due to pH sensitivity of the ligands and the innate pH gradient of the small intestine (acidic in proximal, neutral in distal).[189,191] Both ligands are reported to be released from the epithelium due to changes in luminal stimuli such as intestinal tonicity—a variable that would be expected to change post-resection as transporters recalibrate and intestinal volume changes.[192] Based on the known role of GC-C signaling in the intestine and the

77
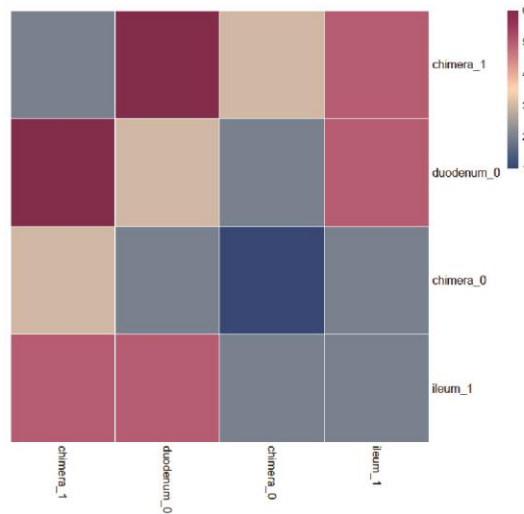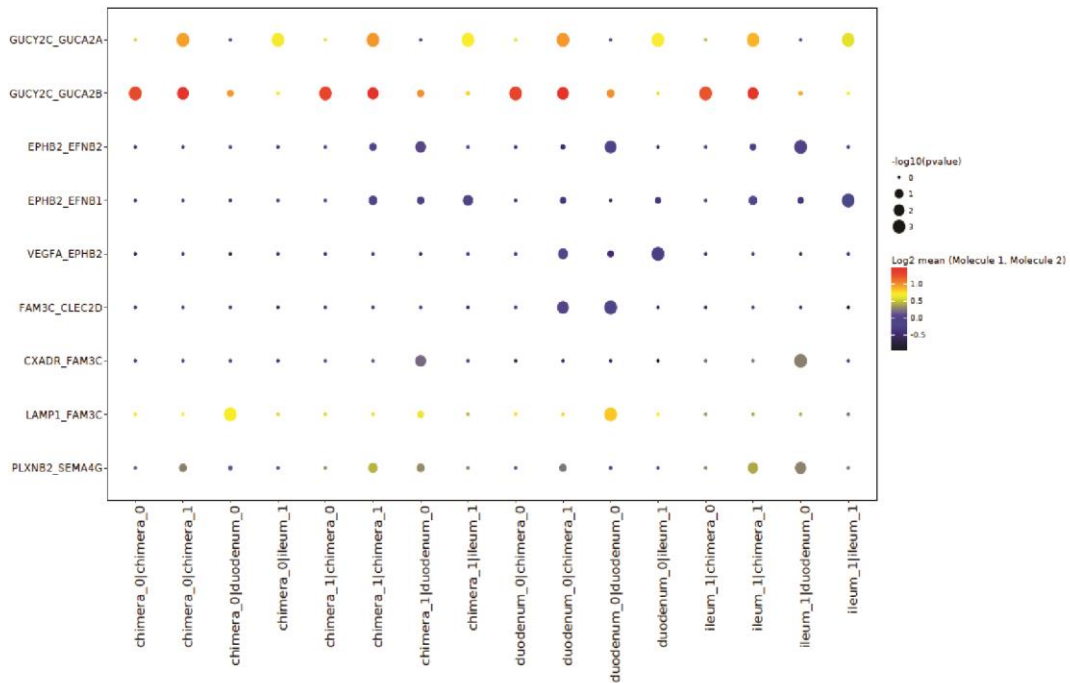
Figure 3.8: CellPhoneDB analysis identifies GUCY2C signaling changes in chimeras. *(A)* Heatmap of total interaction counts by sample type. *(B)* Plot of significant interactions; x-axis denotes interacting samples with first listed expressing the receptor and second expressing the ligand, y-axis denotes interacting pair receptor_ligand; *P* values are denoted by dot size, colors indicate means of average expression values of interacting molecules. Samples: chimera_1 = distally derived chimeric cells, chimera_0 = proximally derived chimeric cells, duodenum_0 = duodenal proximal controls, ileum_1 = ileal distal controls.

distinct patterns of ligand expression by region, we further explored the interaction changes with CellPhoneDB.[107]

First, looking at the interaction of GUCY2C and distal ligand GUCA2A, we noted that the interactions of GUCY2C and GUCA2A were only considered significant in samples containing ileal derived cells (chimera_1 and ileum_1) (Figure 3.9*A*). This was unsurprising as *Guca2a* is known to be predominantly produced in the distal intestine and would likely come from these cells. More interesting, perhaps, is that the mean interaction value between GUCY2C and GUCA2A was increased in samples in which the distally derived chimera cells expressed the ligand. We postulate that this increased value is due to an overall increase in *Guca2a* ligand expression in distally derived chimera cells (Figure 3.9*C*).

Looking at the interaction of GUCY2C with proximal ligand GUCA2B, we were surprised to see that GUCY2C and GUCA2B were only significantly interacting in comparisons containing chimeric samples. Additionally, the interaction value was increased in samples in which the distally derived chimera cells (chimera_1) expressed the ligand (Figure 3.9*B*). This data indicates, again, that an overall increase in *Guca2b* ligand expression, a proximal ligand, in the distally derived chimeras may be driving the significant interactions (Figure 3.9*C*). This is especially interesting as it indicates that chimerization (regional reprogramming) increases expression of a proximal ligand in distally derived cells. The increased expression of *Guca2b* in the distal chimeric cells represents a response to chimerization resulting in increased production of proximal signals.

Based on our findings in CellPhoneDB, we conclude that, via *in silico* analysis, guanylate cyclase-C signaling is altered in and playing a role in the changes associated with the preliminary *in vitro* chimeric model. Based on literature reports, GC-C signaling is present in and can be

79
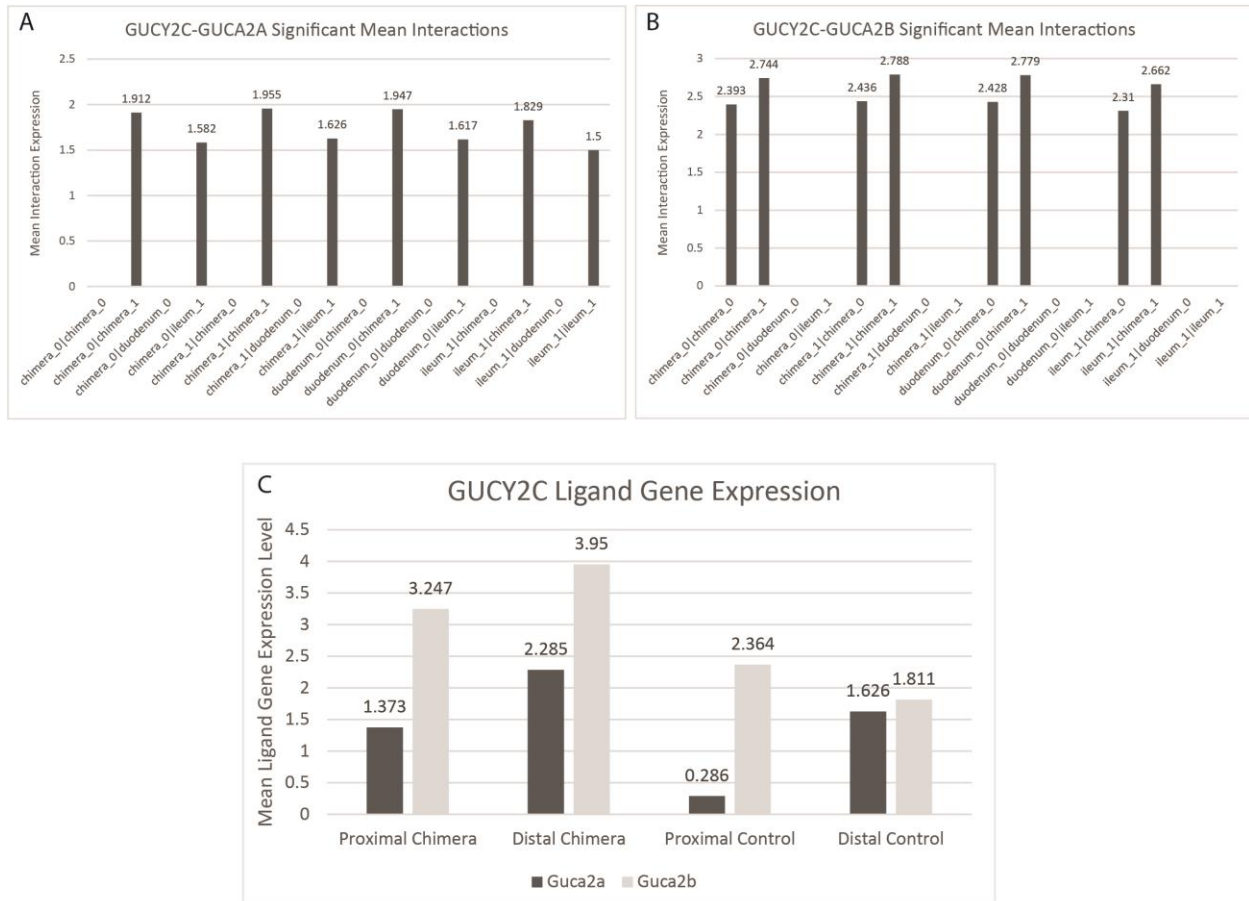
Figure 3.9: GC-C signaling and expression changes in chimeras. *(A)* Mean significant interactions of GUCY2C and GUCA2A, means shown fit $P < 0.05$. *(B)* Mean significant interactions of GUCY2C and GUCA2B means shown fit $P < 0.05$. x-axis in *(A)* and *(B)* denotes interacting samples with first listed expressing the receptor and second expressing the ligand, samples: chimera_1 = distally derived chimeric cells, chimera_0 = proximally derived chimeric cells, duodenum_0 = duodenal proximal controls, ileum_1 = ileal distal controls. *(C)* Mean expression of each GUCY2C ligand gene by sample. (No statistics available for *(C)* due to n = 1 and small sample sizes).

stimulated in *in vitro* organoids, providing support to our *in silico* findings.[193] In particular, we note that the exposure of the distally derived chimeric cells to the proximally derived cells upregulates expression of both the *Guca2a* (distal) and *Guca2b* (proximal) ligands in the distally-derived populations (Figure 3.9*C*). These changes indicate definite changes in regionally-regulated gene expression upon exposure to a proximal signaling environment in the chimera. Our position that the novel exposure to a region-specific internal environment in this

chimeric organoid (and the *in vivo* model) is driving regional reprogramming is supported by the fact that these ligands are known to release based on exposure to luminal (internal) environmental stimuli.[192] It is reasonable to conclude that alterations in the internal environment due to increased exposure to proximal signaling stimuli alters the expression of and release of proximal and distal ligands in our model system and likely in the *in vivo* system as well.

3.7 Discussion

In this chapter, we report on the development of a preliminary *in vitro* model of regional reprogramming we deem the "chimeric organoid" or "chimera" and the consequent downstream analysis of identity changes in cells within the chimeric organoid using scRNA-seq. We show in both a pilot and core experiment that aggregation of cells from two sources (organoids and/or cultured 2-D cells) induces changes in cellular transcriptional identity detectable by scRNA-seq. Because the distal cells in the chimeric organoid take on higher proximal enterocyte identity scores and cell classifications indicative of regional reprogramming, we propose further use and development of the chimeric organoid model to validate findings and for study of adaptation signals after SBR in mice and treatment applications in humans. Due to further computational analysis using CellPhoneDB highlighting changes in the GC-C signaling pathway in chimeric organoids, we propose that diffusible signals in the internal environment of the chimeric organoid play a role in regional reprogramming, a proposal supported by our findings regarding Retinoic Acid/Vitamin A in Chapter 2.[107,183]

While others have explored chimerization of cell types to produce organoids to mature *in vitro*, our model is novel due to the completely *in vitro* maturation cycle which relies on the innate structure and function of intestinal organoids.[68,69] Our model is stand-alone and offers

independence from traditional animal models. These findings offer significant progress for the future study of adaptation post bowel resection through isolation of an *in vitro* model which can easily be adapted to patient samples simply through using patient-derived organoids. In addition, this study gives insight into how findings from scRNA-seq studies can be used to provide insight into disease response pathology. Similar to the findings in Chapter 2, we noted that the distally-derived cells in the chimeric populations did not shift completely to non-ileal regions of the UMAP visualization, but still exhibited greater proximal identity scores (Figure 3.6*A*). These results support our previous finding suggesting a "hybrid proximal-distal" identity of the ileum after resection in which enterocytes gain proximal functions without fully losing their distal functions.[183]

In our pilot study utilizing chimerization of mouse organoids and iEPs, we found that exposure of iEPs to the signaling environment within the chimeric organoid system induced transcriptional shifts in some cells moving them closer to intestinal identities (Figure 3.2). Consistent with our previous reports regarding the adoption of intestinal identity in iEPs, we found that some chimerized iEPs began expressing intestine-specific stem cell markers such as *Lgr5* and *Prom1*, indicating not only their ability to reprogram to intestinal cell types but also the ability of the chimeric system to induce identity changes directing cells toward intestinal fates.[77,184] We conclude from this that intestinal organoids contain strong niche signaling to promote intestinal identity and iEPs are a progenitor with multiple endodermal fate choices. The continued investigation of iEPs as they differentiate down their varied fate pathways (liver and intestine) can provide insight into multiple reprogramming and differentiation processes both by engraftment *in vivo* and through chimerization models *in vitro*.

The imaging, especially brightfield, brings an interesting qualitative observation to light. We noted that the resulting organoids (control and experimental) were similar in size and shape of crypt-like structures (Figure 3.4). This observation indicates that the well-known structural adaptation (characterized by lengthening of villi and cellular proliferation) accompanying regional reprogramming may not occur in the chimeric *in vitro* model.[31] If structural adaptation occurs in the chimeric model, we expect to see changes in the length of the crypt-like outcroppings and overall size of the organoids due to increased cellular proliferation. While our initial qualitative image analysis indicates, upon visual inspection of samples, that the overall size of organoids and crypts appear no different in chimeric organoids, a more quantitative analysis should be carried out to truly make a conclusion. To determine if the chimera undergoes structural adaptation, in-depth image analysis in optimized media should be carried out to measure crypts and perform statistics. In addition, proliferation and differentiation could be investigated through labeling of the stem cells and tracking progeny differentiation and turnover.

Utilization of scRNA-seq to investigate gene expression differences between distal controls and distally-derived chimera cells showed not only increased proximalization of the chimera cells via proximal enterocyte identity scores, but also indicated changed proximal identities based on differential gene expression between the samples in Replicate 2. Three highly differentially expressed genes between sham and SBR populations in Chapter 2 were also expressed at significantly higher levels in distally-derived chimera cells when compared to distal controls.[183] These genes, *Fabp1* (avg_log2FC = 2.01)*, Fapb2* (avg_log2FC = 0.86)*,* and *Rbp2* (avg_log2FC = 0.81) are all associated with proximal intestinal nutrient or lipid transport and are upregulated in the distally-derived chimera cells with $P < 0.05$. These results support the finding

of increased proximal enterocyte identity scores by providing a more population-level significant difference in proximal nutrient transporter expression levels.

Additionally, use of an alternative analytic tool, Capybara, to classify cells as distinct identities, identified the chimeric-derived distal sample as having a higher incidence of mature proximal enterocytes and an expansion of total enterocyte populations.[100] These findings support our previous work and provide evidence for the success of the chimeric model. Use of Capybara also provided us with information regarding the maturity of our organoids due to media formulation alterations. In initial 10x experiments (data not shown), cells cultured in IntestiCult (a commercially available organoid growth media) showed a nearly ubiquitous stem cell identity and failure to mature and differentiate, this finding was confirmed by a representative of StemCell Technologies. Because of this, we changed from using IntestiCult to grow cells for the 10x experiments to using a custom media formulation with controlled amounts of Noggin, R-spondin 1, and EGF—factors known to promote maturity. We chose a formulation that indicated, via qPCR, that organoid maturity was increasing (*Alpi* expression) and stemness was decreasing (*Lgr5* expression) but cells still proliferated well (consistent growth) (data not shown). While we note that there is the presence of differentiated cells in the chimeric and control data, stem cells make up 32.5% of the distally derived chimeric cells and 53.9% of the distal controls. This abundance of stem cells may indicate that future media formulation tweaks are required to more fully mature the organoids.[194] Cellular composition can be further validated through staining.

Our use of scRNA-seq allowed us to identify GC-C signaling as being active during regional reprogramming. GC-C signaling provides an interesting avenue for future study as intestinal GC-C is not only implicated in homeostasis, but also in inflammatory intestinal diseases like IBD and colitis. Interestingly, both ligands (GUCA2A and GUCA2B) and receptor

(GUCY2C) have been shown to be down-regulated in IBD and it is proposed that this down-regulation is to promote tissue regeneration through proliferation.[188] As it seems that the production of ligands is increased in chimeras, we hypothesize that in order to prevent mere tissue hyperplasia and to promote expansion of only certain differentiated cell types, in our case, enterocytes, ligand production increases to discourage proliferation.[183,189] GC-C signaling is also downstream of Retinoic Acid signaling, as increased RA signaling increases CDX2 protein expression in some systems.[195] *Cdx2* is known to regulate *Gucy2c* expression in the intestine, so it is possible that an increase in RA signaling due to resection (as discussed in Chapter 2) increases *Cdx2* expression and thereby alters GC-C signaling after resection.[183,189] Future studies should investigate changes in GC-C signaling *in vivo* after resection to validate these hypotheses.

A shortcoming of this scRNA-seq study of the chimeras relates to the small sample sizes achieved in the datasets, especially in distally-derived chimeric cells, as mentioned when discussing the underpowered nature of the analysis. This shortcoming comes from unexpected difficulties in detecting tdTomato mRNA in the datasets, likely due to the structure of the tdTomato mRNA. This resulted in decreased isolation of distally-derived chimeric cells from the greater chimeric population due to drop-out of tdTomato reads. Methodologies to increase the power of this study and avoid this shortcoming in the future will be discussed in Chapter 5.

Overall, the chimeric model provides support to the proposition of Chapter 2, that diffusible signals whose gradients are altered by resection, such as RA, play a role in triggering regional reprogramming. Our chimeric organoid model is a preliminary *in vitro* system in which proximal signaling from proximally-derived organoid tissue influences distally-derived organoid tissue. Further study utilizing this model on mouse tissue will provide additional insights into signaling associated with regional reprogramming and should look into the possibility of

structural adaptation in the model as well. Based on our analysis, GC-C signaling is changing in the chimeric model. As GC-C signaling plays large roles in intestinal response to diseases and homeostasis, this pathway deserves further investigation relating to SBR and adaptation. *In vivo* studies looking for GC-C alterations in mice will validate these findings of a new pathway involved in the adaptive response. By further developing and investigating the mouse chimeric model, we can gain further insight into successful adaptation triggers and signaling for applications in humans. The mouse chimeric model should be applied to human organoids in subsequent studies. By comparing the chimeric model and regional reprogramming occurring in mouse chimeras vs. human chimeras, we can glean more information regarding differences in the successful adaptation process between the species.

3.8 Conclusions

Here, we have developed a novel model of mouse regional reprogramming using *in vitro* organoids and analyzed cellular identity changes and signaling pathway alterations using scRNA-seq. Our analysis of the chimeric organoid model reveals that distally-derived organoid cells are capable of taking on more proximal enterocyte-like identities, indicative of regional reprogramming, when exposed to a proximally-derived organoid environment. Pathways like GC-C signaling are implicated in the chimeric setting, and reveal the importance of changed availability of diffusible signals in the intestinal environment during regional reprogramming. This model should be further developed for applications to human organoid systems and further studied in mice in order to dissect the signaling pathways activated in regional reprogramming to better understand the adaptive response after SBR.

3.9 Materials and Methods

3.9.1 Mice

Mice utilized for organoid derivation were of C57BL/6J background (The Jackson Laboratory #000664). Mice utilized for mT/mG organoid derivation were of C57BL/6J background with homozygous mutations for mT/mG (The Jackson Laboratory #007676).[186] All animal care procedures were approved by the Washington University Institutional Animal Care and Use Committee.

3.9.2 Organoid Derivation

Organoids from mouse small intestine were isolated and derived as intestinal crypts as described in the StemCell Technologies Technical Bulletin.[187] Modifications for regional isolation were as follows: tissue for derivation of proximal organoids (duodenal) was isolated as the ~5 cm of duodenum just distal to the pyloric sphincter; tissue for derivation of distal organoids (ileal) was isolated as the ~5 cm of ileum just proximal to the ileocecal valve. Briefly: tissue was collected and placed in cold PBS, tissue was flushed and minced. Minced tissue was agitated and rinsed 20 times and supernatant filtered through a 70 µM filter to isolate intestinal crypts for culture.

3.9.3 Cell Culture

iEPs were generated as previously described.[78,79] Briefly, MEFs (from E13.5 C57BL/6J mouse embryos, The Jackson Laboratory #000664) were cultured on gelatin and serially transduced every 12h with *Hnf4a*-t2a-*Foxa1* retrovirus 5 times over 3 days. The reprogrammed cells were then cultured on collagen in HepBase Media consisting of: DMEM:F-12 (Gibco) supplemented with 10% FBS, 1% penicillin/streptomycin, 55 µM 2-mercaptoethanol, 10 mM Nicotinamide (Sigma-Aldrich), 100 nM dexamethasone (Sigma-Aldrich), 1 µg/mL insulin

(Sigma-Aldrich), and 20 ng/mL epidermal growth factor (Sigma-Aldrich). Cells were maintained at 37°C, 5% CO_2 in an incubator.

Mouse organoids were maintained as previously described in the StemCell Technologies Technical Bulletin with modifications.[187] Briefly, organoids were plated in 48-well cell culture plates (VWR 10062-898) in 20 µL droplets consisting of 50% Mouse IntestiCult Organoid Growth Media (StemCell Technologies #06005) and 50% Growth-Factor Reduced Matrigel (Corning 354230). Droplets were polymerized for 10 minutes at 37°C, 5% CO_2 in an incubator before 250 µL of Mouse IntestiCult Organoid Growth Media was added to each well. Organoids were maintained at 37°C, 5% CO_2 in an incubator. Organoids were passaged as described in the StemCell Technologies Technical Bulletin.[187] Briefly, Matrigel domes were broken down using Gentle Cell Dissociation Reagent (StemCell Technologies #07174) and gentle rocking at room temperature for 10 minutes. Cells were washed and centrifuged 2x before replating in 20 µL droplets consisting of 50% Mouse IntestiCult Organoid Growth Media (StemCell Technologies #06005) and 50% Growth-Factor Reduced Matrigel (Corning 354230) in 250 µL of Mouse IntestiCult Organoid Growth Media.

3.9.4 Chimerization

Chimeras were formed using an adapted protocol based on aspects of Xinaris et al., 2012, Xinaris et al., 2016, and Sato et al., 2011.[49,68,69] Normal proximal and mT/mG distal organoids were each collected and split to small clusters of cells using Gentle Cell Dissociation Reagent (StemCell Technologies #07174) for 10 minutes on a rocker at room temperature. Cells were counted and placed in 100 µL of growth media in 96-well low-bind U-bottom plates (Greiner Bio-One 650970). Samples consisted of controls—2000 cells of a single type in a well; experimental—1000 cells of each type in a well. Wells were gently pipetted 10x to mix and

plates placed on ice for 10 minutes to allow organoids to settle. Plates were carefully centrifuged for 5 minutes at 300xg, 4°C. Finally, 10 µL of Matrigel (Corning 354230) was added to each well before the plates were allowed to grow for 5 days at 37°C, 5% CO_2 in an incubator. Growth media consists of 21.75 µL basal media, 0.5 mL B27 (Gibco 17504044), 250 µL N-2 (Gibco 17502048), 65 µL 500 mM N-acetylcysteine (Sigma-Aldrich A9165), 12.5 µL 100 µg/mL EGF, 25 µL 100 µg/mL stock mouse Noggin (PeproTech 250-38), 25 µL 100 µg/mL stock R-spondin 1 (R&D Systems 7150-RS0010/CF). Basal media consists of 48.5 mL Advanced DMEM:F12 (Gibco 12634010), 0.5 mL GlutaMax 100x (Gibco 35050061), 0.5 mL HEPES 1M (Gibco 15630106), 0.5 mL penicillin/streptomycin.

3.9.5 scRNA-seq Cell Collection

Controls and chimeras for scRNA-seq were collected and prepared as follows. Each control or experimental sample was pipetted and combined in a 15 mL tube. Samples were centrifuged at 500xg for 3 minutes. Media was carefully aspirated and 3 mL TrypLE Express (Gibco 12605036) was added. Organoids were agitated through pipetting up and down 20-30x with a p200 pipette and placed in a 37°C bath. Every 5 minutes, cells were agitated via pipetting 20-30x and checked for breakdown to single cells. When cells achieved single-cell state, 5 mL of DMEM:F12 (Gibco) was added, and the cells were filtered through a 40 µm FlowMi strainer (Bel-Art H136800040). Cells were counted and viability checked using Trypan Blue exclusion (viability > 90% required). Final samples were centrifuged at 500xg for 3 minutes at 4°C. Replicate 1 cells were used fresh while Replicate 2 cells were subsequently methanol fixed as described in Alles et al., 2017 by resuspending cells in 200 µL PBS and adding 800 µL ice cold 100% methanol dropwise, then stored at -80°C until use.[177]

3.9.6 scRNA-seq Library Preparation

For scRNA-seq on iEP chimeras using Drop-seq, Drop-seq was performed as previously described (http://mccarrolllab.com/dropseq/).[185] In brief, cells and beads were diluted to an estimated co-occupancy rate of 5% upon co-encapsulation: 1×10^5 cells/ml and 1.2×10^5 beads/mL. Emulsions were collected and broken using 1 mL of Perfluorooctanol (Sigma) for 15 ml of emulsion, followed by washing in $6\times$ saline-sodium citrate (SSC) buffer to recover beads. Reverse transcription was then performed using the Maxima H Minus Reverse Transcriptase kit (Life Tech EP0752). After treatment with 2,000 U/mL of ExonucleaseI (New England Biolabs), aliquots of beads were amplified by PCR using Kapa HiFi Hotstart Readymix (Kapa Biosystems). The PCR product resulting from this reaction was purified by addition of $0.6\times$ AMPure XP beads (Beckman Coulter). Purified cDNA product was tagmented using Nextera XT according to the manufacturer's protocol (Illumina). The resulting cDNA library was again purified using $0.6\times$ AMPure XP beads, followed by $1\times$ AMPure XP beads. cDNA concentrations were assessed by Agilent Tapestation. Libraries were sequenced on an Illumina HiSeq 2500.

For single-cell library preparation of Replicate 1 on the 10x Genomics Chromium platform, we used the 3'CellPlex Kit Set A (PN-1000261), Chromium Next GEM Single Cell 3' Kit v3.1 (PN-1000268), 3' Feature Barcode Kit (PN-1000262), Chromium Next GEM Chip G Single Cell Kit (PN-1000120), and Dual Index Kit NN, Set A (PN-1000243), according to the manufacturer's instructions in the Chromium Next GEM Single Cell 3' v3.1 Cell Surface Protein User Guide (Rev. B). Replicate 2 was prepared as follows: methanol-fixed cells were placed on ice, then spun at 3000 rpm for 5 minutes at 4°C, followed by resuspension and rehydration in PBS, as previously described.[177] Then, on the 10x Genomics Chromium platform, we used the

Chromium Next GEM Single Cell 3' Kit v3.1 (PN-1000268), Chromium Next GEM Chip G Single Cell Kit (PN-1000120), and Dual Index Kit TT, Set A (PN-1000215), according to the manufacturer's instructions in the Chromium Next GEM Single Cell 3' v3.1 Gene Expression Dual Index User Guide (Rev. B). All resulting cDNA libraries were quantified on an Agilent Tapestation and sequenced on an Illumina NextSeq500.

3.9.7 scRNA-seq Analysis

The Cell Ranger 6.0 pipeline was used to align 10x reads to a custom mm10 genome build containing *tdTomato*, demultiplex at a confidence level of 0.75 for multiplexed R1, and generate a digital gene expression (DGE) matrix: (https://support.10xgenomics.com/single-cell-gene-expression/software/downloads/latest). For initial filtering of all DGE matrices, we first excluded cells with a low number (<200) of unique detected genes. We then excluded cells for which the total number of unique molecules (UMIs) (after $\log_{10}$ transformation) was not within 3 standard deviations of the mean. This was followed by the exclusion of outlying cells with an unusually high or low number of UMIs/genes given their number of reads by fitting a Loess curve (span = 0.5, degree = 2) to the number of UMIs/genes with number of reads as predictor (after $\log_{10}$ transformation), removing cells with a residual more than 3 standard deviations the mean. Finally, we excluded cells in which the proportion of the UMI count attributable to mitochondrial genes was >20%. After filtering and normalization of the DGE, the R package Seurat (Version 4) was used to cluster and analyze the single-cell transcriptomes.[102] Highly variable genes were identified and used as input for dimensionality followed by clustering and visualization via UMAP.[103] Module scores for proximal identity detection were calculated using Seurat function *AddModuleScore*.[102] R package Capybara was used to classify cells.[100] CellPhoneDB was used on raw data as described in Efremova et al., 2020.[107]

3.9.8 Imaging

Bright field images were taken on a Nikon Eclipse Ts2 with Nikon DS-Fi2 camera. For confocal imaging, chimeric organoids were collected for imaging by gentle pipetting. Organoids were fixed for 20 minutes in 4% paraformaldehyde and stained for 30 seconds in 300 nM DAPI in PBS. Organoids were subsequently embedded in 0.75% agarose in a 96-well glass bottom plate for imaging. Confocal imaging took place on a ZEISS LSM 880 II Airyscan FAST Confocal Microscope using ZEISS Zen Software (ZEISS Group).

3.9.9 Statistical Analysis via Permutation Testing

To determine whether distally-derived chimeric cells had higher proximal enterocyte identities than distal controls, we used permutation testing (also called randomized testing) to evaluate whether the mean proximal enterocyte identity score of the experimental population was greater than that of the control relative to a randomly selected null distribution of the data. Here, the proximal enterocyte identity score is as defined by Seurat using the *AddModuleScore* function.[102] The methodology is adapted from Biddy et al., 2018.[79] Let N represent the number of cells in the experimental population and M represent the number of cells in the distal control population. We pool the two groups of cells (size $= N + M$) and resample N random cells, without replacement, from the pooled cells 100000 times such that many possible separations with ending groups of size N can be sampled and captured. During this process, the mean is calculated based on the N randomly sampled cells. With the mean calculated, P values can be evaluated based on the proportion of randomly sampled cells with a mean greater than or equal to the null percentage. Using the P value of <0.05, we evaluated proximal identity acquisition differences. The calculations were performed using a custom R-based script.

3.10 Acknowledgements

3.11 Authors' Contributions

Sarah Elizabeth Waye: study concept and design, data acquisition, analysis, interpretation, figure preparation; Samantha A. Morris: study concept and design, obtained funding, provided overall study supervision.

Chapter 4: Effects of Direct Reprogramming on Modeling Intestinal Regional Reprogramming in Organoids

4.1   Abstract

Utilization of an *in vitro* model of disease alleviates many issues surrounding the use of animal models from high cost of maintenance and experimentation to limitations in animal number. In the case of modeling SBS and the resulting regional reprogramming, while the mouse model has provided valuable information regarding changes in the intestinal epithelium post resection (see Chapter 2), and our novel chimeric organoid system provides insight into endogenous signaling and new opportunities to discover signals playing a role in the process (Chapter 3), we desire a model in which we can test specific signals for their effects on regional reprogramming. In this chapter, we explore the use of upstream transcription factors in inducing regional reprogramming via direct reprogramming of organoids in order to study the effects of specific, known signals of interest on the activation of regional reprogramming. We identify three candidate transcription factors, *Creb3l3, Klf4,* and *Gata4,* for inducing regional reprogramming via direct reprogramming. We generate three overexpression organoid lines to study the effects of these transcription factors on induction of regional reprogramming. Focusing our efforts on the effects of *Gata4,* we find that overexpression of *Gata4* alone is not sufficient to induce regional reprogramming in distal organoids based on scRNA-seq data and computational analysis. However, we do find that overexpression of *Gata4* in distal organoids induces changes in the balance of cell types expressed in the organoids toward more stem-like identities. We conclude that overexpression of *Gata4* alone is insufficient to induce regional reprogramming *in*

vitro but based on noted changes in the epithelial composition after *Gata4* overexpression, hypothesize that inclusion of *Gata4* in a cocktail of reprogramming factors will likely produce the desired regional reprogramming *in vitro* response.

4.2 Introduction

As discussed in Chapter 3, the growing importance of personalized medicine necessitates a shift in focus and experimental modeling in stem cell biology and regenerative medicine. Animal models have provided the field with invaluable opportunities to study disease and responses *in vivo,* however the drive to scale up and focus on human biology requires new, *in vitro,* models to be built. In shifting our focus to *in vitro* modeling, we must apply what we know about disease from our animal-based systems in a dish to recapitulate an *in vivo* process for further study and new applications.

In the study of intestinal disease, an *in vitro* model of the small intestinal epithelium exists as the intestinal organoid (for further discussion of organoids, see Subsection 1.1.2). Intestinal organoids provide a valuable system for study as they can be easily derived from animals or humans from a simple intestinal biopsy, making personalized patient organoids easily attainable.[196] Additionally, intestinal organoids maintain much of the biology of the *in vivo* intestine in their genetics, cellular composition, and even functional signaling and responses to signaling.[1,7,43,47,48] A large body of literature provides evidence supporting the use of intestinal organoids as models and drives us to utilize intestinal organoid systems to model regional reprogramming after SBR. However, a large issue in the use of organoids to model regional reprogramming *in vitro* is identifying upstream components responsible for the process and activating those signals in the organoids.

In Chapter 3, we built the preliminary chimeric organoid system which allowed us to utilize endogenous signaling within the organoid environment to induce regional reprogramming effects in distal organoids. This methodology didn't require previous knowledge of the specific signals at play and allowed us to utilize the resulting data to later identify GC-C signaling as being involved. While this model is useful for the study of regional reprogramming and discovery of new pathways, it does not allow for the specific testing of factors of interest for their involvement in regional reprogramming.

During our study in Chapter 2, we utilized scRNA-seq to identify regional reprogramming in the distal remnant intestine after resection. We can identify regional reprogramming through increased expression of proximal-related tissue markers in the distal intestine post-adaptation. In addition to Vitamin A metabolism and diffusible signals upstream of regional reprogramming (Chapter 3), our study identified an increase in proximal transcription factors *Creb3l3* and *Klf4* during and after adaptation.[183] We hypothesized that *Creb3l3* (a putative downstream target of RA) was an upstream transcription factor driving and maintaining the regional reprogramming response after resection. Unfortunately, in the *in vivo* system, it is difficult to study the effects of a single transcription factor on a tissue as the complex biological interactions of the *in vivo* system cannot be controlled as well as the environment in a dish.

Based on our findings, we generated a general hypothesis: regional reprogramming is controlled by activation of transcription factor signaling in the bowel. We further hypothesized that using direct reprogramming to test the effects of transcription factors on regional reprogramming would provide us not only with information regarding which transcription factors play a role in the process, but also allow us to dissect their downstream effects without

the confounding interactions that can occur *in vivo*. We aim for this study to ultimately culminate in a direct reprogramming model of regional reprogramming after the factors are delineated.

Our development of a direct reprogramming system for regional reprogramming provides tighter molecular control over the regional reprogramming process. The addition of this model to our collection of models provides a system in which putative regulators of regional reprogramming can be studied for their effects on inducing adaptation in patient tissue. The combination of this model with the chimeric system provides a model for signal discovery (the chimera) and for pathway effect elucidation (the direct reprogramming model), the data from which can improve our understanding of regional reprogramming after SBR in mice and the effects of applying our findings to human organoids *in vitro*.

We initiated our transition from our traditional direct reprogramming of iEPs (a 2-dimensional system) to direct reprogramming of organoids (a 3-dimensional system) through first utilizing a direct reprogramming system from 2-D human umbilical vein endothelial cells (HUVEC cells) to 3-D human intestinal organoids as published by Miura & Suzuki in 2017.[80] This system proved difficult to scale up and mature the resulting spheroids for study, so we refocused on direct reprogramming of existing intestinal organoids for use in modeling regional reprogramming. Through gene regulatory network (GRN) reconstruction, we identified two potential upstream regulators of regional reprogramming (*Klf4* and *Creb3l3*) and searching the literature provided us with an additional target (*Gata4*). Using these transcription factors, we generated an organoid direct reprogramming system to study the effects of overexpression of each factor on inducing regional reprogramming. We identified *Gata4* as holding the most promise in ability to proximalize distal organoids, but found that overexpression of *Gata4* alone was insufficient to recapitulate regional reprogramming for *in vitro* modeling. Interestingly,

Gata4 does appear to induce differentiation changes in the epithelial composition of organoids. We conclude that use of direct reprogramming on organoids allows for investigation of the effects of gene overexpression on cellular identity. Additionally, we conclude that while *Gata4* alone does not induce regional reprogramming, it is a likely candidate for this process if used in combination with other transcription factors in future studies.

4.3 Direct Reprogramming to Human Induced Intestinal Progenitors Pilot Indicates Inefficiency of Reprogramming

We began our work in organoid reprogramming through using an adaptation of the iEP direct reprogramming process to produce human intestinal organoids from other cells. We hypothesized that the production of mature human intestinal organoids from non-intestinal cells could assist in the derivation of patient-derived organoids for use in disease and treatment modeling purposes if the process proved efficient at generating mature organoids.

As discussed in Subsection 1.1.3, direct reprogramming was utilized to make induced hepatocytes (iHeps) from mouse embryonic fibroblasts (MEFs) through overexpression of *Hnf4a* and *Foxa1/2/3*.[78] However, when iHep cells were further analyzed using a GRN analysis tool called CellNet, it was discovered that these cells were not fully hepatocytes in identity but had additional colonic identity potential. This was validated through *in vivo* transplantation of iHeps into damaged mouse colon, where they were able to functionally engraft.[77] This study showed a wider endodermal potential for these cells, which we now refer to as induced Endoderm Progenitors (iEPs).

After the discovery of colonic identity in iEPs, the Suzuki group generated reprogrammed mouse organoids from MEFs capable of regenerating intestinal epithelium and colonic epithelium. Overexpression of four factors, *Hnf4a, Foxa3 (*the two iEP factors), as well as *Gata6* and *Cdx2,* resulted in formation of budding organoids from induced intestinal stem

cells (iISCs). However, while expression of these four factors in human umbilical vein endothelial cells (HUVECs) produced expression of intestinal markers, it formed only immature intestinal spheroids, failing to reach an intestinal stem cell state capable of mature differentiation.[80] These human cells are what we refer to as human induced intestinal progenitors (HIIPs) but fail to fully mature and differentiate like their mouse counterparts (Figure 4.1). We set out to enhance this model system through improving the maturity of the final organoids for use in modeling patient specific intestinal disease and treatments.
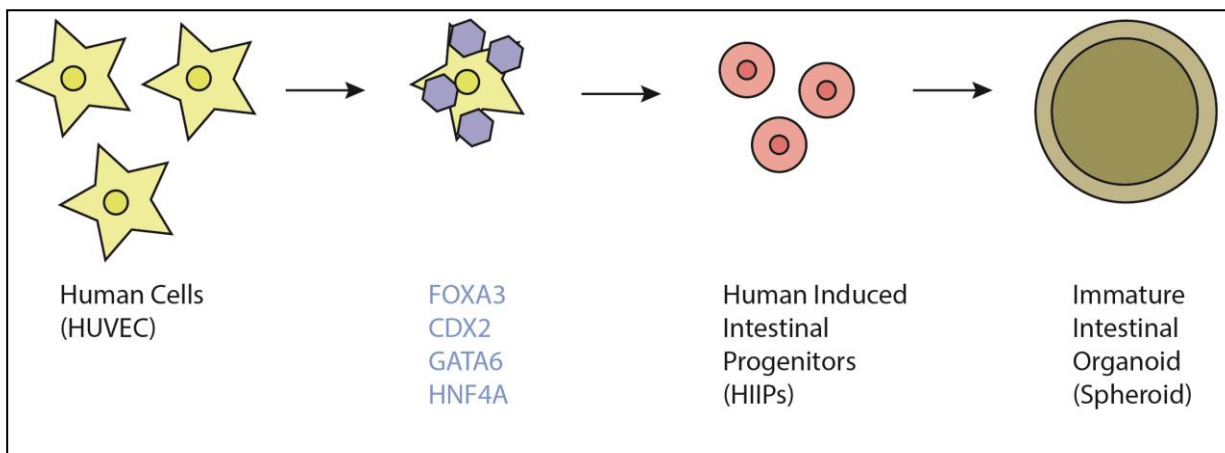
Figure 4.1: Schematic of HIIP reprogramming. HUVEC cells are transduced with retroviral constructs for *FOXA3, CDX2, GATA6,* and *HNF4A* (blue) every 12 hours for 8 times. Transduced cells are plated in 3-dimensions in Matrigel and allowed to reprogram for 20-21 days, resulting in immature intestinal spheroid structures.

While this model was reported to be tractable in the literature, our attempts to recapitulate the cell type in order to develop methods to further mature the end state suffered from one of the main issues surrounding *in vitro* reprogramming, low reprogramming efficiency.[197,198] Using constructs cloned in our laboratory as well as the original constructs obtained from the Suzuki Lab, we performed numerous rounds of reprogramming but struggled to achieve reported results (Figure 4.1).[80]

Initially, with our cloned constructs, we produced small spheroid-like clusters which we analyzed for markers of the developing and mature intestine via qPCR (Figure 4.2). We compared the 3-D cultured reprogrammed HIIPs to the parent cell type, HUVECs, grown in the same 3-D conditions to control for the effects of culture on the cell identity. In doing so, we aimed to determine if the resulting cells were taking on any sort of reprogrammed identities. We found that the analyzed HIIPs showed increased expression of general stem cell markers (e.g.: *SOX9* and *KLF5)* and even intestinal stem cell (ISC) markers, notably *LGR5,* when compared to HUVEC controls.[199–201] However, markers of mature intestine, such as *VIL1*, a marker of mature enterocytes, failed to be detected in either sample via qPCR. Based on these findings, we concluded that the reprogramming process was working on some level. It was producing a more progenitor-like state in the HIIPs than in the control HUVECs that, similar to Miura & Suzuki's results, failed to mature to a full, differentiated intestinal state.

The organoids produced in our experimentation using the constructs obtained from the Suzuki Lab closely resembled immature intestinal organoids, sometimes referred to as fetal spheroids or enteroids (Figure 4.3). These cultures consisted of round, epithelialized layers surrounding a hollow lumen, similar to those expected based on the results in Miura & Suzuki, 2017.[80] However, these resulting spheroids failed to represent a tractable system to achieve our goals of an efficient, mature *in vitro* model. Primarily, the formation rate of spheroids was less efficient than the expected 0.1%, with only 3 spheroids forming from a population of 100,000 initial cells. This initial inefficiency of reprogramming may have been able to be overlooked had the cells been able to be passaged and expanded after the reprogramming for use in further maturational studies. However, the cells failed to expand—making the ultimate goal of increasing the maturity of the final cell types unlikely in this context.
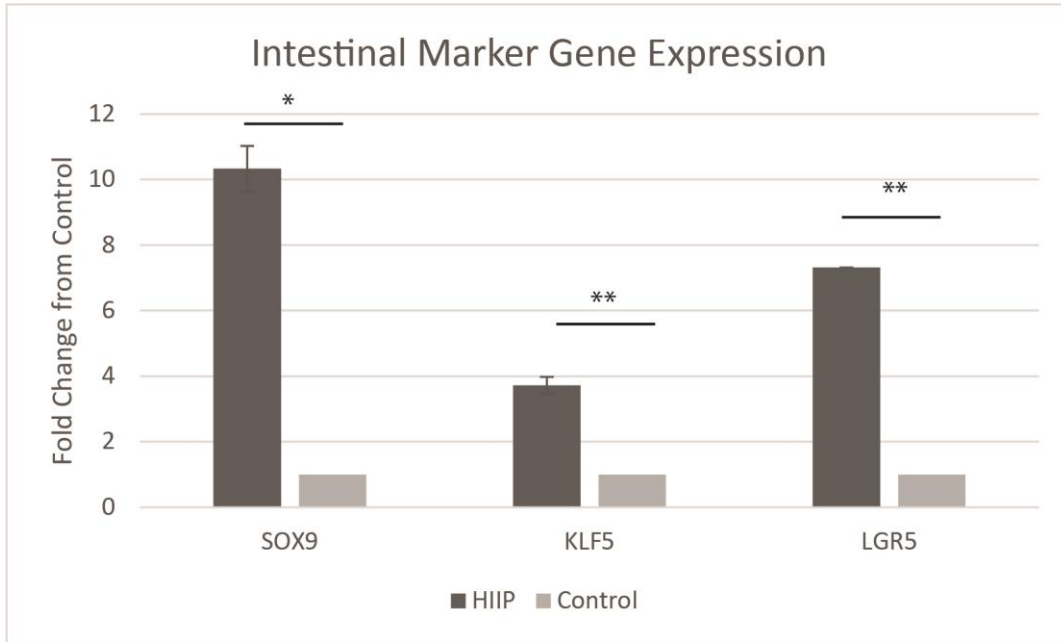
Figure 4.2: qPCR results indicate increased progenitor-like state in HIIPs. qPCR analysis shows upregulation of *SOX9* (stem cell marker), *KLF5* (stem cell marker), and *LGR5* (intestinal stem cell marker) when compared to HUVEC controls. qPCR failed to detect *VIL1* (mature enterocyte marker) in either sample. Fold change is calculated vs. control sample (value = 1) using *GAPDH* as a normalization control. Graph presented as fold change $\pm$ SD. $*P < 0.05$, $**P < 0.01$. Statistics shown represent mean delta CT of technical duplicates.
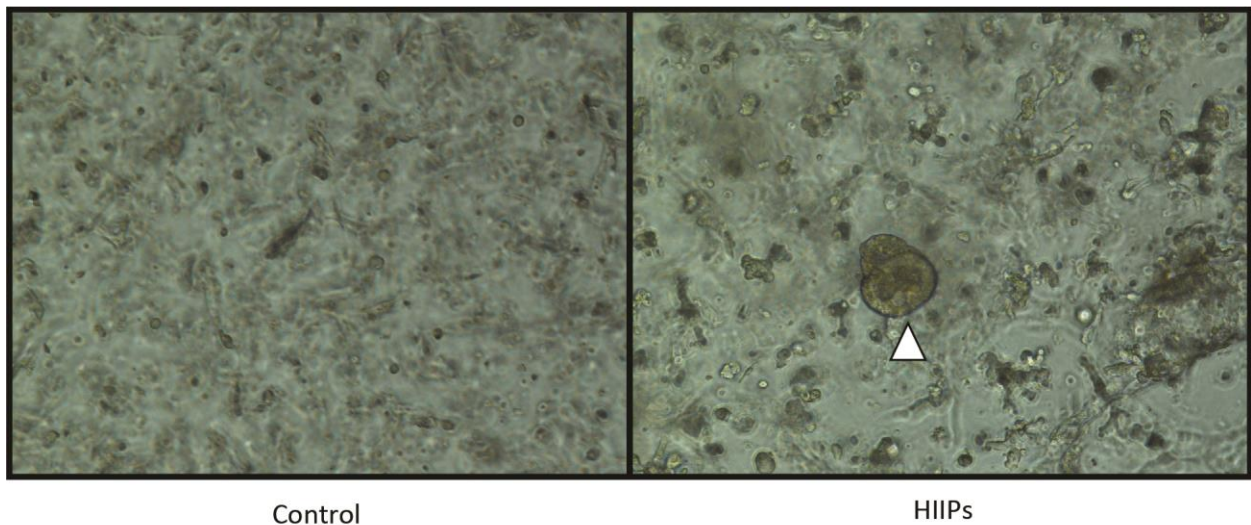
Figure 4.3: HIIP spheroids *in vitro*. Left: Control HUVECs cultured in 3-dimensions for 21 days; Right: Reprogrammed HIIPs cultured in 3-dimensions for 21 days. White arrowhead denotes epithelialized spheroids. Images taken on Nikon Eclipse Ts2 with Nikon DS-Fi2 camera with unknown objective.

Overall, this study highlighted the distinct challenges surrounding achieving efficient direct reprogramming and maturation of reprogrammed cells for modeling purposes. While we were able to generate cell types that were becoming more stem-like compared to controls, we

101

couldn't generate these cells at a high enough efficiency to merit continued use. Because of our difficulties in expanding the cultures, we were ultimately unable to promote maturation. However, having learned from the difficulties surrounding this reprogramming process, we altered our approach moving forward to improve both the potential maturity of the final cells and the expandability of the cultures after reprogramming. To improve on these fronts, we decided to utilize existing organoids as the starting cell for reprogramming. By doing so, we were able to ensure the genetic and regional origin of the starting organoid and promote the starting cell maturity.

4.4 Determination of Potential Transcription Factor Regulators of Regional Reprogramming

In Chapter 2 we discussed the discovery that *Creb3l3* and *Klf4*, transcription factors, were upregulated post-resection. We noted the likelihood that other transcription factors could be highly involved but undetectable by scRNA-seq due to transient temporal expression.[183] Therefore, in order to predict transcription factors likely to be playing a role in regional reprogramming that we could test *in vitro*, we again turned to *in silico* means, using a tool called CellOracle, to identify (or validate) our targets.

CellOracle is a GRN reconstruction algorithm that additionally allows for simulation of the effects of gene overexpression or knockout on scRNA-seq data. In brief, CellOracle uses the differential gene expression data found in a scRNA-seq dataset to infer which transcription factors are active to create the unique gene signatures in the dataset. Then it provides visualization and prediction of what transitions the cells in the scRNA-seq dataset are likely to undergo if the expression of different transcription factors is altered.[101]

To identify transcription factors likely involved in regional reprogramming, we analyzed the SBR datasets from Chapter 2. Initial analysis of the datasets provided a number of potentially interesting factors involved in proximal enterocyte identity including: *Fos, Fosb, Egr1, Klf4, Jun, Creb3l3,* and *Atf3.* To determine which had the most potential in our system to induce the intended alterations, we used the perturbation simulation function of CellOracle on each factor. Some factors, such as *Egr1* and *Atf3,* did not show large induction of proximal enterocyte identity when overexpressed and were dismissed (data not shown). *Fos* and *Jun* showed promise when overexpressed (Figure 4.4*A-B*) as large portions of the enterocyte population shifted toward proximal enterocyte identities. *Creb3l3* and *Klf4* also had large shifts toward proximal enterocyte identities when overexpression was simulated (Figure 4.4*C-D*).

Fos and *Jun* are members of the Activator Protein1 (AP-1) complex which is known to play roles in nearly every cellular process from developmental proliferation and differentiation to tumorigenesis.[202] We chose not to pursue *Fos* and *Jun* in our initial study due to their widespread involvement across tissue and processes making their specific involvement in regional reprogramming unlikely.

In lieu of *Fos* and *Jun,* we chose to target *Creb3l3* and *Klf4* based on their known actions and the results of CellOracle. Interestingly, both *Creb3l3* and *Klf4* were detected as being upregulated post-resection in Chapter 2, making them stand out as potential targets.[183] *Creb3l3* is known to be proximally intestinally expressed and to control lipid metabolism and the response to inflammatory infection. The Creb3 family of proteins are additionally known to be regulated by nutrient and stress signals. The changes in nutrient absorption, inflammation, and innate stress response following SBR will likely alter *Creb3l3* expression levels, making it a very promising target for regulation of regional reprogramming.[126] *Klf4,* similar to *Creb3l3,* is reported to have

highest expression levels in the proximal intestine. Functions of *Klf4* in the intestine include homeostatic regulation of growth, cellular maturation, and differentiation as well as tumor suppression.[129] We deemed *Klf4* a promising target for regional reprogramming because of its regional expression pattern and roles in homeostasis and cellular differentiation—processes that are disrupted by SBR.

While researching potential regional regulators, we noted that an additional factor, *Gata4*, had a very interesting relationship with the development and maintenance of regional identity in the small intestine. *Gata4* is a particular type of transcription factor known as a pioneer factor, or a transcription factor capable of binding to nucleosomal sites and increasing their accessibility to other transcription factors.[203] According to the intestinal literature, GATA4 is expressed in duodenal and jejunal (proximal) enterocytes and is not expressed in ileal (distal) enterocytes. While it plays roles in development, most notable are the identity maintenance roles attributed to *Gata4*. *Gata4* is both necessary to maintain jejunal gene expression (loss of *Gata4* induces ileal gene expression profiles) and sufficient to drive jejunal (proximal) identity in ileal tissue.[204–206] In addition, *Gata4* is a transcriptional repressor of *Fgf15*, a marker of distal enterocyte identity in the intestine.[204] Because of its specific role in delineating distal vs. proximal enterocyte identity in the small intestine, we chose to study *Gata4* as well, as its functions are closely related to the process of regional reprogramming. Unfortunately, CellOracle was unable to utilize *Gata4* for perturbation simulations as its detected expression in the dataset was too low, which is unsurprising as transcription factors are often transiently or lowly expressed and difficult to detect by scRNA-seq (data not shown).
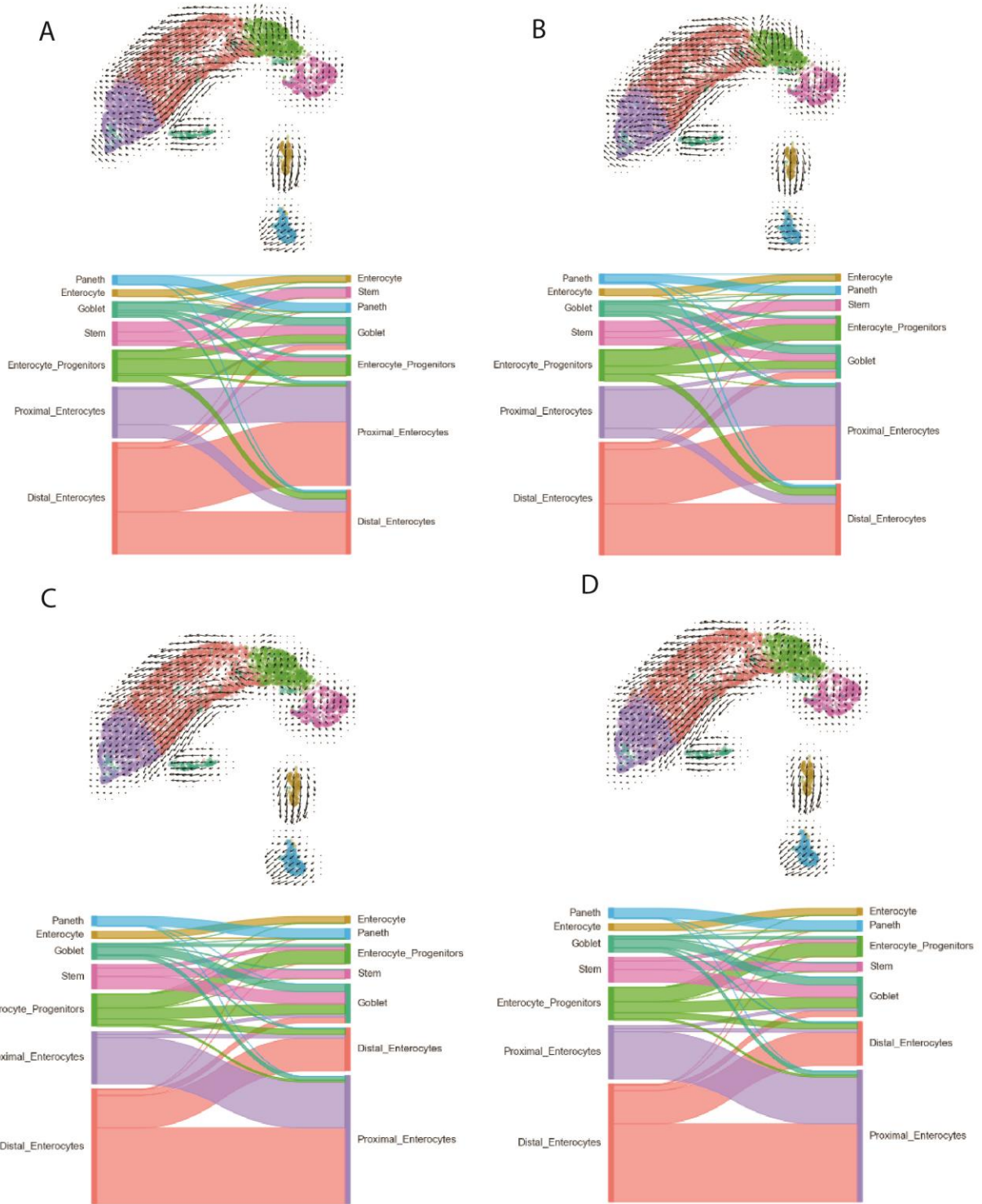
Figure 4.4: CellOracle perturbation simulations for regional reprogramming transcription factors. Top image of each panel depicts a trajectory graph showing predicted transitions in cell identity. Bottom image in each panel depicts a Sankey diagram of predicted transitions in cell identity (starting identity on left, ending on right). (*A*) *Fos* perturbation analysis. (*B*) *Jun* perturbation analysis. (*C*) *Creb3l3* perturbation analysis. (*D*) *Klf4* perturbation analysis. Visualization examples from dataset "SBR2".

In summation, using both *in silico* predictive tools and literature searching, we identified *Creb3l3, Klf4,* and *Gata4* as likely candidates for induction of regional reprogramming after SBR. We moved forward to test the effects of these transcription factors on induction of regional reprogramming using direct reprogramming methods on intestinal organoids.

4.5 *Gata4* Overexpression Induces Molecular Changes in Ileal Organoids Mimicking More Proximal Identities

To generate an organoid system in which to overexpress our factors of interest (*Creb3l3, Klf4,* and *Gata4)*, we inserted our genes of interest into an established retroviral vector system for intestinal organoid transduction, pMSCV-loxp-dsRed-loxp-(GENE)eGFP-Puro-WPRE.[43] The use of this vector, when inserted into Cre-expressing organoids, allows for the inducible recombination of the backbone such that dsRed and a stop codon are cut out, resulting in the transcription of the gene of interest and eGFP. This makes for a simple, visual system to monitor transduction efficiency (dsRed expression) and expression of the gene(s) of interest (eGFP expression). We expressed each gene of interest containing vector, and an empty control vector, in *Villin*-CreERT2 distal (ileal) intestinal organoids. The use of the *Villin*-CreERT2 organoids allows for targeting of the Cre-recombinase to *Villin* expressing intestinal epithelial cells, a marker of mature epithelium, and additionally allows for the activation of Cre-recombinase using Tamoxifen.

While set up the same as in Koo et al., 2012, our system experienced leaking of the Cre-recombinase, resulting in baseline levels of target gene expression (and eGFP) before Tamoxifen was added to cultures (Figure 4.5). Gene expression of transgenes was significantly increased compared to controls, but unpredictable in terms of expression by vehicle or Tamoxifen (Figure

106

4.5*B)*. In attempts to eliminate this leakage, we moved the organoids to Phenol Red-free Matrigel, as phenol red can act as a weak estrogen and activate CreERT2 (estrogen responsive) systems and maintained the cells in IntestiCult.[207] However, this change did not lower the baseline level of target gene expression. Further research revealed that CreERT2 systems can experience spontaneous recombination if the cells express high endogenous levels of CreERT2.[208] We postulate that the leakiness in our system is due to naturally high CreERT2 expression in the mouse line the cells were derived from. As our system was initially set up merely to test the effects of overexpression of transcription factors on regional reprogramming and not necessarily to test temporal expression, we chose to continue using the lines we generated for further experiments. In the case that administration of Tamoxifen induced additional expression above baseline, we included Tamoxifen treatment in our experiments. This decision was made with the understanding that future experiments in which a temporal aspect was of interest would require use of a less leaky Cre-recombinase organoid line.
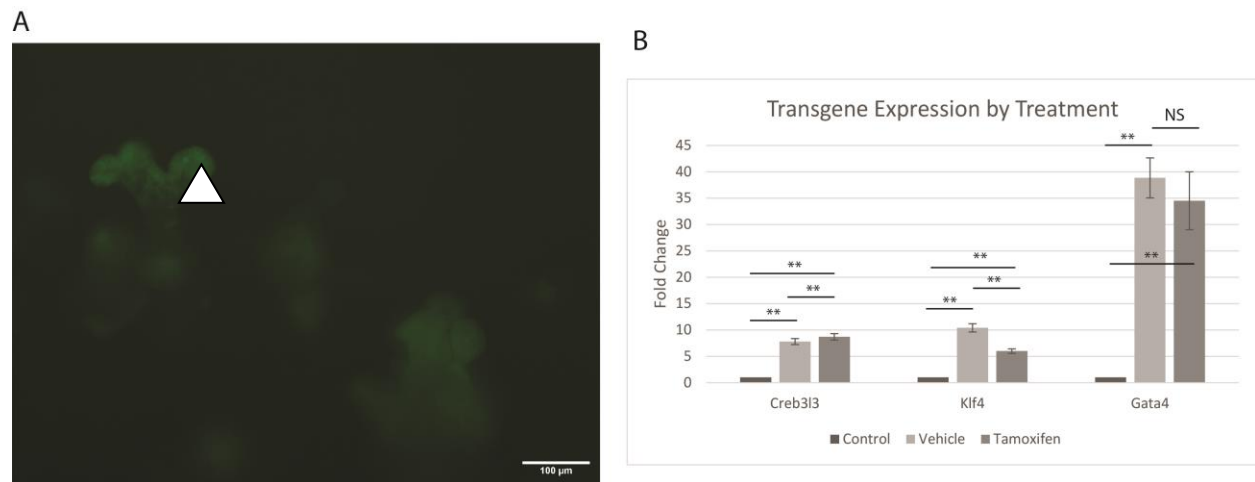
Figure 4.5: Organoids exhibit leaky CreERT2 expression. *(A)* Representative image of empty control Villin-CreERT2 organoid expressing eGFP after vehicle administration (white arrowhead). Scale bar = 100 µm. *(B)* qPCR analysis shows increased expression of all genes of interest when compared to empty controls. Samples indicate that genes are expressed without the administration of tamoxifen (+T) and with only vehicle added to the wells (+V). Fold change is calculated vs. control sample (value = 1) using *Hprt* as a normalization control. Images taken on Nikon Eclipse Ts2 with Nikon DS-Fi2 camera with a 10× objective. Graph presented as fold change ± SD. *P < 0.05, **P < 0.01, NS < not significant. Statistics shown represent mean delta CT of technical triplicates.

After establishing 4 stable distal organoid lines expressing the 3 genes of interest and the control empty vector, we decided to narrow our focus to the most promising factor to start so that future studies could be more easily optimized. We induced gene expression in our cells using Tamoxifen then used brightfield imaging to investigate for phenotypic differences between the gene overexpression organoids and the empty vector containing distal control (hereby called "empty control"). We anecdotally noted slight visual differences including appearance of extended (long) crypt structures in *Klf4, Gata4,* and *Creb3l3* and the appearance of more immature/stem-like spheroids in *Klf4, Creb3l3,* and *Gata4* in organoids grown in IntestiCult (Figure 4.6*A*). Based on these observations, we posit that all three genes result in increased stem cell proliferation, leading to the occurrence of deeper crypts and non-differentiated spheroids, but we could not make any conclusions regarding differentiated regional identity.[209]

To obtain a more quantitative result, we then compared the samples (including a proximal organoid control derived from *Villin*-CreERT2 duodenum treated with vehicle) via qPCR to look for several markers (*Fabp1, Fabp6, Apoa1, Apoc3,* and *Rbp2*) defining regional reprogramming/proximal enterocyte identity as determined in Chapter 2 in organoids grown in our custom growth media to promote differentiation (Figure 4.6*B*).[183] Our proximal control displayed the expected gene expression pattern of increased *Fabp1, Apoa1, Apoc3,* and *Rbp2* and decreased *Fabp6* when compared to the empty distal control. Only *Gata4* recapitulated the expected gene expression pattern for proximal tissue with increased *Fabp1, Apoa1, Apoc3,* and *Rbp2* and decreased *Fabp6* when compared to the empty distal control, though all values were still statistically significant in differences between the pairs (Figure 4.6*B*).
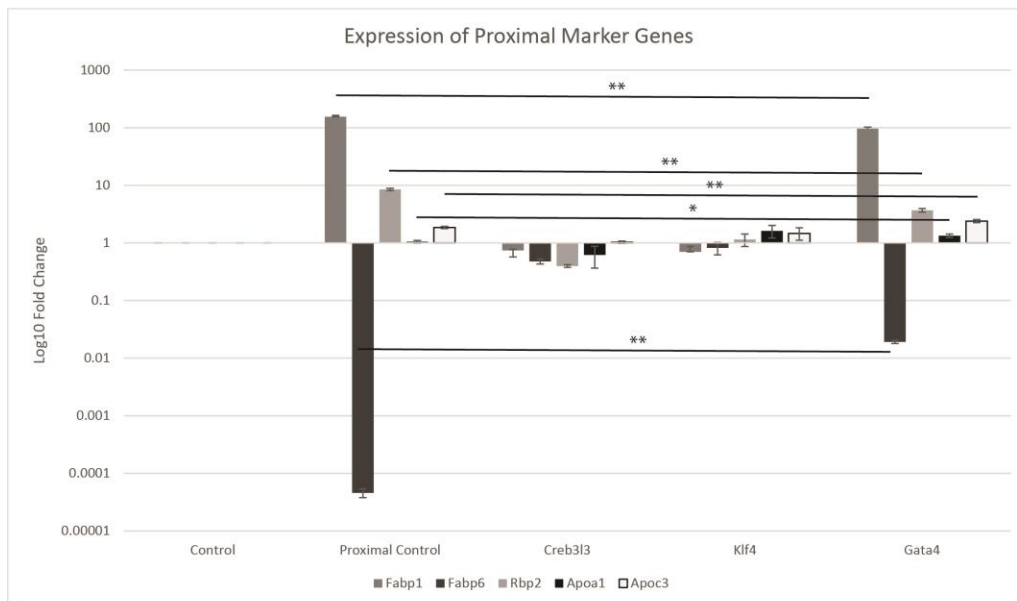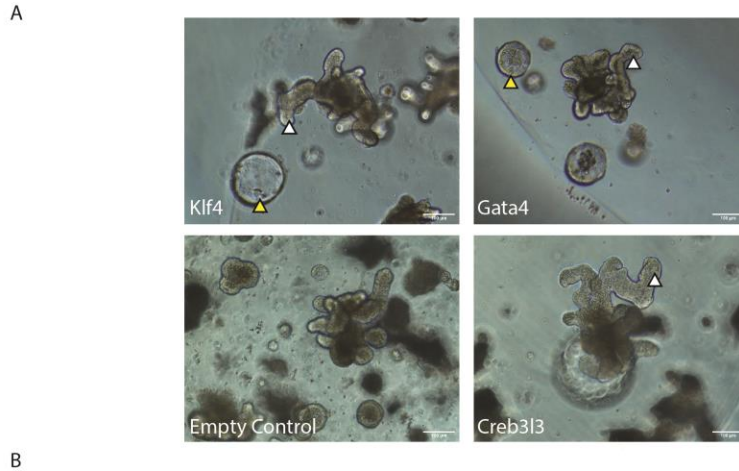
Figure 4.6: *Gata4* expressing organoids show most likely proximal identity acquisition. *(A)* Representative images of phenotypic changes in gene expressing organoids and empty controls grown in IntestiCult. White arrowheads indicate extended crypts and yellow arrowheads indicate stem-like spheroids. *(B)* qPCR analysis for *Fabp1, Apoa1, Apoc3, Rbp2,* and *Fabp6* when compared to the empty distal control. Fold change is calculated vs. empty control sample (value = 1) using *Hprt* as a normalization control. Images taken on Nikon Eclipse Ts2 with Nikon DS-Fi2 camera with a 10× objective. Graph presented as fold change ± SD. *$P < 0.05$, **$P < 0.01$. Statistics shown represent mean delta CT of technical triplicates of only *Gata4* sample and proximal control due to space constraints.

Therefore, we chose to move forward and focus on *Gata4* in our scRNA-seq studies, as it

gave the greatest evidence of taking on more proximal identities. The failure of *Creb3l3* and *Klf4*

to recapitulate the expected patterns of gene expression do not reflect a lack of potential

involvement in regional reprogramming, there is still logical reasoning for their involvement. Instead, we postulate that these transcription factors may not be responsible alone for regional reprogramming but instead act in concert with other signals and a more combined approach is necessary to investigate their effects in the future.

4.6 scRNA-seq Reveals *Gata4* is Insufficient to Induce Proximal Enterocyte Identity in Organoids

We used scRNA-seq to analyze the *Gata4* overexpressing organoids alongside the empty distal controls and proximal controls grown in custom growth media. The data was aligned and run through quality control measures before utilizing Seurat to perform dimensional reduction and investigate differential gene expression between populations.[102] Seurat analysis resulted in a UMAP visualization consisting of 12 clusters (Figure 4.7*A*). The analyzed data consisted of 1926 empty control cells, 3646 proximal control cells, and 1914 *Gata4* expressing cells.

Visualization of the UMAP by original sample identity of the cells showed clear delineation between the control proximal and control distal populations, similar to those discussed in Chapter 3 (Figure 4.7*B*). The proximal controls (blue) are predominantly distinct from the distal controls (red), which indicates clear transcriptional differences between control populations, as previously described.[3,7] The *Gata4* expressing organoids (green) overlap with the empty control populations and additionally extend out above the empty control populations into a region populated only by *Gata4* derived cells. This distribution does not support our hypothesis that overexpression of *Gata4* in distal organoid cells will induce a transition toward proximal enterocyte cell identities representative of regional reprogramming. While we wouldn't necessarily expect the *Gata4* cells to move to the proximal control area on the UMAP due to a potential "hybrid" identity (discussed in Chapters 2 and 3), we did not expect to see *Gata4* cells

outside the controls on the UMAP. The appearance of these cells in their own region may indicate expression profile transitions toward unknown cellular identities.
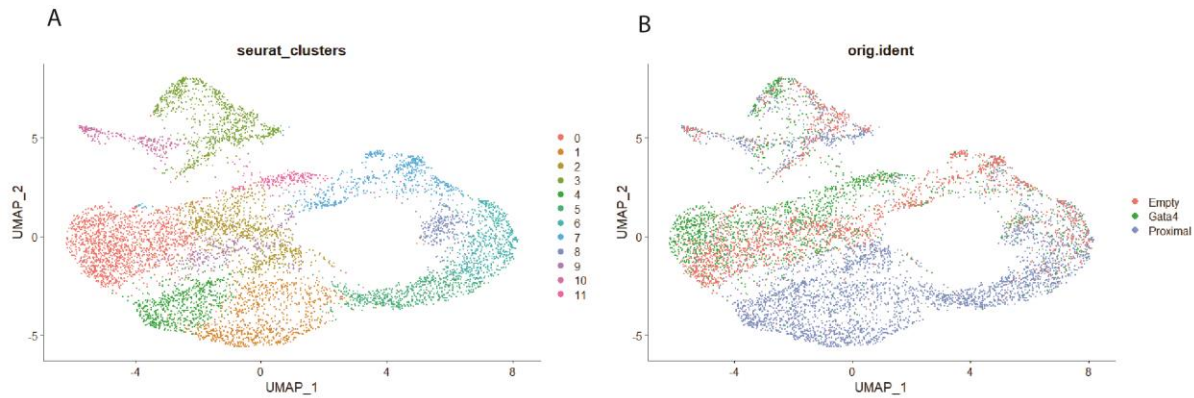
Figure 4.7: Single cell RNA-seq analysis visualizations of *Gata4* experiment. *(A)* UMAP visualization of the 12 clusters. *(B)* UMAP visualizations original sample identities.

To verify that these results did not support our hypothesis, we compared the *Gata4* expressing cells to the empty distal control populations to look for differences in proximal enterocyte identity acquisition. As in Chapter 3, using Haber et al., 2017, we compiled a list of markers of proximal enterocyte identity (Table 3.1).[106] Seurat was utilized to generate a proximal enterocyte identity score for each cell in the *Gata4* and empty control populations based on their respective expression of the genes in the list. Permutation testing was utilized to compare the mean proximal identity score of the populations to determine if the difference between proximal enterocyte identity scores was statistically significant. We found that *Gata4* cells had significantly lesser proximal enterocyte identity scores than empty distal controls ($P = 0.01006$). This result was surprising, as the detection of proximal enterocytes in the empty distal control is expected to be very low. Interested in this finding, we used Haber et al., 2017 to compile a list of distal enterocyte markers to test if there was a difference in distal identity between the two

samples (Table 4.1).[106] We found that *Gata4* cells had significantly lesser distal enterocyte identity scores than empty distal controls as well (P = 0).

Table 4.1: Distal Enterocyte Gene List for Identity Scores

Reg3g	Gsdmc4	Prss32	Krt8	Tmigd1	Fabp6	Slc51b
Slc51a	Mep1a	Fam151a	Naaladl1	Slc34a2	Plb1	Nudt4
Dpep1	Pmp22	Xpnpep2	Muc3	Neu1	Clec2h	Phgr1
2200002D01Rik	Prss30	Cubn	Plec	Fgf15	Crip1	Krt20
Dhcr24	Myo15b	Amn	Enpep	Anpep	Slc7a9	Ocm
Anxa2	Aoc1	Ceacam20	Arf6	Xpnpep1	Abcb1a	Vnn1
Cndp2	Nostrin	Slc13a1	Aspa	Maf	Myh14	

We were initially confused by this seemingly confounding result. First, we hypothesized that the lesser detection of proximal identity could have to do with duodenal vs. jejunal proximal identities. The proximal markers from the Haber et al., 2017 publication are not defined as "duodenal" or "jejunal".[106] Most work on *Gata4* indicates that it is specifically responsible for jejunal identity, so we theorized that the markers used for our identity classification may not align with specific markers of jejunal identity, skewing our detection of proximal identity. We cross-referenced our list (Table 3.1) with the jejunal-enriched transcripts in Thompson et al., 2017 and found that several factors overlapped (*DnaseI, Fabp1, Hsd17b6,* and *Lct*), indicating that our marker identity list was likely not completely off the mark.[204] This indicated that *Gata4* overexpression did not increase proximal enterocyte identity in our system, but we wondered, based on the outlying population in the UMAP, if the *Gata4* population was seeing changes in cellular composition outside of enterocyte identities. We hypothesized that our seemingly confounding enterocyte identity scores may represent a loss of enterocytes in the *Gata4* samples compared to the empty distal controls, resulting in the control organoids having higher proximal *and* distal enterocyte identities due to cell composition changes.

We conclude, based on this analysis, that the *Gata4* overexpression model as developed in this study fails to recapitulate regional reprogramming as detected by scRNA-seq analysis. It seems that in an *in vitro* system, *Gata4* is not sufficient to induce proximal enterocyte identity in organoids. It is apparent, however, that the overexpression of *Gata4* in organoids produces identity changes of an unknown type, so we set out to determine what compositional identity changes were occurring due to *Gata4* overexpression.

4.7 *Gata4* Reprogrammed Organoids Take on More Stem-Like Identities

To dissect the cellular composition of our preliminary datasets to look for transitions due to *Gata4* overexpression, we again employed the quadratic programming approach Capybara.[100] This computational method uses quadratic programming to compare the gene expression of query datasets against annotated reference datasets, in our case, the Haber et al., 2017 intestinal atlas.[106] Uniquely, Capybara allows for classification of each cell in the dataset based on the reference and additionally allows for the identification of cells in transition states expressing intermediate identities ("multi-ID" cells).

We first utilized Capybara to generate specific cell classifications for the *Gata4*, empty distal control, and proximal control populations (Figure 4.8). The majority population of each sample consisted of stem cells ($\geq$57.7%). While the presence of stem cells is normal and indicates continued turnover of the organoids just like the endogenous epithelium, this large population is a relative overabundance of what would be expected, similar to the results in Chapter 3. Again, it is likely due to the composition of the growth medium and the abundance of factors such as Noggin and R-spondin 1 as it has been reported that titrating small molecules in the growth media of organoids can direct production of certain lineages.[194] Future experiments

113

using this system should optimize the media composition to promote differentiation, however, we consider these results valid as all organoids for 10x were grown in the same medium composition, therefore any cellular composition changes compared to one another are due to the organoids themselves.

There were clear shifts in compositional identities between populations (Figure 4.8). There was an increase in the tuft cell population in the *Gata4* (0.8%) and empty distal controls (1.2%) compared to proximal controls (0.2%). This finding supports literature reports of higher incidence of tuft cells in the distal small intestine.[210] Additional observations included a decrease in goblet cells in *Gata4* cultures compared to empty controls (6.9% vs. 9.4%). This finding supports literature reports of the effects of *Gata4* expression on goblet cell populations, supporting our overexpression of *Gata4* as having results in the *in vitro* system comparable to reports in *in vivo* systems.[123,205] Notably, *Gata4* cultures showed higher stem cell populations than the empty controls (68.1% vs. 57.7%).

Pertinent to our amended hypothesis, *Gata4* cultures did show an overall decrease in enterocyte identities (immature and mature, proximal and distal) when compared to distal and proximal controls (combined 8.5% vs. 18.7% and 16.2% respectively). This finding supports our hypothesis that *Gata4* overexpression resulted in a loss of enterocyte identity in the *in vitro* system and a shift toward more undifferentiated (stem cell) identities. As *Gata4* is reported sufficient to repress ileal enterocyte identity, we hypothesize that the loss in enterocyte number may reflect an initial loss of distal enterocytes from the population (perhaps via dedifferentiation to stem cells), a reported function of *Gata4*.[204] We would expect that an increase in proximal enterocytes would accompany this loss, but our dataset does not indicate this. We hypothesize two possible explanations: 1) the accompanying upregulation of proximal enterocytes is a later

114

temporal process and given more development time *in vitro*, we would see an increase in proximal enterocyte in *Gata4* organoids or 2) *Gata4* depends on other *in vivo* signals and variables to promote proximal identity after dedifferentiation which are not active in our *in vitro* system. While not as underpowered as the experimentation in Chapter 3, to verify these preliminary results and generate statistics regarding population composition changes, the *Gata4* scRNA-seq should be replicated and additional staining experiments used as validation.

As we would expect our reprogramming via *Gata4* would induce a transitional state in which some cells have hybrid identities, we investigated the multi-ID cell populations as well. A goblet-stem cell intermediate occurred at a frequency of 39.2% of multi-ID cells in *Gata4* populations while only appearing at a frequency of 20% in empty controls. We hypothesize that this intermediate may represent dedifferentiation of goblet cells toward stem cells in *Gata4* populations, accounting for the overall loss of goblet cells in that treatment. Additionally, we noted an increase in enterocyte progenitor early-stem cell intermediates in *Gata4* populations compared to empty distal controls (35.3% vs. 15% of multi-ID cells). We hypothesize that this population accounts for some of the low levels of enterocyte populations we report in *Gata4* overexpression. Based on our findings, we hypothesize that *Gata4* overexpression is resulting in a dedifferentiation of enterocyte progenitors to stem cells, resulting in the differentiation of fewer enterocytes overall.[211] This is supported by our qualitative deduction that *Gata4* expressing organoids had increased stem cell activity (Figure 4.6*A*). It is also possible that this population represents stem cells differentiating toward enterocyte progenitors. In summation, we conclude via Capybara analysis that *Gata4* overexpression altered cellular composition of the organoids such that there is a loss of enterocyte identities and an increase in stem-like states indicating roles for *Gata4* in cellular differentiation in the *in vitro* organoid system.
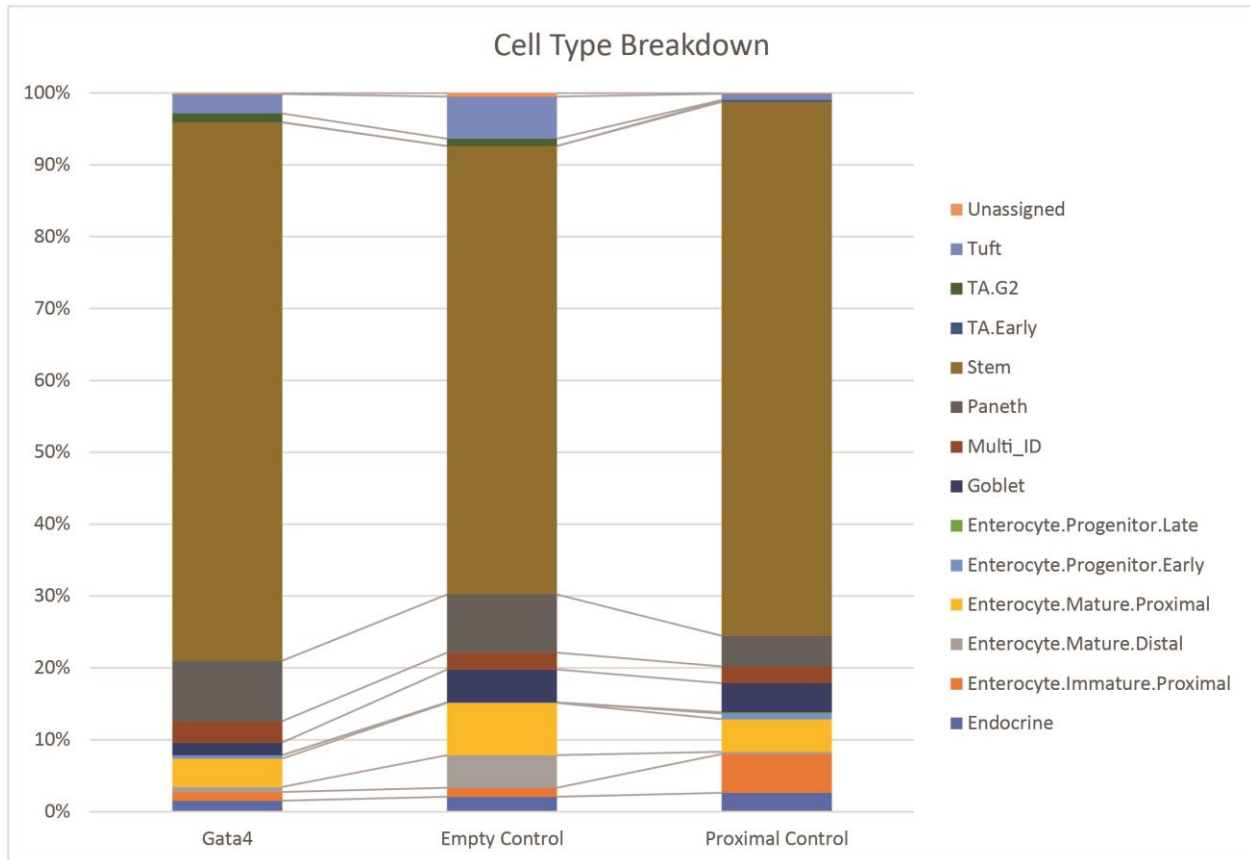
Figure 4.8: Capybara analysis of *Gata4* experiment. Capybara classifications of cells in each sample as percent total cells. (No statistics available due to n = 1 and small sample sizes).

4.8   Discussion

In this chapter, we report on our generation of directly reprogrammed intestinal organoids for the study of transcription factor effects on regional reprogramming *in vitro* and the downstream analysis of cellular identity using scRNA-seq. We learn, in a pilot experiment, that maturation of organoids is difficult to achieve and maintain *in vitro* and therefore utilize adult intestinal organoids with custom growth media to promote maturity in additional experiments. Our direct reprogramming model using *Gata4* showed cellular identity changes detectable by scRNA-seq and thus we propose the continued use and optimization of this model to look at effects of various transcription factors on the intestinal epithelium. Because our analysis of *Gata4* overexpression indicates that *Gata4* alone is insufficient to induce regional

reprogramming *in vitro* but is sufficient to alter cellular differentiation and composition, we propose that combinations of transcription factors will be required *in vitro* to mimic regional reprogramming. Our findings in this chapter fail to support our hypothesis that *Gata4* induces regional reprogramming *in vitro* but provide valuable information on utilization of direct reprogramming and *in vitro* organoids models for the future.

The effects of *Gata4* on *in vivo* intestinal systems have been well studied, but our model allows us to zoom in on how *Gata4* functions in the intestine without the complication of other systems and signals *in vivo*. Our demonstration of the use of direct reprogramming to monitor the effects of transcription factors on organoids provides the field with the groundwork for investigating the potential roles of different transcription factors during regional reprogramming. It is a valuable system to accompany our preliminary model from Chapter 3, the chimera, in that while the chimera provides opportunity to discover new signals at play in regional reprogramming, the direct reprogramming model provides the opportunity to test known signals *in vitro*. Together, the combined use of these systems can help clearly delineate the upstream elements triggering regional reprogramming in mice and their downstream effects. Improved delineation of these triggers will better inform future studies of induction of regional reprogramming in human populations.

In a pilot study attempting to address the immaturity of directly reprogrammed human intestinal organoids, we found that generating organoids through direct reprogramming was so inefficient that it prevented us from achieving our goal of increasing organoid maturity *in vitro*. Both low efficiency of reprogramming (reported 0.1%-20% efficiency) and immature gene expression patterns are issues that plague reprogramming experiments.[72,77,197,198] However, since improving efficiency was not a target of these experiments, we used the failure of the HIIP pilot

study to inform our future experiments in intestinal direct reprogramming. Based on the low efficiency in reprogramming HUVECs into intestinal organoids, we chose to start with intestinal organoids as our starting cell type moving forward. We additionally hoped that use of adult intestinal organoids would eliminate issues of cell maturity.

The use of CellOracle allowed us to easily identify transcription factors of interest without the extended time needed to search the literature for potential factors. While the analysis pointed toward *Creb3l3* and *Klf4* as interesting targets we chose to pursue, CellOracle alerted us to the potential involvement of *Fos* and *Jun* as well. As previously mentioned, *Fos* and *Jun* are part of the AP-1 Complex and are widely activated and expressed across systems. AP-1 is activated by a number of stimuli including growth factors and stress, making it a likely responder post intestinal resection.[202] In fact, it has been found that the AP-1 complex is an "immediate early" response post intestinal ischemia, meaning expression is immediately increased.[212,213] Because of this immediate early response and the general roles attributed to AP-1, we chose to forgo *Fos* and *Jun* as initial targets for direct reprogramming as we were interested in looking for proximal-related transcription factors with sustained expression to maintain cellular identity changes after resection. However, it is possible that the immediate early response of AP-1 to the altered growth factor and stress signals after resection is critical to the full regional reprogramming response, and we may need to target these in combination with other transcription factors in the future to achieve regional reprogramming *in vitro*.

We initially chose to only pursue *Gata4* based on the similarities between proximal and distal gene expression patterns in *Gata4* overexpressing organoids and proximal controls via qPCR (Figure 4.6*B*). However, both *Creb3l3* and *Klf4* are still promising targets that merit investigation. In the qPCR results from Figure 4.6*B*, *Creb3l3* and *Klf4* both indicate potential

loss of some distal enterocyte identity based on *Fabp6* expression decreasing, but it is worth noting that fold change of genes detected in both populations is relatively small. We don't believe these results indicate that *Creb3l3* and *Klf4* are not players in regional reprogramming. It is possible that these factors do not directly affect the downstream targets we measured or that they require the assistance of other factors to induce the expected downstream effects. Ultimately we conclude that *Creb3l3* and *Klf4* merit exploration based on the CellOracle results and their reported functions from the literature. It is likely, however, that they do not act alone in induction of regional reprogramming, just as *Gata4* is unlikely to act alone. Testing of direct reprogramming with combinations of transcription factors will be an important future direction (discussed in Section 5.2).

Our use of scRNA-seq analysis indicated that *Gata4* overexpression is insufficient to induce proximal enterocyte identity in distal organoids but is sufficient to induce cellular composition changes affecting differentiation. We saw an overall increase in stem cells and decrease in enterocytes in the *Gata4* samples. This was surprising, as while it is predicted that *Gata4* acts in intestinal stem cells to promote differentiation of enterocytes, it is not reported that it promotes stem cell identities.[205] However, we noted that *Olfm4,* a marker of intestinal stem cells (as well as tumorigenesis), was significantly upregulated in the *Gata4* samples compared to the empty controls via scRNA-seq (avg_log2FC = 0.67, P = 1.11E-33).[214–216] This increase in *Olfm4* expression in the *Gata4* populations supports our preliminary findings of expansion of stem cells after *Gata4* overexpression. In an interesting connection between the effects of *Gata4* and the chimera in Chapter 3, we noted that expression of both *Guca2a* (avg_log2FC = -2.06, P = 2.14 × 10^{-49}) and *Guca2b* (avg_log2FC = -1.17, P = 4.04 × 10^{-18}) was downregulated when comparing *Gata4* to empty controls—a known hallmark of tumorigenesis in the colon.[217] Many

of the other highly upregulated genes in *Gata4* populations compared to empty controls were associated with increased risk of intestinal and gastric cancers (e.g.: *Pgk1, Mif1, Ldha, Fabp5,* and *Aldoa*), indicating that *Gata4* overexpression has a large effect on cellular metabolism, differentiation, and proliferation in this system and may need to be carefully controlled during reprogramming to avoid cancerous cell transformations.[218–222]

We were additionally surprised that there was no detectable acquisition of proximal identity in the *Gata4* organoids based on the reported function of *Gata4* in maintenance of jejunal identity.[204] We postulate that there are several factors that could be playing into this contrast to the literature. Most of the studies in the literature focus on deletion or knock-in of *Gata4,* not overexpression. It is possible that our use of overexpression elevated the expression of *Gata4* far beyond normal levels, inducing unexpected changes in the system. In this vein, we must bring up the identity of *Gata4* as a pioneer factor as a consideration. Pioneer factors are a unique class of transcription factor capable of binding to inaccessible DNA sequences and opening them up to make them accessible to other transcription factors and further recruiting such factors.[72,223] It is additionally proposed that pioneer factors, when activated, activate gene networks associated with early developmental patterning, a conclusion that makes sense based on *Gata4*'s known role in regional patterning of the intestine. Overexpression of pioneer factors is proposed to allow "promiscuous" binding of transcription factors to chromatin, which could additionally account for our unexpected results due to off target effects and the host cell being unable to control the pioneer factors accurately.[72] We hypothesize that in maintenance of jejunal identity, *Gata4* acts as a pioneer factor, binding to the chromatin in the *in vivo* system and allowing the activity of other transcription factors to express proximal enterocyte identities, potentially through encouraging cells to pass back through a less mature state and then recommit.

By overexpressing *Gata4* in our *in vitro* system, we may have caused large perturbations to high-level circuits regulating fate and maturity, explaining the immaturity of our cell population. It is possible that *Gata4* must be expressed in a much more controlled manner to achieve the intended perturbations and to achieve the intended transcription factor binding. As it is likely that *Gata4* requires the synchronized expression of other transcription factors to exert its effects, it is likely that *Gata4* may need to be combined with other factors and carefully controlled during direct reprogramming to elicit regional reprogramming.

As in Chapter 3, Capybara analysis revealed unusually high stem cell abundance in all three samples. We again attribute this lack of maturity to the abundance of small molecules in the growth media. Downstream use of the proposed organoid models will require additional optimization of the media formulation to promote a balance of maturity and proliferative capabilities when continuing experiments.

We also noted in the Capybara analysis that there was an unexpectedly high number of cells identified as mature proximal enterocytes in the empty distal controls (Figure 4.8*A*) as well as the distal controls from the chimera experiment (Figure 3.7). As these organoids (two separate lines) were derived from terminal ileum, this finding was unusual, especially based on the fact that in the qPCR analyses in Chapter 4, the proximal controls showed increased levels of markers of proximal enterocyte identity compared to the empty distal controls via fold change calculation (Figure 4.6*B*). This result may indicate that further curation of the reference dataset for identification of proximal enterocytes is necessary as some of the markers may be less specific than expected. When investigating this possibility, we noted that the mature proximal enterocyte score generated by Capybara for the ileum in the reference dataset was somewhat high, which could account for the higher than expected detection in our samples as they are compared to the

reference. Conversely, this result may indicate that distal organoids grown *in vitro* for extended periods undergo a more proximal-like enterocyte conversion due to *in vitro* conditions. While the literature reports maintenance of regional identity *in vitro*, we know culture conditions (media) can affect cell identity, and perhaps this is playing a role in the aberrant appearance of mature proximal enterocytes in distal controls. Future experiments should look for the presence of cells identified as mature proximal enterocytes from freshly isolated intestinal crypts and utilize additional reference datasets for Capybara to explore these possibilities.[100]

Technical difficulties in achieving quality cell multiplexing and demultiplexing resulted in low levels of cells attributed to each sample. Due to this, the analysis is somewhat underpowered and preliminary. Findings reported should be further supported and analyzed by additional replicates of the study to increase cell number and power and confidence of the analysis.

Overall, our direct reprogramming organoid model of regional reprogramming provides valuable information on the effects of transcription factor overexpression on cells of the intestinal epithelium *in vitro*. Investigation of *Gata4* using this system provided evidence that *Gata4* alone is not responsible for regional reprogramming in the mouse epithelium. Based on our preliminary analysis, overexpression of pioneer factor *Gata4* induced a more stem-like identity in cells. Further study utilizing the *Gata4* model should investigate titration of *Gata4* levels to attempt to better induce on-target effects. Based on our findings, future efforts should optimize the culture conditions for our *in vitro* model such that cellular composition is distributed in more normal ratios and to promote further differentiation. Finally, future work should focus on the combined effects of transcription factors in inducing regional reprogramming, as combinatorial effects are more likely to drive this complex response. By

further optimizing and developing the direct reprogramming organoid model, we provide a system in which the direct effects of transcription factor signaling can be investigated. This will allow for identification of the factors involved in the process in mouse and better inform our study of human adaptation through comparing the factors involved between species. Through doing so, we may identify factors inactive in humans that, if activated, could promote adaptation. When the cocktail of factors required for mouse regional reprogramming is determined, the direct reprogramming model will provide a valuable model of regional reprogramming in a dish for further study and can be utilized to test activity of signals in patient tissue to predict patient response to resection.

4.9 Conclusions

Here, we have set up a system for using direct reprogramming of intestinal organoids to test the effects of specific transcription factors on regional reprogramming. Using this system, we analyzed the contribution of *Gata4* to regional reprogramming *in vitro* using scRNA-seq. Our analysis of the *Gata4* overexpression system indicates that *Gata4* alone is not sufficient to induce regional reprogramming *in vitro*. Regardless, the continued study of the direct reprogramming system indicates a role for *Gata4* in cell differentiation and dedifferentiation in the organoid system. It is possible that *Gata4* plays a role in regional reprogramming, but is unlikely to act alone. The direct reprogramming model can be further used to investigate additional transcription factors for roles in regional reprogramming as well as the effects of combinations of transcription factors. Further delineation of the combination of transcription factors at play can result in creation of a new model of regional reprogramming *in vitro*.

4.10 Materials and Methods

4.10.1 Cell Culture

HUVEC cells were obtained from Lonza (CC-2519) and maintained according to manufacturer's instructions at 37°C, 5% CO_2 in an incubator with EGM-Plus Endothelial Cell Growth Media (Lonza CC-5035).

4.10.2 Organoid Maintenance

I-CreERT2 regional mouse organoids were derived and provided by the Digestive Diseases Research Core Center (DDRCC) Precision Animal Models and Organoids Core (PAMOC) at Washington University in St. Louis Medical School. The organoids were maintained as previously described in the StemCell Technologies Technical Bulletin with modifications.[187] Briefly, organoids were plated in 48-well cell culture plates (VWR 10062-898) in 20 µL droplets consisting of 100% Growth-Factor Reduced, Phenol Red-free Matrigel (Corning 356231). Droplets were polymerized for 10 minutes at 37°C, 5% CO_2 in an incubator before 250 µL of Mouse IntestiCult Organoid Growth Media (StemCell Technologies #06005)was added. Organoids were maintained at 37°C, 5% CO_2 in an incubator. Organoids were passaged as described in the StemCell Technologies Technical Bulletin.[187] Briefly, Matrigel domes were broken down using Gentle Cell Dissociation Reagent (StemCell Technologies #07174) and gentle rocking at room temperature for 10 minutes. Cells were washed and centrifuged 2x before replating in 20 µL droplets consisting of 100% Growth-Factor Reduced, Phenol Red-free Matrigel (Corning 356231) in 250 µL of IntestiCult Organoid Growth Media (StemCell Technologies #06005).

4.10.3 Cloning of Genes

Genes of interest for HIIP reprograming (*CDX2, GATA6, HNF4A,* and *FOXA3*) were obtained as a gift from the Atsushi Suzuki Lab and were in the pGCDNsam-IRES-eGFP backbone (Miura & Suzuki, 2017). Regional reprogramming genes of interest (*Gata4* (gift from Wenjun Kong, Samantha Morris Lab)*, Klf4* (gift from Barak Cohen, Addgene plasmid #66656)*, and *Creb3l3* (gift from Dong-Yan Jin, Addgene plasmid #99509)) were cloned into backbone pMSCV-loxp-dsRed-loxp-eGFP-Puro-WPRE (gift from Hans Clevers, Addgene plasmid #32702) using restriction enzyme HpaI (New England BioLabs R0105S). The intervening fusion-protein sequence between the gene insertion site and the eGFP sequence was replaced with a t2a sequence to avoid complications of gene fusion on folding.

4.10.4 Retrovirus Production

Retrovirus for both HIIPs and regional reprogramming experiments was produced in 293T-17 cells (ATCC CRL-11268) and packaged with pCL-Ampho (HIIPs) (Novus Biologicals NBP2-29541) or pCL-Eco (regional reprogramming) (Novus Biologicals NBP2-29540) using X-tremeGENE 9 DNA Transfection Reagent (Sigma Aldrich XTG9-RO). Retrovirus for HIIP reprogramming was collected 48 and 72 hours after transfection, filtered through a 0.45 µm filter and concentrated using 50% PEG solution before application to cells. Retrovirus for regional reprogramming was collected 72 hours after transfection, filtered through a 0.45 µm filter and concentrated by centrifugation at 4500xg for 16-24 hours at 4°C before application to cells.

4.10.5 HIIP Transduction

HIIPs were transduced with retrovirus (production methodology in Subsection 4.10.4) using an adaptation of the methods described in Miura & Suzuki, 2017.[80] Briefly, HUVEC cells were grown to 20-30% confluence in a 6-well plate. Every 12 hours for a total of 8 transductions, the endothelial growth medium was replaced with 1.86 mL media, 25 µL of each virus was added, and 4 µL of 500x protamine sulfate was added. 12 hours after the final (8th) transduction, cells were trypsinized (TrypLE Express (Gibco 12605036)) and resuspended at a concentration of $1x10^5$ cells/25 µL of Matrigel (Corning 543234) and plated as a 25 µL droplet in the wells of a 48-well culture plate. Matrigel was allowed to polymerize for 10 minutes at 37°C, 5% CO_2 in an incubator before medium was added and cells cultured for 20-21 days before passaging or analysis at 37°C, 5% CO_2 in an incubator. For days 0-12, media consisted of a 4:1 ratio of endothelial media:WCENR media. For days 12-16, media consisted of a 2:3 ratio of endothelial media:WCENR media. For the final days 16-21, media consisted of 100% WCENR.

WCENR media consists of: Advanced DMEM/F-12 (Gibco 12634010), 1x B27 (Gibco 17504044), 1x N-2 (Gibco 17502048), 1 mM N-acetylcysteine (Sigma-Aldrich A9165), 2mM GlutaMax 100x (Gibco 35050061), 10 mM HEPES 1M (Gibco 15630106), 1x penicillin/streptomycin, 10 µM Y-27632 (Sigma Aldrich Y0503), 0.5 µM A8301 (Tocris 2939), 10 µM SB202190 (Sigma Aldrich S7067), 10 mM Nicotinamide (Sigma Aldrich), 10 nM [Leu[15]]-Gastrin1 (Sigma Aldrich G9145), 2.5 µM prostaglandin E2 (Sigma Aldric P5640), 50 ng/mL human recombinant EGF (Sigma Aldrich E9644), 100 ng/mL human recombinant Noggin (PeproTech 250-38), 1 mg/mL human recombinant R-spondin1 (Miltenyi Biotec 130-114-824), 100 ng/mL human recombinant Wnt3a (R&D Systems 5036-WN), and 3 µM CHIR99021 (Cayman Chemicals 13122).

4.10.6 Regional Reprogramming Transduction and Induction

Villin-CreERT2 ileal organoids were transduced with retrovirus (production methodology in Subsection 4.10.4) using an adaptation of the methods described in Koo et al., 2012; Koo, Sasselli, & Clevers, 2013; and Andersson-Rolf et al., 2014.[43,224,225] Briefly, 2 days prior to transduction, organoids were passaged and cultured in 20 µL droplets of Growth-Factor Reduced, Phenol Red-free Matrigel (Corning 356231) in 250 µL IntestiCult Organoid Growth Medium (StemCell Technologies #06005) supplemented with 5 µM CHIR99021 (Cayman Chemicals 13122) and 10 mM Nicotinamide (Sigma Aldrich) to expand stem cells. On the day of transduction, organoids were transferred from the plate and mechanically disrupted by pipetting up and down 30-50 times using a P-200 pipette tip. Cells were centrifuged 5 minutes at 900xg at 4°C. The pellet was resuspended in 500 µL TrypLE Express (Gibco 12605036) and incubated at 37°C for 5 minutes until small cell fragments formed. 500 µL of IntestiCult (StemCell Technologies #06005) was added and the cells centrifuged 5 minutes at 900xg at 4°C. To infect, the prepared organoid cells were combined with the retroviral suspension (concentrated retrovirus in IntestiCult (StemCell Technologies #06005) with 5 µM CHIR99021 (Cayman Chemicals 13122), 10 mM Nicotinamide (Sigma Aldrich), 10 µM Y-27632 (Sigma Aldrich Y0503), and 10 mg/mL protamine sulfate) in wells of a 48-well culture plate. The cells and virus were mixed well using a P-1000 pipette tip then the plate sealed with Parafilm (Sigma Aldrich P7793) before centrifugation for 60 minutes at 600xg at 32°C ("spinoculation"). After centrifugation, the Parafilm was discarded and the plate incubated for 6 hours at 37°C, 5% CO_2 in an incubator. After 6 hours, the cells were collected and centrifuged for 5 minutes at 900xg at 4°C. The infection medium was then aspirated, and the pellet left on ice for 5 minutes. Cells were resuspended in 100 µL of Growth-Factor Reduced, Phenol Red-free Matrigel (Corning 356231)

and plated in 25 µL droplets in the wells of a 48-well plate. The droplets were allowed to polymerize for 10 minutes at 37°C, 5% CO_2 in an incubator before addition of 250 µL of infection medium (IntestiCult (StemCell Technologies #06005) with 5 µM CHIR99021 (Cayman Chemicals 13122), 10 mM Nicotinamide (Sigma Aldrich), and 10 µM Y-27632 (Sigma Aldrich Y0503). After 2-3 days of culture, medium was changed to pure IntestiCult, and we introduced 10 µL/mL puromycin (Sigma-Aldrich P8833) to select for successfully transduced organoids and continued to passage and culture transduced organoids normally with puromycin selection until use.

To induce overexpression of genes of interest, organoids were passaged and replated in growth media + 10 µL/mL puromycin (excluding puromycin in non-transduced controls) for 10x experiment. 1 µL of vehicle was added to proximal control wells and 1 µM Tamoxifen ((Z)-4-hydroxy Tamoxifen) (Cayman Chemical 14854) was added to empty control and experimental wells. After 48 hours of treatment, cells were monitored for appearance of eGFP as the dsRed is cut out through recombination out and the gene overexpression induced. Media was refreshed as needed. After 5 days, cells and controls were collected for scRNA-seq. Growth media consists of: 21.75 µL basal media, 0.5 mL B27 (Gibco 17504044), 250 µL N-2 (Gibco 17502048), 65 µL 500 mM N-acetylcysteine (Sigma-Aldrich A9165), 12.5 µL 100 µg/mL EGF, 25 µL 100 µg/mL stock mouse Noggin (PeproTech 250-38), 25 µL 100 µg/mL stock R-spondin 1 (R&D Systems 7150-RS0010/CF). Basal media consists of 48.5 mL Advanced DMEM:F12 (Gibco 12634010), 0.5 mL GlutaMax 100x (Gibco 35050061), 0.5 mL HEPES 1M (Gibco 15630106), 0.5 mL penicillin/streptomycin.

4.10.7 Quantitative PCR

HIIPs and organoids for qPCR were collected and prepared as follows. Each Matrigel dome was scraped and dissociated by pipetting before pipetting into 15 mL tubes. Samples were centrifuged at 500xg for 3 minutes. Media was carefully aspirated, and RNA was collected using the RNeasy Mini Kit (QIAGEN 74104). cDNA was synthesized from equal amounts RNA using the Maxima First Strand cDNA Synthesis Kit (Thermo Scientific K1641). cDNA was amplified using TaqMan Gene Expression Master Mix (Applied Biosystems 4369016) and the primer probes on the StepOnePlus Real-Time PCR System (Applied Biosystems 4376600) using StepOne Software v2.3. Primer probes were: *SOX9* (Hs00165814_m1)*, KLF5* (Hs00156145_m1)*, LGR5* (Hs00969422_m1)*, VIL1* (Hs01031724_m1)*, GAPDH* (Hs02758991_g1)*, Creb3l3* (Mm00520279_m1), *Gata4* (Mm00484689_m1), *Klf4 (*Mm00516104_m1), *Apoa1* (Mm00437569_m1)*, Apoc3* (Mm00445670_m1)*, Fabp1* (Mm00444340_m1)*, Fabp6* (Mm00434315_m1)*, Hprt* (Mm03024075_m1), and *Rbp2* (Mm00436300_m1) all from Applied Biosystems.

4.10.8 Organoid Collection for scRNA-seq

Controls and experimental organoids for scRNA-seq were collected and prepared as follows. Each control or experimental Matrigel dome was scraped and dissociated by pipetting before pipetting into 15 mL tubes. Samples were centrifuged at 500xg for 3 minutes. Media was carefully aspirated and 3 mL TrypLE Express (Gibco 12605036) was added. Organoids were agitated through pipetting up and down 20-30x with a p200 pipette and placed in a 37°C bath. Every 5 minutes, cells were agitated via pipetting 20-30x and checked for breakdown to single cells. When cells achieved single-cell state, 5 mL of DMEM:F12 (Gibco) was added, and the cells were filtered through a 40 µm FlowMi strainer (Bel-Art H136800040). Cells were counted

and viability checked using Trypan Blue exclusion (viability > 90% required). Final samples were centrifuged at 500xg for 3 minutes at 4°C.

4.10.9 scRNA-seq Library Preparation

For single-cell library preparation on the 10x Genomics Chromium platform, we used the 3'CellPlex Kit Set A (PN-1000261), Chromium Next GEM Single Cell 3' Kit v3.1 (PN-1000268), 3' Feature Barcode Kit (PN-1000262), Chromium Next GEM Chip G Single Cell Kit (PN-1000120), and Dual Index Kit NN, Set A (PN-1000243), according to the manufacturer's instructions in the Chromium Next GEM Single Cell 3' v3.1 Cell Surface Protein User Guide (Rev. B). Resulting cDNA libraries were quantified on an Agilent Tapestation and sequenced on an Illumina NextSeq500.

4.10.10 scRNA-seq Analysis

The Cell Ranger 6.0 pipeline was used to align reads to the mm10 genome build, demultiplex at a confidence level of 0.75, and generate a digital gene expression (DGE) matrix: (https://support.10xgenomics.com/single-cell-gene-expression/software/downloads/latest). For initial filtering of these DGE matrices, we first excluded cells with a low number (<200) of unique detected genes. We then excluded cells for which the total number of unique molecules (UMIs) (after $\log_{10}$ transformation) was not within 3 standard deviations of the mean. This was followed by the exclusion of outlying cells with an unusually high or low number of UMIs/genes given their number of reads by fitting a Loess curve (span = 0.5, degree = 2) to the number of UMIs/genes with number of reads as predictor (after $\log_{10}$ transformation), removing cells with a residual more than 3 standard deviations the mean. Finally, we excluded cells in which the proportion of the UMI count attributable to mitochondrial genes was >15%. After filtering and normalization of the DGE, the R package Seurat (Version 4) was used to cluster and analyze the

single-cell transcriptomes.[102] Highly variable genes were identified and used as input for dimensionality followed by clustering and visualization via UMAP.[103] CellOracle was used for Gene Regulatory Network analysis and overexpression simulations.[101] R package Capybara was used for cell type identification and multi-id cell type identification.[100]

4.10.11 Imaging

Bright field images were taken on a Nikon Eclipse Ts2 with Nikon DS-Fi2 camera

4.10.12 Statistical Analysis

Statistical comparison of the qPCR datasets was performed using an unpaired Student's t test. Graphs with error bars represent fold change $\pm$ SD. $P < 0.05$ was considered significant. Figure 4.2 statistics are based on the mean delta CT values (n = 2 technical replicates) between controls and HIIPs (LGR5 is a separate experiment). Figure 4.5 statistics are based on the mean delta CT values (n = 3 technical replicates) between groups. Figure 4.6 statistics are based on the mean delta CT values (n = 3 technical replicates) between *Gata4* and proximal control treatments, only this comparison is included for space constraints.

4.11 Acknowledgements

Animal Models and Organoids Core (PAMOC) at Washington University in St. Louis Medical School. Special thanks go to Kunal Jindal for assistance in data alignment and analysis, and to Wenjun Kong for assistance in data analysis.

4.12 Authors' Contributions

Sarah Elizabeth Waye: study concept and design, data acquisition, analysis, interpretation, figure preparation; Samantha A. Morris: study concept and design, obtained funding, provided overall study supervision.

Chapter 5: Conclusions and Future Directions

5.1 Conclusions

In conclusion, this dissertation utilized scRNA-seq alongside computational and experimental analysis to contribute to the field of work surrounding the study of Short Bowel Syndrome and molecular-level adaptation after small bowel resection. The findings of this work, both preliminary and published, greatly advance the available information regarding cellular changes after adaptation and provide future researchers with two modeling tools for use in continued study and discovery in SBS. These findings additionally lead the author to pose many new questions and research directions for continued advancement of the field and this research.

In this work we addressed three major questions, the first of which concerned the molecular changes associated with structural adaptation after SBR on a single cell level. To approach this question, we utilized scRNA-seq to dissect cellular identities after resection on an established *in vivo* model of SBR. We identified the novel occurrence of a proximalization of a population of enterocytes of the remnant distal intestine which we deemed "regional reprograming". This regional reprogramming process consists of the upregulation of a variety of nutrient and lipid transporters and markers typically associated with the proximal bowel. Our scRNA-seq findings regarding changes in transporters such as *Fabp1* in these cells were well supported by previous literature reports of bulk gene expression changes after resection and we performed additional validations of these findings using *in situ* detection methods.[183] Based on these results, we hypothesized that regional reprogramming is a functional compensatory effect of adaptation. While structural adaptation compensates directly for loss of intestinal surface area, regional reprogramming compensates for the loss of expression of critical nutrient and lipid

transporters in order to maintain regional functionality of the bowel in nutrient absorption. The elucidation of the regional reprogramming process accompanying structural adaptation answers our initial question through defining distinct molecular changes in cell populations associated with the adaptation process. The discovery of regional reprogramming and the accompanying changes in gene expression provide researchers in the field with novel markers and pathways of mouse adaptation to study for potential applications in human SBR treatments to encourage adaptation.

Our interest in regional reprogramming and the molecular side of adaptation did not end with determining molecular effects of structural adaptation, but also included diving into the scRNA-seq data to answer our second question of how we could apply scRNA-seq datasets to determine potential upstream drivers of adaptation and regional reprogramming. While structural adaptation is well described, the actual pathways inducing adaptive changes (structural adaptation and regional reprogramming) are ill-defined. The determination of upstream drivers can identify specific triggers of successful adaptation which could be used to encourage adaptation in humans. We applied *in silico* interactome analysis to our SBR dataset and determined that Retinoic Acid metabolism and signaling was a likely driver of adaptation based on our downstream gene activation. This finding was well supported by literature reports of the effects of exogenous RA on structural adaptation.[183]

Our analysis in Chapter 2 indicated an upregulation of transcription factors *Creb3l3* and *Klf4* after adaptation, sparking interest in upstream transcription factor drivers of regional reprogramming that could be used *in vitro* to help model regional reprogramming. In subsequent studies into modeling of regional reprogramming *in vitro*, we again turned to *in silico* analysis to assist in picking out transcription factors likely to play a role in the process. We utilized

CellOracle to identify potential upstream transcription factors regulating regional reprogramming, pursuing *Klf4* and *Creb3l3*.[101] Using the scRNA-seq findings that there was increased proximal enterocyte identity after SBR in our Chapter 2 dataset, we identified *Gata4* as a potential driver of regional reprogramming from the literature. Through these methods, we combined *in silico* analysis of scRNA-seq data with our previous findings to identify potential upstream drivers of regional reprogramming for further study, addressing our second question and providing the field with novel likely candidates for induction of regional reprogramming in mice.

Our *in silico* findings regarding RA signaling and upstream transcription factors led our final drive to generate models of regional reprogramming *in vitro* for use in future study of regional reprogramming. In doing so, we approached our final question of how we could model adaptive responses using organoids, cellular reprogramming, and scRNA-seq analysis. In following up on the findings indicating a role for RA in regional reprogramming, we hypothesized that disruption of an endogenous Vitamin A/RA gradient induced regional reprogramming of the distal bowel though exposure to higher levels of Vitamin A/RA.[183] We used this hypothesis to build our chimeric organoid model in which proximal and distal intestinal organoids are combined to create a unique signaling environment between proximal and distal tissue capable of reprogramming distal enterocytes through exposure to signaling factor gradients.

We again utilized scRNA-seq to analyze the outcomes of this modeling system. In doing so, we found preliminarily that the exposure of distal intestinal organoid tissue to a proximal signaling environment through chimerization activates more proximal identities in distal cells (detected by assignment of computational proximalization scores and cellular identities via

Capybara).[100] In doing so, we uniquely combined scRNA-seq analysis and cellular reprogramming of organoids to generate a preliminary *in vitro* model of regional reprogramming that can greatly increase the ability of labs to study regional reprogramming because of the departure from an *in vivo* system.

We proposed utility for this model through using further computational analysis (CellPhoneDB) on the scRNA-seq data.[107] This analysis indicates a role for GC-C signaling during chimerization-related regional reprogramming. This pathway has not previously been associated with regional reprogramming and provides a future target for study and treatment and highlights the utility of the chimeric model for discovery of novel pathways activated in regional reprogramming.

We even further developed our modeling capabilities through applying our findings on upstream transcription factor activation through use of direct reprogramming on distal organoids. In doing so, we studied the effects of transcription factor overexpression in order to better delineate effects of different transcription factors during the regional reprogramming process. In utilizing *Gata4* for direct reprogramming, we found that *Gata4* alone was insufficient to induce proximalization, but our findings led us to hypothesize that combinations of transcription factors will produce the intended effect. Again, Capybara and proximalization scores were used to identify proximalization or regional reprogramming in this system, highlighting how scRNA-seq datasets can be used for high-level analysis of *in vitro* organoid and reprogramming systems.

Overall, our findings advance the field and promote future, more well informed studies of the response of the bowel after resection. In Chapter 2, we identified a process inducing molecular changes in adaptation–regional reprogramming–and mined our scRNA-seq dataset to identify potential upstream drivers of regional reprogramming. These findings will assist in the

further expansion of knowledge regarding successful adaptation in mice and can better inform us of mouse triggers that could be employed to induce adaptation in human systems. In addition, the scRNA-seq dataset we built of SBR is available online and can be mined by other labs for new information. In Chapter 3, we introduced a preliminary model for *in vitro* regional reprogramming and downstream discovery of pathways activated in regional reprogramming, the chimera. This model can greatly contribute to the prediction of human response to resection through using patient organoids for chimerization and measuring their endogenous response and ability to proximalize. Ultimately this can better target treatments for specific patients. It also provides a system to discover pathways affected by regional reprogramming for future study, providing valuable targets for future treatment of SBS. Finally, in Chapter 4, we coupled direct reprogramming with scRNA-seq analysis to provide a system capable of providing insight into the effects of transcription factor signaling on regional reprogramming. This system will prove valuable for testing which transcription factors are playing a role in successful mouse adaptation and also can inform clinicians of the effects and potential utility of those transcription factors on inducing regional reprogramming in humans through use of patient-derived organoids in the direct reprogramming.

While our study provided us with new findings, models, and methods for investigating and addressing SBS, it also piques interest in many follow-up studies and experiments. Some questions we believe should be addressed in future studies include: How can we further mature organoids *in vitro*? How can we improve the power of our scRNA-seq experiments? Is RA signaling upstream of the chimeric changes? Can we use functional testing to validate regional reprogramming in organoids and chimeras? What is the role of GC-C signaling in regional reprogramming? Is proximalization accompanied by reciprocal distalization of remnant proximal

tissue? What are the effects of *Creb3l3* and *Klf4* on distal organoids? And finally, how can we use direct reprogramming to test combinations of transcription factors for induction of regional reprogramming? In our final section, Section 5.2, we will propose experiments and methodologies to approach these questions in the future.

5.2 Future Directions

The results and conclusions presented in this dissertation lay the groundwork for many additional studies to follow up on and expand on the findings regarding regional reprogramming and SBR. In this final section, we briefly outline future experimental approaches for pursuit of these questions as are justified by our findings.

How can we further mature organoids in vitro? In Chapters 3 and 4 we utilized organoid systems for modeling regional reprogramming *in vitro*, however, as discussed, we struggled with formulating a media to promote differentiation of stem cells in the organoids. This resulted in a high incidence of stem cells in all cultures. While expansion of stem cells is necessary to culture organoids effectively, encouraging differentiation for endpoint experiments is justified as it will better inform researchers about changes to end-point cellular composition in the organoid models we present by enlarging the population of fully differentiated cells.

To approach this issue, we propose testing and formulation of multiple media conditions to modulate the Wnt, Notch, and BMP pathways, as these are reported to control stem cell renewal and differentiation in organoids. We suggest modulating the culture concentrations of R-spondin 1, Noggin, and Wnt3a to promote differentiation of cells.[194] In particular, the concentrations of R-spondin 1 and Noggin should be titrated down to promote differentiation but still be present to promote stem cell proliferation for organoid growth. In order to determine the optimized culture conditions, qPCR should be performed as before to check for increase in

mature markers such as *Alpi* and decrease in stem markers such as *Lgr5* while organoid growth characteristics are monitored to maintain some cellular expansion. Once a promising formulation (or several) is identified, the resulting organoids should be analyzed for differentiated cell identities via scRNA-seq and for imaging to gain a full, consistent view of the population. This experiment is critical to future use of the organoid models as it will allow for the organoids to better recapitulate the cell distribution of the intestine and will optimize the models to best mimic the *in vivo* environment for *in vitro* testing for further use in the field.

How can we improve the power of our scRNA-seq experiments? A shortcoming of the experiments performed in Chapters 3 and 4 was underpowered analysis of scRNA-seq data due to small sample sizes being of high enough quality to analyze. The causes of this were due to failure to detect mT/mG cells and failure to demultiplex combined experiments. To improve the power of the findings, future work should make two major changes: 1) chimeras should switch from mT/mG organoids to UBC-GFP ad 2) cell multiplexing should be carefully titrated to increase demultiplexing output.

Chapter 3 mentioned the use of mT/mG organoids expressing tdTomato to detect organoid origins within a chimera. Unexpectedly, scRNA-seq analysis detected only low numbers of cells identifiable by tdTomato expression. Research into the topic suggests this is a common problem among users and may be due to the 3-D structure of tdTomato mRNA being difficult to sequence. This results in what we call "drop out" or the failure of scRNA-seq to detect expression of a gene even though it is expressed in the cell, often due to low levels of expression. Ultimately, this resulted in our analysis having very few distally-derived chimera cells to test and also meant that the remaining proximally-derived chimera cells could not be

reciprocally tested for identity changes due to the likelihood of presence of "dropped out" distal cells in their population.

Moving forward, the chimeric system must be optimized to avoid this detection issue. Distal organoids should be newly derived from a GFP expressing mouse, such as the UBC-GFP mouse, as GFP is well-detected by scRNA-seq. The chimeric experiments should be replicated using these GFP organoids instead of mT/mG in order to increase detected populations of distally derived organoids and allow for reciprocal proximal analyses.

In Chapter 3, R1, and Chapter 4, cell multiplexing methods were used to ameliorate potential batch effects in the datasets.[226] However, the multiplexing methods require more optimization to enhance cell capture. The 10x Genomics CellPlex kit was utilized but a majority of cells failed to be identified by their oligo-tag as the background detection was too high. According to 10x Genomics representatives, future experiments can lower this background level to allow for higher cell capture and identification by increasing wash steps after tagging of the cells. Optimization must be carried out to test the effects of additional washes on background detection and cell viability/quality such that the two can be balanced and enhance the datasets. By using GFP in the chimera and amending the multiplexing, future studies will increase sample sizes and the confidence of the analysis such that the results can be better interpreted and applied to the field and the chimeric model can be used and studied with greater confidence.

Is RA signaling upstream of the chimeric changes? In Chapter 2, we identified RA signaling as a likely upstream trigger of regional reprogramming. The actual presence of RA signaling in the upstream cascades merits further investigation so that it can be determined the role it truly plays. This question arises as RA signaling is reportedly related to many of our investigations. RA is a putative activator of *Creb3l3, Klf4*, *Gata4*, and *Cdx2*—known to regulate

140

Gucy2c expression.[147,227,228] Because of this potential role of RA signaling upstream of the three transcription factors of interest from direct reprogramming and the GC-C pathway activated by chimerization, we ask what role RA signaling plays in regional reprogramming during chimerization.

To address this question, we suggest utilizing RA inhibitors (inhibitors of RARs and RXRs) and administration of exogenous RA (all-trans Retinoic Acid and/or 9-cis Retinoic Acid) in organoid cultures during chimerization. After the growth period, we propose to follow with scRNA-seq to analyze changes in proximalization compared to traditional chimera models as well as look into expression of regional reprogramming related genes. If RA promotes regional reprogramming, we'd expect administration of exogenous RA to chimeras to enhance proximalization and gene expression while administration of inhibitors would lower proximalization and gene expression. By performing this follow-up, the lab can better determine how much of a role RA signaling is actually playing during the process. If RA is truly a high-level upstream activator of regional reprogramming as we propose, it could be used for future therapeutic advances for SBS, and the chimeric system could be used to titrate treatment.

Can we use functional testing to validate regional reprogramming in organoids and chimeras? The conclusions of this dissertation regarding the proximal identity of resulting organoids from chimerization or direct reprogramming relied on *in silico* modeling and analysis. We recommend additional validations be performed on these models to validate that the altered gene expression in the organoids relates to altered functions such as proximal nutrient transporter activity. Such validation will assist in determining how well chimerization or other reprogramming methods truly mimic regional reprogramming in a dish and their utility as models.

141

To investigate the functionality of the adaptive gene expression changes noted in chimeras, we propose use of a labeling and transport approach. We suggest that the lab labels various nutrients absorbed in the proximal intestine and monitors for increased transport in the organoid system. One method to do this would be to focus on FABP1 and FABP2, proximally expressed fatty acid binding proteins shown to increase after resection.[183] These studies could employ radiolabeled long chain fatty acids or BODIPY FL C16, a fluorescent fatty acid, to test for functionally increased transport of fatty acids after chimerization.[229] In doing so, the lab will validate and strengthen the chimeric model for use in the field, increasing its applicability to treatment-related studies as a *functional* model.

What is the role of GC-C signaling in regional reprogramming? The findings in Chapter 3 encourage future study the GC-C pathway for involvement in SBR and SBS—a previously unstudied and novel SBR-related pathway. As the *Gucy2c* discovery was first made in the preliminary chimera model using *in silico* analysis, the place to start looking at *Gucy2c* involvement in adaptation is to ask whether it is involved *in vivo*, then to investigate direct effects. First, we recommend using the available *Gucy2c* knockout mouse model for 50% SBR surgery as described and utilized in Chapter 2. Subsequently, scRNA-seq should be carried out on the tissue after the adaptation period and compared to SBR in normal mouse to investigate for regional reprogramming-related changes. If GC-C signaling increases play a role in the regional reprogramming response, we would expect that loss of *Gucy2c* would have a detrimental effect on detectable adaptation and regional reprogramming in the system. Investigation and validation of *Gucy2c* signaling and its potential role in regional reprogramming will advance knowledge of signals leading to regional reprogramming and gives way to potentially using modulations of such signaling pathways to enhance adaptive responses in humans.

142

Is proximalization accompanied by reciprocal distalization of remnant proximal tissue?
The CellPhoneDB investigation in Chapter 3 raised an interesting question due to the observation that proximally derived chimeric tissue showed higher levels of GUCA2A and GUCA2B expression compared to proximal controls.[107] As previously mentioned, analysis of the proximal portion of the chimera is confounded by the inclusion of dropped-out distal cells in the population, but these increased levels of GUCA2A and GUCA2B in proximal cells begs the question: does proximal intestine distalize via regional reprogramming in response to resection? It seems likely, as the hypothesis that gradients change due to disrupted intestinal length would affect proximal intestinal gradients as well.

To look into this possibility, we suggest to, in essence, repeat the work of Chapter 2, but focus on the proximal intestine. After the traditional SBR and adaptation period, tissue *proximal* to the anastomosis should be collected and analyzed for changes between sham and SBR mice. If the proximal tissue is distalizing, we would expect to see an upregulation in expression of distal-related transport genes in the cell populations, like *Fabp6*. This study would help generate a more well-rounded picture of the adaptive response in the intestine as a whole instead of focusing only on the distal intestine and could better inform how clinicians target future treatments and even re-design regional bowel resections such that the remnant regions are more likely to adapt.

What are the effects of Creb3l3 and Klf4 on distal organoids? Based on our findings in Chapters 2 and 4, we remain inquisitive about the roles that *Creb3l3* and *Klf4* play in the regional reprogramming process. We suggest that the lab continue use of the regional reprogramming organoid model to delineate effects of these transcription factors *in vitro*. However, we suggest that prior to these investigations, new organoid lines capable of being induced with minimal

leakage are derived.[208] To do so, the lab should source several organoid lines from different litters of mice expressing Cre-recombinase in intestinal cells (e.g.: *Villin*-CreERT2, *ROSA26*-CreERT2). The levels of Cre-recombinase in each line could be determined using Western Blotting to identify a less-highly expressed but still active Cre-recombinase line for use. Deriving new organoid lines will strengthen the direct reprogramming system for use as a temporal system of reprogramming. Combination of this system with *Creb3l3* and *Klf4* will provide us with greater knowledge of the roles of these upregulated transcription factors in regional reprogramming, extending the knowledge from Chapter 2.

How can we use direct reprogramming to test combinations of transcription factors for induction of regional reprogramming? Finally, our work implies that combinations of signals and factors are necessary to induce regional reprogramming. We seek to answer the major question: what combination of transcription factors induces regional reprogramming *in vitro*? Testing combinations of transcription factors used to rely on a classical "pool and pull" approach in which groups of factors were pooled and expressed and then removed one by one to determine effects.[70] The process was lengthy and prone to errors.

To overcome shortcomings of this process, we suggest future study utilize new technology from the lab, CellTags, to greatly enhance this approach. We suggest using the CellTag as a barcoded gene expression vector that can be used to detect what transcripts are expressed in a cell. For this experiment, the researchers should clone transcription factors of interest into individually barcoded CellTags, such that each transcription factor has a unique barcode. The CellTags can then be pooled and utilized for direct reprogramming of organoids at an MOI which encourages different combinations and amounts of viral particles infect each cell. Following reprogramming, the organoids should be collected and analyzed for regional

144

reprogramming-related changes via scRNA-seq. When cells are identified with promising changes, the barcodes present in that cell can be used to determine which transcription factors are expressed and involved in the identity changes. Further testing of these factors will lead to the optimal combination of factors for induction of regional reprogramming.[79] This experiment in particular provides much to the field of SBS research, it will not only greatly advance our knowledge of the upstream factors involved in regional reprogramming and their downstream effects for targeted treatments to promote adaptation in humans, but also will provide a complete *in vitro* model for investigation of regional reprogramming and modulation of the process.

In summation, the data and conclusions provided by this dissertation provide not only important findings for the field of study around SBS and bowel resection, but also provide the groundwork, models, and ideas for extensive future studies on regional reprogramming and adaptation. Continued research based on the provided future directions will greatly increase familiarity with what factors and pathways are driving regional reprogramming and adaptation, how we can optimize our modeling of this process, and ultimately will directly impact human health through promoting better, personalized testing and treatment of SBS.

References

1.   Watson CL, Mahe MM, Múnera J, et al. An in vivo model of human small intestine using pluripotent stem cells. *Nat Med*. 2014;20(11):1310-1314. doi:10.1038/nm.3737

2.   Gehart H, Clevers H. Tales from the crypt: new insights into intestinal stem cells. *Nat Rev Gastroenterol Hepatol*. 2019;16(1):19-34. doi:10.1038/s41575-018-0081-y

3.   Cheng H, Leblond CP. Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine V. Unitarian theory of the origin of the four epithelial cell types. *Am J Anat*. 1974;141(4):537-561. doi:10.1002/aja.1001410407

4.   Boonekamp KE, Dayton TL, Clevers H. Intestinal organoids as tools for enriching and studying specific and rare cell types: advances and future directions. *J Mol Cell Biol*. 2020;12(8):562-568. doi:10.1093/jmcb/mjaa034

5.   Date S, Sato T. Mini-Gut Organoids: Reconstitution of the Stem Cell Niche. *Annu Rev Cell Dev Biol*. 2015;31(1):269-289. doi:10.1146/annurev-cellbio-100814-125218

6.   Howell JC, Wells JM. Generating intestinal tissue from stem cells: potential for research and therapy. *Regen Med*. 2011;6(6):743-755. doi:10.2217/rme.11.90

7.   Sato T, Vries RG, Snippert HJ, et al. Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature*. 2009;459(7244):262-265. doi:10.1038/nature07935

8.   Beumer J, Clevers H. Cell fate specification and differentiation in the adult mammalian intestine. *Nat Rev Mol Cell Biol*. 2021;22(1):39-53. doi:10.1038/s41580-020-0278-0

9.   Tian H, Biehs B, Warming S, et al. A reserve stem cell population in small intestine renders Lgr5-positive cells dispensable. *Nature*. 2011;478(7368):255-259. doi:10.1038/nature10408

10. Snippert HJ, Schepers AG, van Es JH, Simons BD, Clevers H. Biased competition between Lgr5 intestinal stem cells driven by oncogenic mutation induces clonal expansion. *EMBO Rep*. 2014;15(1):62-69. doi:10.1002/embr.201337799

11. Schalamon J, Mayr JM, Höllwarth ME. Mortality and economics in short bowel syndrome. *Best Pract Res Clin Gastroenterol*. 2003;17(6):931-942. doi:10.1016/s1521-6918(03)00079-9

12. Wales PW, de Silva N, Kim J, Lecce L, To T, Moore A. Neonatal short bowel syndrome: population-based estimates of incidence and mortality rates. *J Pediatr Surg*. 2004;39(5):690-695. doi:10.1016/j.jpedsurg.2004.01.036

13.   Sommovilla J, Warner BW. Surgical Options to Enhance Intestinal Function in Patients with Short Bowel Syndrome. *Curr Opin Pediatr*. 2014;26(3):350-355. doi:10.1097/MOP.0000000000000103

14.   McMellen ME, Wakeman D, Longshore SW, McDuffie LA, Warner BW. Growth Factors: Possible Roles for Clinical Management of the Short Bowel Syndrome. *Semin Pediatr Surg*. 2010;19(1):35-43. doi:10.1053/j.sempedsurg.2009.11.010

15.   Thompson JS, Rochling FA, Weseman RA, Mercer DF. Current Management of Short Bowel Syndrome. *Curr Probl Surg*. 2012;49(2):52-115. doi:10.1067/j.cpsurg.2011.10.002

16.   Shirafkan A, Montalbano M, Mcguire J, Rastellini C, Cicalese L. New approaches to increase intestinal length: Methods used for intestinal regeneration and bioengineering. *World J Transpl*. 2016;206(69).

17.   Kitano K, Schwartz DM, Zhou H, et al. Bioengineering of functional human induced pluripotent stem cell-derived intestinal grafts. *Nat Commun*. 2017;8(1). doi:10.1038/s41467-017-00779-y

18.   Spurrier RG, Grikscheit TC. Tissue Engineering the Small Intestine. *Clin Gastroenterol Hepatol*. 2013;11(4):354-358. doi:10.1016/j.cgh.2013.01.028

19.   Evans GS, Flint N, Somers AS, Eyden B, Potten CS. The development of a method for the preparation of rat intestinal epithelial cell primary cultures. *J Cell Sci*. 1992;101 (Pt 1):219-231.

20.   Vacanti JP, Morse MA, Saltzman WM, Domb AJ, Perez-Atayde A, Langer R. Selective cell transplantation using bioabsorbable artificial polymers as matrices. *J Pediatr Surg*. 1988;23(1 Pt 2):3-9. doi:10.1016/s0022-3468(88)80529-3

21.   Meran L, Massie I, Campinoti S, et al. Engineering transplantable jejunal mucosal grafts using patient-derived organoids from children with intestinal failure. *Nat Med*. 2020;26(10):1593-1601. doi:10.1038/s41591-020-1024-z

22.   Clevers H, Conder RK, Li VSW, et al. Tissue-Engineering the Intestine: The Trials before the Trials. *Cell Stem Cell*. 2019;24(6):855-859. doi:10.1016/j.stem.2019.04.018

23.   Finkbeiner SR, Freeman JJ, Wieck MM, et al. Generation of tissue-engineered small intestine using embryonic stem cell-derived human intestinal organoids. *Biol Open*. 2015;4(11):1462-1472. doi:10.1242/bio.013235

24.   Sala FG, Matthews JA, Speer AL, Torashima Y, Barthel ER, Grikscheit TC. A multicellular approach forms a significant amount of tissue-engineered small intestine in the mouse. *Tissue Eng Part A*. 2011;17(13-14):1841-1850. doi:10.1089/ten.TEA.2010.0564

25.   Rubin DC, Levin MS. Mechanisms of Intestinal Adaptation. *Best Pract Res Clin Gastroenterol*. 2016;30(2):237-248. doi:10.1016/j.bpg.2016.03.007

26.   Schall KA, Holoyda KA, Grant CN, et al. Adult zebrafish intestine resection: a novel model of short bowel syndrome, adaptation, and intestinal stem cell regeneration. *Am J Physiol - Gastrointest Liver Physiol*. 2015;309(3):G135-G145. doi:10.1152/ajpgi.00311.2014

27.   Ziegler TR, Mantell MP, Chow JC, Rombeau JL, Smith RJ. Intestinal adaptation after extensive small bowel resection: differential changes in growth and insulin-like growth factor system messenger ribonucleic acids in jejunum and ileum. *Endocrinology*. 1998;139(7):3119-3126. doi:10.1210/endo.139.7.6097

28.   Warner BW. Adaptation: Paradigm for the Gut and an Academic Career. *J Pediatr Surg*. 2013;48(1):20-26. doi:10.1016/j.jpedsurg.2012.10.014

29.   Fowler KL, Wieck MM, Hilton AE, Hou X, Schlieve CR, Grikscheit TC. Marked stem/progenitor cell expansion occurs early after murine ileostomy: a new model. *J Surg Res*. 2017;220:182-196. doi:10.1016/j.jss.2017.06.079

30.   Taylor JA, Martin CA, Nair R, Guo J, Erwin CR, Warner BW. Lessons learned: optimization of a murine small bowel resection model. *J Pediatr Surg*. 2008;43(6):1018-1024. doi:10.1016/j.jpedsurg.2008.02.025

31.   Helmrath MA, VanderKolk WE, Can G, Erwin CR, Warner BW. Intestinal adaptation following massive small bowel resection in the mouse. *J Am Coll Surg*. 1996;183(5):441-449.

32.   Schall KA, Thornton ME, Isani M, et al. Short bowel syndrome results in increased gene expression associated with proliferation, inflammation, bile acid synthesis and immune system activation: RNA sequencing a zebrafish SBS model. *BMC Genomics*. 2017;18. doi:10.1186/s12864-016-3433-4

33.   Gillingham MB, Dahly EM, Murali SG, Ney DM. IGF-I treatment facilitates transition from parenteral to enteral nutrition in rats with short bowel syndrome. *Am J Physiol-Regul Integr Comp Physiol*. 2003;284(2):R363-R371. doi:10.1152/ajpregu.00247.2002

34.   Lund PK. Molecular Basis of Intestinal Adaptation: The Role of the Insulin-like Growth Factor System. *Ann N Y Acad Sci*. 1998;859(1):18-36. doi:10.1111/j.1749-6632.1998.tb11108.x

35.   Wang L, Tang Y, Rubin DC, Levin MS. Chronically administered retinoic acid has trophic effects in the rat small intestine and promotes adaptation in a resection model of short bowel syndrome. *Am J Physiol-Gastrointest Liver Physiol*. 2007;292(6):G1559-G1569. doi:10.1152/ajpgi.00567.2006

36.   Knott AW, Erwin CR, Profitt SA, Juno RJ, Warner BW. Localization of postresection EGF receptor expression using laser capture microdissection. *J Pediatr Surg*. 2003;38(3):440-445. doi:10.1053/jpsu.2003.50076

37.   Chaet MS, Arya G, Ziegler MM, Warner BW. Epidermal growth factor enhances intestinal adaptation after massive small bowel resection. *J Pediatr Surg*. 1994;29(8):1035-1039. doi:10.1016/0022-3468(94)90274-7

38.   Cohran VC, Prozialeck JD, Cole CR. Redefining short bowel syndrome in the 21st century. *Pediatr Res*. 2017;81(4):540-549. doi:10.1038/pr.2016.265

39.   Huch M, Knoblich JA, Lutolf MP, Martinez-Arias A. The hope and the hype of organoid research. *Development*. 2017;144(6):938-941. doi:10.1242/dev.150201

40.   Lancaster MA, Knoblich JA. Organogenesis in a dish: Modeling development and disease using organoid technologies. *Science*. 2014;345(6194):1247125-1247125. doi:10.1126/science.1247125

41.   Quadrato G, Nguyen T, Macosko EZ, et al. Cell diversity and network dynamics in photosensitive human brain organoids. *Nature*. 2017;545(7652):48-53. doi:10.1038/nature22047

42.   Dekkers JF, Wiegerinck CL, de Jonge HR, et al. A functional CFTR assay using primary cystic fibrosis intestinal organoids. *Nat Med*. 2013;19(7):939-945. doi:10.1038/nm.3201

43.   Koo B-K, Stange DE, Sato T, et al. Controlled gene expression in primary Lgr5 organoid cultures. *Nat Methods*. 2012;9(1):81-83. doi:10.1038/nmeth.1802

44.   O'Rourke KP, Ackerman S, Dow LE, Lowe SW. Isolation, Culture, and Maintenance of Mouse Intestinal Stem Cells. *Bio-Protoc*. 2016;6(4):e1733.

45.   Leushacke M, Barker N. Ex vivo culture of the intestinal epithelium: strategies and applications. *Gut*. 2014;63(8):1345-1354. doi:10.1136/gutjnl-2014-307204

46.   Frum T, Spence JR. hPSC-derived organoids: models of human development and disease. *J Mol Med*. 2021;99(4):463-473. doi:10.1007/s00109-020-01969-w

47.   Zachos NC, Kovbasnjuk O, Foulke-Abel J, et al. Human Enteroids/Colonoids and Intestinal Organoids Functionally Recapitulate Normal Intestinal Physiology and Pathophysiology. *J Biol Chem*. 2016;291(8):3759-3766. doi:10.1074/jbc.R114.635995

48.   Middendorp S, Schneeberger K, Wiegerinck CL, et al. Adult Stem Cells in the Small Intestine Are Intrinsically Programmed with Their Location-Specific Function. *STEM CELLS*. 2014;32(5):1083-1091. doi:10.1002/stem.1655

49.   Sato T, van Es JH, Snippert HJ, et al. Paneth cells constitute the niche for Lgr5 stem cells in intestinal crypts. *Nature*. 2011;469(7330):415-418. doi:10.1038/nature09637

50.   Sato T, Clevers H. Growing self-organizing mini-guts from a single intestinal stem cell: mechanism and applications. *Science*. 2013;340(6137):1190-1194. doi:10.1126/science.1234852

51. Nakamura T, Sato T. Advancing Intestinal Organoid Technology Toward Regenerative Medicine. *Cmgh*. 2018;5(1):51-60. doi:10.1016/j.jcmgh.2017.10.006

52. Fukuda M, Mizutani T, Mochizuki W. Small intestinal stem cell identity is maintained with functional Paneth cells in heterotopically grafted epithelium onto the colon Small intestinal stem cell identity is maintained with functional Paneth cells in heterotopically grafted epithelium onto t. Published online 2014:1752-1757. doi:10.1101/gad.245233.114

53. Yui S, Nakamura T, Sato T, et al. Functional engraftment of colon epithelium expanded in vitro from a single adult Lgr5 + stem cell. *Nat Med*. 2012;18(4):618-623. doi:10.1038/nm.2695

54. Ladd MR, Nino DF, March JC, Sodhi CP, Hackam DJ. Generation of an Artificial Intestine for the Management of Short Bowel Syndrome. *Curr Opin Organ Transpl*. 2016;21(2):129-139. doi:10.5588/ijtld.16.0716.Isoniazid

55. Lukonin I, Serra D, Meylan LC, et al. Phenotypic landscape of intestinal organoid regeneration. *Nature*. 2020;586(7828):275-280. doi:10.1038/s41586-020-2776-9

56. Blutt SE, Klein OD, Donowitz M, Shroyer N, Guha C, Estes MK. Use of organoids to study regenerative responses to intestinal damage. *Am J Physiol-Gastrointest Liver Physiol*. 2019;317(6):G845-G852. doi:10.1152/ajpgi.00346.2018

57. Serra D, Mayr U, Boni A, et al. Self-organization and symmetry breaking in intestinal organoid development. *Nature*. 2019;569(7754):66-72. doi:10.1038/s41586-019-1146-y

58. Wallach TE, Bayrer JR. Intestinal Organoids: New Frontiers in the Study of Intestinal Disease and Physiology. *J Pediatr Gastroenterol Nutr*. 2017;64(2):180-185. doi:10.1097/MPG.0000000000001411

59. Fair KL, Colquhoun J, Hannan NRF. Intestinal organoids for modelling intestinal development and disease. *Philos Trans R Soc B Biol Sci*. 2018;373(1750):20170217. doi:10.1098/rstb.2017.0217

60. Waddington CH, Kacser H. *The Strategy of the Genes: A Discussion of Some Aspects of Theoretical Biology*. Allen & Unwin; 1957.

61. Takahashi K. Cellular Reprogramming. *Cold Spring Harb Perspect Biol*. 2014;6(2):a018606. doi:10.1101/cshperspect.a018606

62. Gurdon JB, Elsdale TR, Fischberg M. Sexually Mature Individuals of Xenopus laevis from the Transplantation of Single Somatic Nuclei. *Nature*. 1958;182(4627):64-65. doi:10.1038/182064a0

63. Gurdon JB, Byrne JA. The first half-century of nuclear transplantation. *Proc Natl Acad Sci*. 2003;100(14):8048-8052. doi:10.1073/pnas.1337135100

64. Blau HM, Pavlath GK, Hardeman EC, et al. Plasticity of the differentiated state. *Science*. 1985;230(4727):758-766. doi:10.1126/science.2414846

65. Morris SA, Daley GQ. A blueprint for engineering cell fate: current technologies to reprogram cell identity. *Cell Res*. 2013;23(1):33-48. doi:10.1038/cr.2013.1

66. Spemann H, Mangold H. Induction of embryonic primordia by implantation of organizers from a different species. 1923. *Int J Dev Biol*. 2001;45(1):13-38.

67. Mascetti VL, Pedersen RA. Contributions of Mammalian Chimeras to Pluripotent Stem Cell Research. *Cell Stem Cell*. 2016;19(2):163-175. doi:10.1016/j.stem.2016.07.018

68. Xinaris C, Benedetti V, Rizzo P, et al. In vivo maturation of functional renal organoids formed from embryonic cell suspensions. *J Am Soc Nephrol JASN*. 2012;23(11):1857-1868. doi:10.1681/ASN.2012050505

69. Xinaris C, Benedetti V, Novelli R, et al. Functional Human Podocytes Generated in Organoids from Amniotic Fluid Stem Cells. *J Am Soc Nephrol*. 2016;27(5):1400-1411. doi:10.1681/ASN.2015030316

70. 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. doi:10.1016/j.cell.2006.07.024

71. Paige SL, Osugi T, Afanasiev OK, Pabon L, Reinecke H, Murry CE. Endogenous Wnt/beta-catenin signaling is required for cardiac differentiation in human embryonic stem cells. *PloS One*. 2010;5(6):e11134. doi:10.1371/journal.pone.0011134

72. Morris SA. Direct lineage reprogramming via pioneer factors; a detour through developmental gene regulatory networks. *Development*. 2016;143(15):2696-2705. doi:10.1242/dev.138263

73. Aydin B, Mazzoni EO. Cell Reprogramming: The Many Roads to Success. *Annu Rev Cell Dev Biol*. 2019;35(1):433-452. doi:10.1146/annurev-cellbio-100818-125127

74. Pevny L, Simon MC, Robertson E, et al. Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. *Nature*. 1991;349(6306):257-260. doi:10.1038/349257a0

75. Davis RL, Weintraub H, Lassar AB. Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell*. 1987;51(6):987-1000. doi:10.1016/0092-8674(87)90585-X

76. Cahan P, Li H, Morris SA, da Rocha EL, Daley GQ, Collins JJ. CellNet: Network Biology Applied to Stem Cell Engineering. *Cell*. 2014;158(4):903-915. doi:10.1016/j.cell.2014.07.020

77. Morris SA, Cahan P, Li H, et al. Dissecting engineered cell types and enhancing cell fate conversion via Cellnet. *Cell*. Published online 2014. doi:10.1016/j.cell.2014.07.021

78. Sekiya S, Suzuki A. Direct conversion of mouse fibroblasts to hepatocyte-like cells by defined factors. *Nature*. 2011;475(7356):390-393. doi:10.1038/nature10263

79. Biddy BA, Kong W, Kamimoto K, et al. Single-cell mapping of lineage and identity in direct reprogramming. *Nature*. 2018;564(7735):219-224. doi:10.1038/s41586-018-0744-4

80. Miura S, Suzuki A. Generation of Mouse and Human Organoid-Forming Intestinal Progenitor Cells by Direct Lineage Reprogramming. *Cell Stem Cell*. 2017;21(4):456-471.e5. doi:10.1016/j.stem.2017.08.020

81. Mazzoni EO, Mahony S, Closser M, et al. Synergistic binding of transcription factors to cell-specific enhancers programs motor neuron identity. *Nat Neurosci*. 2013;16(9):10.1038/nn.3467. doi:10.1038/nn.3467

82. Son EY, Ichida JK, Wainger BJ, et al. Conversion of mouse and human fibroblasts into functional spinal motor neurons. *Cell Stem Cell*. 2011;9(3):205-218. doi:10.1016/j.stem.2011.07.014

83. Heinrich C, Bergami M, Gascón S, et al. Sox2-Mediated Conversion of NG2 Glia into Induced Neurons in the Injured Adult Cerebral Cortex. *Stem Cell Rep*. 2014;3(6):1000-1014. doi:10.1016/j.stemcr.2014.10.007

84. Mertens J, Marchetto MC, Bardy C, Gage FH. Evaluating cell reprogramming, differentiation and conversion technologies in neuroscience. *Nat Rev Neurosci*. 2016;17(7):424-437. doi:10.1038/nrn.2016.46

85. Lim SM, Choi WJ, Oh K-W, et al. Directly converted patient-specific induced neurons mirror the neuropathology of FUS with disrupted nuclear localization in amyotrophic lateral sclerosis. *Mol Neurodegener*. 2016;11:8. doi:10.1186/s13024-016-0075-6

86. Santambrogio P, Dusi S, Guaraldo M, et al. Mitochondrial iron and energetic dysfunction distinguish fibroblasts and induced neurons from pantothenate kinase-associated neurodegeneration patients. *Neurobiol Dis*. 2015;81:144-153. doi:10.1016/j.nbd.2015.02.030

87. Drouin-Ouellet J, Pircs K, Barker RA, Jakobsson J, Parmar M. Direct Neuronal Reprogramming for Disease Modeling Studies Using Patient-Derived Neurons: What Have We Learned? *Front Neurosci*. 2017;0. doi:10.3389/fnins.2017.00530

88. Femino AM, Fay FS, Fogarty K, Singer RH. Visualization of Single RNA Transcripts in Situ. 1998;280:7.

89. Eberwine J, Yeh H, Miyashiro K, et al. Analysis of gene expression in single live neurons. *Proc Natl Acad Sci*. 1992;89(7):3010-3014. doi:10.1073/pnas.89.7.3010

90. Kamme F, Salunga R, Yu J, et al. Single-Cell Microarray Analysis in Hippocampus CA1: Demonstration and Validation of Cellular Heterogeneity. *J Neurosci*. 2003;23(9):3607-3615. doi:10.1523/JNEUROSCI.23-09-03607.2003

91.   Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods*. 2008;5(7):621-628. doi:10.1038/nmeth.1226

92.   Picelli S. Single-cell RNA-sequencing: The future of genome biology is now. *RNA Biol*. 2017;14(5):637-650. doi:10.1080/15476286.2016.1201618

93.   Tang F, Barbacioru C, Wang Y, et al. mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods*. 2009;6(5):377-382. doi:10.1038/nmeth.1315

94.   Olsen TK, Baryawno N. Introduction to Single-Cell RNA Sequencing. *Curr Protoc Mol Biol*. 2018;122(1):e57. doi:10.1002/cpmb.57

95.   Grün D, Lyubimova A, Kester L, et al. Single-cell messenger RNA sequencing reveals rare intestinal cell types. *Nature*. 2015;525(7568):251-255. doi:10.1038/nature14966

96.   Hwang B, Lee JH, Bang D. Single-cell RNA sequencing technologies and bioinformatics pipelines. *Exp Mol Med*. 2018;50(8):1-14. doi:10.1038/s12276-018-0071-8

97.   Kolodziejczyk AA, Kim JK, Svensson V, Marioni JC, Teichmann SA. The Technology and Biology of Single-Cell RNA Sequencing. *Mol Cell*. 2015;58(4):610-620. doi:10.1016/j.molcel.2015.04.005

98.   Chromium Single Cell 3' Reagent Kits User Guide (v3.1 Chemistry) -User Guide -Library Prep -Single Cell Gene Expression -Official 10x Genomics Support. Accessed August 5, 2021. https://support.10xgenomics.com/single-cell-gene-expression/library-prep/doc/user-guide-chromium-single-cell-3-reagent-kits-user-guide-v31-chemistry

99.   Zheng GXY, Terry JM, Belgrader P, et al. Massively parallel digital transcriptional profiling of single cells. *Nat Commun*. 2017;8:14049. doi:10.1038/ncomms14049

100.   Kong W, Fu YC, Morris SA. Capybara: A computational tool to measure cell identity and fate transitions. *bioRxiv*. Published online March 26, 2020:2020.02.17.947390. doi:10.1101/2020.02.17.947390

101.   Kamimoto K, Hoffmann CM, Morris SA. CellOracle: Dissecting cell identity via network inference and in silico gene perturbation. *bioRxiv*. Published online February 17, 2020:2020.02.17.947416. doi:10.1101/2020.02.17.947416

102.   Satija R, Farrell JA, Gennert D, Schier AF, Regev A. Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol*. 2015;33(5):495-502. doi:10.1038/nbt.3192

103.   Becht E, McInnes L, Healy J, et al. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotechnol*. 2019;37(1):38-44. doi:10.1038/nbt.4314

104.   Treutlein B, Brownfield DG, Wu AR, et al. Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. *Nature*. 2014;509(7500):371-375. doi:10.1038/nature13173

105. Scialdone A, Tanaka Y, Jawaid W, et al. Resolving early mesoderm diversification through single-cell expression profiling. *Nature*. 2016;535(7611):289-293. doi:10.1038/nature18633

106. Haber AL, Biton M, Rogel N, et al. A single-cell survey of the small intestinal epithelium. *Nature*. 2017;551(7680):333-339. doi:10.1038/nature24489

107. Efremova M, Vento-Tormo M, Teichmann SA, Vento-Tormo R. CellPhoneDB: inferring cell–cell communication from combined expression of multi-subunit ligand–receptor complexes. *Nat Protoc*. 2020;15(4):1484-1506. doi:10.1038/s41596-020-0292-x

108. Squires RH, Duggan C, Teitelbaum DH, et al. Natural history of pediatric intestinal failure: initial report from the Pediatric Intestinal Failure Consortium. *J Pediatr*. 2012;161(4):723-728.e2. doi:10.1016/j.jpeds.2012.03.062

109. Kubal CA, Mangus RS, Tector AJ. Intestine and multivisceral transplantation: current status and future directions. *Curr Gastroenterol Rep*. 2015;17(1):427. doi:10.1007/s11894-014-0427-8

110. Merritt RJ, Cohran V, Raphael BP, et al. Intestinal Rehabilitation Programs in the Management of Pediatric Intestinal Failure and Short Bowel Syndrome. *J Pediatr Gastroenterol Nutr*. 2017;65(5):588–596. doi:10.1097/MPG.0000000000001722

111. Diaz-Miron J, Sun R, Choi P, et al. The effect of impaired angiogenesis on intestinal function following massive small bowel resection. *J Pediatr Surg*. 2015;50(6):948-953. doi:10.1016/j.jpedsurg.2015.03.014

112. 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. doi:10.1038/nbt.4096

113. Treutlein B, Lee QY, Camp JG, et al. Dissecting direct reprogramming from fibroblast to neuron using single-cell RNA-seq. *Nature*. 2016;534(7607):391-395. doi:10.1038/nature18323

114. Haber AL, Biton M, Rogel N, et al. A single-cell survey of the small intestinal epithelium. *Nature*. 2017;551(7680):333-339. doi:10.1038/nature24489

115. Bjerknes M, Cheng H. The stem-cell zone of the small intestinal epithelium. IV. Effects of resecting 30% of the small intestine. *Am J Anat*. 1981;160(1):93-103. doi:10.1002/aja.1001600108

116. Helmrath MA, Fong JJ, Dekaney CM, Henning SJ. Rapid expansion of intestinal secretory lineages following a massive small bowel resection in mice. *Am J Physiol-Gastrointest Liver Physiol*. 2007;292(1):G215-G222. doi:10.1152/ajpgi.00188.2006

117. Jarboe MD, Juno RJ, Stehr W, et al. Epidermal growth factor receptor signaling regulates goblet cell production after small bowel resection. *J Pediatr Surg*. 2005;40(1):92-97. doi:10.1016/j.jpedsurg.2004.09.009

118. Sukhotnik I, Coran AG, Pollak Y, Kuhnreich E, Berkowitz D, Saxena AK. Activated Notch signaling cascade is correlated with stem cell differentiation toward absorptive progenitors after massive small bowel resection in a rat. *Am J Physiol Gastrointest Liver Physiol*. 2017;313(3):G247-G255. doi:10.1152/ajpgi.00139.2017

119. Washizawa N, Gu LH, Gu L, Openo KP, Jones DP, Ziegler TR. Comparative Effects of Glucagon-Like Peptide-2 (GLP-2), Growth Hormone (GH), and Keratinocyte Growth Factor (KGF) on Markers of Gut Adaptation After Massive Small Bowel Resection in Rats. *J Parenter Enter Nutr*. 2004;28(6):399-409. doi:10.1177/0148607104028006399

120. Helmrath MA, Erwin CR, Shin CE, Warner BW. Enterocyte apoptosis is increased following small bowel resection. *J Gastrointest Surg*. 1998;2(1):44-49. doi:10.1016/S1091-255X(98)80102-9

121. Tappenden KA. Intestinal adaptation following resection. *JPEN J Parenter Enteral Nutr*. 2014;38(1 Suppl):23S-31S. doi:10.1177/0148607114525210

122. Rubin DC, Swietlicki EA, Wang JL, Dodson BD, Levin MS. Enterocytic gene expression in intestinal adaptation: evidence for a specific cellular response. *Am J Physiol*. 1996;270(1 Pt 1):G143-152. doi:10.1152/ajpgi.1996.270.1.G143

123. Aronson BE, Stapleton KA, Krasinski SD. Role of GATA factors in development, differentiation, and homeostasis of the small intestinal epithelium. *Am J Physiol - Gastrointest Liver Physiol*. 2014;306(6):G474-G490. doi:10.1152/ajpgi.00119.2013

124. Hebiguchi T, Mezaki Y, Morii M, et al. Massive bowel resection upregulates the intestinal mRNA expression levels of cellular retinol-binding protein II and apolipoprotein A-IV and alters the intestinal vitamin A status in rats. *Int J Mol Med*. 2015;35(3):724-730. doi:10.3892/ijmm.2015.2066

125. Wang JL, Swartz-Basile DA, Rubin DC, Levin MS. Retinoic acid stimulates early cellular proliferation in the adapting remnant rat small intestine after partial resection. *J Nutr*. 1997;127(7):1297-1303. doi:10.1093/jn/127.7.1297

126. Khan HA, Margulies CE. The Role of Mammalian Creb3-Like Transcription Factors in Response to Nutrients. *Front Genet*. 2019;0. doi:10.3389/fgene.2019.00591

127. Nakagawa Y, Satoh A, Tezuka H, et al. CREB3L3 controls fatty acid oxidation and ketogenesis in synergy with PPARα. *Sci Rep*. 2016;6:39182. doi:10.1038/srep39182

128. Nakagawa Y, Shimano H. CREBH Regulates Systemic Glucose and Lipid Metabolism. *Int J Mol Sci*. 2018;19(5):1396. doi:10.3390/ijms19051396

129. Flandez M, Guilmeau S, Blache P, Augenlicht LH. KLF4 REGULATION IN INTESTINAL EPITHELIAL CELL MATURATION. *Exp Cell Res*. 2008;314(20):3712-3723. doi:10.1016/j.yexcr.2008.10.004

130. Chen J, Bardes EE, Aronow BJ, Jegga AG. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res*. 2009;37(Web Server issue):W305-W311. doi:10.1093/nar/gkp427

131. Swartz-Basile DA, Wang L, Tang Y, Pitt HA, Rubin DC, Levin MS. Vitamin A deficiency inhibits intestinal adaptation by modulating apoptosis, proliferation, and enterocyte migration. *Am J Physiol-Gastrointest Liver Physiol*. 2003;285(2):G424-G432. doi:10.1152/ajpgi.00524.2002

132. Swartz-Basile DA, Rubin DC, Levin MS. Vitamin A status modulates intestinal adaptation after partial small bowel resection. *JPEN J Parenter Enteral Nutr*. 2000;24(2):81-88. doi:10.1177/014860710002400281

133. Goncalves A, Roi S, Nowicki M, et al. Fat-soluble vitamin intestinal absorption: Absorption sites in the intestine and interactions for absorption. *Food Chem*. 2015;172:155-160. doi:10.1016/j.foodchem.2014.09.021

134. Villablanca EJ, Wang S, De Calisto J, et al. MyD88 and Retinoic Acid Signaling Pathways Interact to Modulate Gastrointestinal Activities of Dendritic Cells. *Gastroenterology*. 2011;141(1):176-185. doi:10.1053/j.gastro.2011.04.010

135. Blomhoff R, Green MH, Berg T, Norum KR. Transport and storage of vitamin A. *Science*. 1990;250(4979):399-404. doi:10.1126/science.2218545

136. Vilhais-Neto GC, Pourquié O. Retinoic acid. *Curr Biol CB*. 2008;18(5):R191-192. doi:10.1016/j.cub.2007.12.042

137. Grant CE, Bailey TL, Noble WS. FIMO: scanning for occurrences of a given motif. *Bioinformatics*. 2011;27(7):1017-1018. doi:10.1093/bioinformatics/btr064

138. Zhang L, E X, Luker KE, et al. Analysis of human cellular retinol-binding protein II promoter during enterocyte differentiation. *Am J Physiol Gastrointest Liver Physiol*. 2002;282(6):G1079-1087. doi:10.1152/ajpgi.00041.2001

139. Poirier H, Braissant O, Niot I, Wahli W, Besnard P. 9-cis-retinoic acid enhances fatty acid-induced expression of the liver fatty acid-binding protein gene. *FEBS Lett*. 1997;412(3):480-484. doi:10.1016/s0014-5793(97)00830-2

140. Zhong J-C, Huang D-Y, Yang Y-M, et al. Upregulation of angiotensin-converting enzyme 2 by all-trans retinoic acid in spontaneously hypertensive rats. *Hypertens Dallas Tex 1979*. 2004;44(6):907-912. doi:10.1161/01.HYP.0000146400.57221.74

141. Balmer JE, Blomhoff R. Gene expression regulation by retinoic acid. *J Lipid Res*. 2002;43(11):1773-1808. doi:10.1194/jlr.R100015-JLR200

142. Perumal J, Sriram S, Lim HQ, Olivo M, Sugii S. Retinoic acid is abundantly detected in different depots of adipose tissue by SERS. *Adipocyte*. 2016;5(4):378-383. doi:10.1080/21623945.2016.1245817

143. Luo M, Yeruva S, Liu Y, et al. IL-1β-Induced Downregulation of the Multifunctional PDZ Adaptor PDZK1 Is Attenuated by ERK Inhibition, RXRα, or PPARα Stimulation in Enterocytes. *Front Physiol*. 2017;8:61. doi:10.3389/fphys.2017.00061

144. Okano J, Lichti U, Mamiya S, et al. Increased retinoic acid levels through ablation of Cyp26b1 determine the processes of embryonic skin barrier formation and peridermal development. *J Cell Sci*. 2012;125(7):1827-1836. doi:10.1242/jcs.101550

145. Beckman M, Iverfeldt K. Increased gene expression of β-amyloid precursor protein and its homologues APLP1 and APLP2 in human neuroblastoma cells in response to retinoic acid. *Neurosci Lett*. 1997;221(2):73-76. doi:10.1016/S0304-3940(96)13292-4

146. Dechow C, Morath C, Peters J, et al. Effects of all-trans retinoic acid on renin-angiotensin system in rats with experimental nephritis. *Am J Physiol-Ren Physiol*. 2001;281(5):F909-F919. doi:10.1152/ajprenal.2001.281.5.F909

147. Shi J, Zheng B, Chen S, Ma G, Wen J. Retinoic acid receptor α mediates all-trans-retinoic acid-induced Klf4 gene expression by regulating Klf4 promoter activity in vascular smooth muscle cells. *J Biol Chem*. 2012;287(14):10799-10811. doi:10.1074/jbc.M111.321836

148. Wong YF, Wilson PD, Unwin RJ, et al. Retinoic acid receptor-dependent, cell-autonomous, endogenous retinoic acid signaling and its target genes in mouse collecting duct cells. *PloS One*. 2012;7(9):e45725. doi:10.1371/journal.pone.0045725

149. Jenab S, Inturrisi CE. Retinoic acid regulation of mu opioid receptor and c-fos mRNAs and AP-1 DNA binding in SH-SY5Y neuroblastoma cells. *Mol Brain Res*. 2002;99(1):34-39. doi:10.1016/S0169-328X(01)00343-6

150. Doyle TJ, Oudes AJ, Kim KH. Temporal Profiling of Rat Transcriptomes in Retinol-Replenished Vitamin A-Deficient Testis. *Syst Biol Reprod Med*. 2009;55(4):145-163. doi:10.3109/19396360902896844

151. Minkina A, Lindeman RE, Gearhart MD, et al. Retinoic acid signaling is dispensable for somatic development and function in the mammalian ovary. *Dev Biol*. 2017;424(2):208-220. doi:10.1016/j.ydbio.2017.02.015

152. Li C, McFadden SA, Morgan I, et al. All-trans retinoic acid regulates the expression of the extracellular matrix protein fibulin-1 in the guinea pig sclera and human scleral fibroblasts. *Mol Vis*. 2010;16:689-697.

153. Su D, Gudas LJ. Gene expression profiling elucidates a specific role for RARgamma in the retinoic acid-induced differentiation of F9 teratocarcinoma stem cells. *Biochem Pharmacol*. 2008;75(5):1129-1160. doi:10.1016/j.bcp.2007.11.006

154. Liu S-M, Chen W, Wang J. Distinguishing between cancer cell differentiation and resistance induced by all-trans retinoic acid using transcriptional profiles and functional pathway analysis. *Sci Rep*. 2014;4(1):5577. doi:10.1038/srep05577

155. Ma X-Y, Wang J-H, Wang J-L, Ma CX, Wang X-C, Liu F-S. Malat1 as an evolutionarily conserved lncRNA, plays a positive role in regulating proliferation and maintaining undifferentiated status of early-stage hematopoietic cells. *BMC Genomics*. 2015;16(1):676. doi:10.1186/s12864-015-1881-x

156. Umenishi F, Schrier RW. Induction of human aquaporin-1 gene by retinoic acid in human erythroleukemia HEL cells. *Biochem Biophys Res Commun*. 2002;293(3):913-917. doi:10.1016/S0006-291X(02)00316-9

157. Asson-Batres MA, Ryzhov S, Tikhomirov O, et al. Effects of vitamin A deficiency in the postnatal mouse heart: role of hepatic retinoid stores. *Am J Physiol Heart Circ Physiol*. 2016;310(11):H1773-1789. doi:10.1152/ajpheart.00887.2015

158. Napoli JL. Functions of Intracellular Retinoid Binding-Proteins. *Subcell Biochem*. 2016;81:21-76. doi:10.1007/978-94-024-0945-1_2

159. Zeng C, Xu Y, Xu L, et al. Inhibition of long non-coding RNA NEAT1 impairs myeloid differentiation in acute promyelocytic leukemia cells. *BMC Cancer*. 2014;14(1):693. doi:10.1186/1471-2407-14-693

160. Ishijima N, Kanki K, Shimizu H, Shiota G. Activation of AMP-activated protein kinase by retinoic acid sensitizes hepatocellular carcinoma cells to apoptosis induced by sorafenib. *Cancer Sci*. 2015;106(5):567-575. doi:10.1111/cas.12633

161. Larsen FG, Voorhees JJ, Åström A. Retinoic Acid Induces Expression of Early Growth Response Gene-1 (Egr-1) in Human Skin In Vivo and in Cultured Skin Fibroblasts. *J Invest Dermatol*. 1994;102(5):730-733. doi:10.1111/1523-1747.ep12375840

162. Summers JA, Harper AR, Feasley CL, et al. Identification of Apolipoprotein A-I as a Retinoic Acid-binding Protein in the Eye. *J Biol Chem*. 2016;291(36):18991-19005. doi:10.1074/jbc.M116.725523

163. Vu-Dac N, Gervois P, Torra IP, et al. Retinoids increase human apo C-III expression at the transcriptional level via the retinoid X receptor. Contribution to the hypertriglyceridemic action of retinoids. *J Clin Invest*. 1998;102(3):625-632.

164. Iqbal C, Qandeel H, Zheng Y, Duenes J, Sarr M. Mechanisms of Ileal Adaptation for Glucose Absorption after Proximal-Based Small Bowel Resection. *J Gastrointest Surg Off J Soc Surg Aliment Tract*. 2008;12(11):1854-1865. doi:10.1007/s11605-008-0666-9

165. Reboul E. Absorption of vitamin A and carotenoids by the enterocyte: focus on transport proteins. *Nutrients*. 2013;5(9):3563-3581. doi:10.3390/nu5093563

166. Schall KA, Holoyda KA, Isani M, et al. Intestinal adaptation in proximal and distal segments: Two epithelial responses diverge after intestinal separation. *Surgery*. 2017;161(4):1016-1027. doi:10.1016/j.surg.2016.10.033

167. Andorsky DJ, Lund DP, Lillehei CW, et al. Nutritional and other postoperative management of neonates with short bowel syndrome correlates with clinical outcomes. *J Pediatr*. 2001;139(1):27-33. doi:10.1067/mpd.2001.114481

168. Feldman EJ, Dowling RH, McNaughton J, Peters TJ. Effects of Oral Versus Intravenous Nutrition on Intestinal Adaptation After Small Bowel Resection in the Dog. *Gastroenterology*. 1976;70(5):712-719. doi:10.1016/S0016-5085(76)80261-2

169. Dawson MI, Xia Z. The retinoid X receptors and their ligands. *Biochim Biophys Acta BBA - Mol Cell Biol Lipids*. 2012;1821(1):21-56. doi:10.1016/j.bbalip.2011.09.014

170. Danno H, Ishii K, Nakagawa Y, et al. The liver-enriched transcription factor CREBH is nutritionally regulated and activated by fatty acids and PPARα. *Biochem Biophys Res Commun*. 2010;391(2):1222-1227. doi:10.1016/j.bbrc.2009.12.046

171. Chandra V, Huang P, Hamuro Y, et al. Structure of the intact PPAR-γ–RXR-α nuclear receptor complex on DNA. *Nature*. 2008;456(7220):350-356. doi:10.1038/nature07413

172. Mora JR, Iwata M, von Andrian UH. Vitamin effects on the immune system: vitamins A and D take centre stage. *Nat Rev Immunol*. 2008;8(9):685-698. doi:10.1038/nri2378

173. Bünger M, van den Bosch HM, van der Meijde J, Kersten S, Hooiveld GJEJ, Müller M. Genome-wide analysis of PPARα activation in murine small intestine. *Physiol Genomics*. 2007;30(2):192-204. doi:10.1152/physiolgenomics.00198.2006

174. Williamson RC, Buchholtz TW, Malt RA. Humoral stimulation of cell proliferation in small bowel after transection and resection in rats. *Gastroenterology*. 1978;75(2):249-254.

175. Formeister EJ, Sionas AL, Lorance DK, Barkley CL, Lee GH, Magness ST. Distinct SOX9 levels differentially mark stem/progenitor populations and enteroendocrine cells of the small intestine epithelium. *Am J Physiol - Gastrointest Liver Physiol*. 2009;296(5):G1108-G1118. doi:10.1152/ajpgi.00004.2009

176. Dekaney CM, Rodriguez JM, Graul MC, Henning SJ. Isolation and Characterization of a Putative Intestinal Stem Cell Fraction From Mouse Jejunum. *Gastroenterology*. 2005;129(5):1567-1580. doi:10.1053/j.gastro.2005.08.011

177. Alles J, Karaiskos N, Praktiknjo SD, et al. Cell fixation and preservation for droplet-based single-cell transcriptomics. *BMC Biol*. 2017;15(1):44. doi:10.1186/s12915-017-0383-5

178. Barron L, Sun RC, Aladegbami B, Erwin CR, Warner BW, Guo J. Intestinal Epithelial-Specific mTORC1 Activation Enhances Intestinal Adaptation After Small Bowel Resection. *Cell Mol Gastroenterol Hepatol*. 2016;3(2):231-244. doi:10.1016/j.jcmgh.2016.10.006

179. Stappenbeck TS, Virgin HW. Accounting for reciprocal host-microbiome interactions in experimental science. *Nature*. 2016;534(7606):191-199. doi:10.1038/nature18285

180. Schindelin J, Arganda-Carreras I, Frise E, et al. Fiji: an open-source platform for biological-image analysis. *Nat Methods*. 2012;9(7):676-682. doi:10.1038/nmeth.2019

181. Otsu N. A Threshold Selection Method from Gray-Level Histograms. *IEEE Trans Syst Man Cybern*. 1979;9(1):62-66. doi:10.1109/TSMC.1979.4310076

182. Angus HCK, Butt AG, Schultz M, Kemp RA. Intestinal Organoids as a Tool for Inflammatory Bowel Disease Research. *Front Med*. 2020;0. doi:10.3389/fmed.2019.00334

183. Seiler KM, Waye SE, Kong W, et al. Single-Cell Analysis Reveals Regional Reprogramming during Adaptation to Massive Small Bowel Resection in Mice. *Cell Mol Gastroenterol Hepatol*. 2019;0(0). doi:10.1016/j.jcmgh.2019.06.001

184. Guo C, Kong W, Kamimoto K, et al. CellTag Indexing: genetic barcode-based sample multiplexing for single-cell genomics. *Genome Biol*. 2019;20(1):90. doi:10.1186/s13059-019-1699-y

185. Macosko EZ, Basu A, Satija R, et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell*. 2015;161(5):1202-1214. doi:10.1016/j.cell.2015.05.002

186. Muzumdar MD, Tasic B, Miyamichi K, Li L, Luo L. A global double-fluorescent Cre reporter mouse. *genesis*. 2007;45(9):593-605. doi:10.1002/dvg.20335

187. Mouse Intestinal Organoid Culture Protocols. Accessed August 5, 2021. https://www.stemcell.com/intestinal-epithelial-organoid-culture-with-intesticult-organoid-growth-medium-mouse-lp.html

188. Brenna Ø, Bruland T, Furnes MW, et al. The guanylate cyclase-C signaling pathway is down-regulated in inflammatory bowel disease. *Scand J Gastroenterol*. 2015;50(10):1241-1252. doi:10.3109/00365521.2015.1038849

189. Rappaport JA, Waldman SA. The Guanylate Cyclase C-cGMP Signaling Axis Opposes Intestinal Epithelial Injury and Neoplasia. *Front Oncol*. 2018;8:299. doi:10.3389/fonc.2018.00299

190. Grainger S, Savory JGA, Lohnes D. Cdx2 regulates patterning of the intestinal epithelium. *Dev Biol*. 2010;339(1):155-165. doi:10.1016/j.ydbio.2009.12.025

191. Waldman SA, Camilleri M. Guanylate cyclase-C as a therapeutic target in gastrointestinal disorders. *Gut*. 2018;67(8):1543-1552. doi:10.1136/gutjnl-2018-316029

192. Ikpa PT, Sleddens HFBM, Steinbrecher KA, et al. Guanylin and uroguanylin are produced by mouse intestinal epithelial cells of columnar and secretory lineage. *Histochem Cell Biol*. 2016;146(4):445-455. doi:10.1007/s00418-016-1453-4

193.   Pattison AM, Blomain ES, Merlino DJ, et al. Intestinal Enteroids Model Guanylate Cyclase C-Dependent Secretion Induced by Heat-Stable Enterotoxins. *Infect Immun*. 2016;84(10):3083-3091. doi:10.1128/IAI.00639-16

194.   Yin X, Farin HF, van Es JH, Clevers H, Langer R, Karp JM. Niche-independent high-purity cultures of Lgr5+ intestinal stem cells and their progeny. *Nat Methods*. 2014;11(1):106-112. doi:10.1038/nmeth.2737

195.   Grenier E, Maupas FS, Beaulieu J-F, et al. Effect of retinoic acid on cell proliferation and differentiation as well as on lipid synthesis, lipoprotein secretion, and apolipoprotein biogenesis. *Am J Physiol-Gastrointest Liver Physiol*. 2007;293(6):G1178-G1189. doi:10.1152/ajpgi.00295.2007

196.   Sato T, Stange DE, Ferrante M, et al. Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium. *Gastroenterology*. 2011;141(5):1762-1772. doi:10.1053/j.gastro.2011.07.050

197.   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. doi:10.1016/j.cell.2006.07.024

198.   Vierbuchen T, Ostermeier A, Pang ZP, Kokubu Y, Südhof TC, Wernig M. Direct conversion of fibroblasts to functional neurons by defined factors. *Nature*. 2010;463(7284):1035-1041. doi:10.1038/nature08797

199.   Nandan MO, Ghaleb AM, Bialkowska AB, Yang VW. Krüppel-like factor 5 is essential for proliferation and survival of mouse intestinal epithelial stem cells. *Stem Cell Res*. 2015;14(1):10-19. doi:10.1016/j.scr.2014.10.008

200.   Barker N, van Es JH, Kuipers J, et al. Identification of stem cells in small intestine and colon by marker gene Lgr5. *Nature*. 2007;449(7165):1003-1007. doi:10.1038/nature06196

201.   Bastide P, Darido C, Pannequin J, et al. Sox9 regulates cell proliferation and is required for Paneth cell differentiation in the intestinal epithelium. *J Cell Biol*. 2007;178(4):635-648. doi:10.1083/jcb.200704152

202.   Bejjani F, Evanno E, Zibara K, Piechaczyk M, Jariel-Encontre I. The AP-1 transcriptional complex: Local switch or remote command? *Biochim Biophys Acta BBA - Rev Cancer*. 2019;1872(1):11-23. doi:10.1016/j.bbcan.2019.04.003

203.   Donaghey J, Thakurela S, Charlton J, et al. Genetic determinants and epigenetic effects of pioneer factor occupancy. *Nat Genet*. 2018;50(2):250-258. doi:10.1038/s41588-017-0034-3

204.   Thompson CA, Wojta K, Pulakanti K, Rao S, Dawson P, Battle MA. GATA4 Is Sufficient to Establish Jejunal Versus Ileal Identity in the Small Intestine. *Cell Mol Gastroenterol Hepatol*. 2017;3(3):422-446. doi:10.1016/j.jcmgh.2016.12.009

205. Beuling E, Baffour-Awuah NYA, Stapleton KA, et al. GATA factors regulate proliferation, differentiation, and gene expression in small intestine of mature mice. *Gastroenterology*. 2011;140(4):1219-1229.e1-2. doi:10.1053/j.gastro.2011.01.033

206. Battle MA, Bondow BJ, Iverson MA, et al. GATA4 is essential for jejunal function in mice. *Gastroenterology*. 2008;135(5):1676-1686.e1. doi:10.1053/j.gastro.2008.07.074

207. Berthois Y, Katzenellenbogen JA, Katzenellenbogen BS. Phenol red in tissue culture media is a weak estrogen: implications concerning the study of estrogen-responsive cells in culture. *Proc Natl Acad Sci U S A*. 1986;83(8):2496-2500. doi:10.1073/pnas.83.8.2496

208. Hove HV, Antunes ARP, Vlaminck KD, Scheyltjens I, Ginderachter JAV, Movahedi K. Identifying the variables that drive tamoxifen-independent CreERT2 recombination: Implications for microglial fate mapping and gene deletions. *Eur J Immunol*. 2020;50(3):459-463. doi:10.1002/eji.201948162

209. Bues J, Biočanin M, Pezoldt J, et al. Deterministic scRNA-seq of individual intestinal organoids reveals new subtypes and coexisting distinct stem cell pools. *bioRxiv*. Published online May 21, 2020:2020.05.19.103812. doi:10.1101/2020.05.19.103812

210. Lindholm HT, Parmar N, Drurey C, et al. *Developmental Pathways Regulate Cytokine-Driven Effector and Feedback Responses in the Intestinal Epithelium*. Immunology; 2020. doi:10.1101/2020.06.19.160747

211. Jones JC, Dempsey PJ. Enterocyte progenitors can dedifferentiate to replace lost Lgr5+ intestinal stem cells revealing that many different progenitor populations can regain stemness. *Stem Cell Investig*. 2016;3(10). Accessed August 5, 2021. https://sci.amegroups.com/article/view/12004

212. Itoh H, Yagi M, Fushida S, et al. ACTIVATION OF IMMEDIATE EARLY GENE, c-fos, AND c-jun IN THE RAT SMALL INTESTINE AFTER ISCHEMIA/REPERFUSION1. *Transplantation*. 2000;69(4):598–604.

213. Santos MM, Tannuri ACA, Coelho MCM, et al. Immediate expression of c-fos and c-jun mRNA in a model of intestinal autotransplantation and ischemia-reperfusion in situ. *Clinics*. 2015;70(5):373-379. doi:10.6061/clinics/2015(05)12

214. Liu W, Li H, Hong S-H, Piszczek GP, Chen W, Rodgers GP. Olfactomedin 4 deletion induces colon adenocarcinoma in ApcMin/+ mice. *Oncogene*. 2016;35(40):5237-5247. doi:10.1038/onc.2016.58

215. Grover PK, Hardingham JE, Cummins AG. Stem cell marker olfactomedin 4: critical appraisal of its characteristics and role in tumorigenesis. *Cancer Metastasis Rev*. 2010;29(4):761-775. doi:10.1007/s10555-010-9262-z

216. van der Flier LG, Haegebarth A, Stange DE, van de Wetering M, Clevers H. OLFM4 is a robust marker for stem cells in human intestine and marks a subset of colorectal cancer cells. *Gastroenterology*. 2009;137(1):15-17. doi:10.1053/j.gastro.2009.05.035

217. Pattison AM, Merlino DJ, Blomain ES, Waldman SA. Guanylyl cyclase C signaling axis and colon cancer prevention. *World J Gastroenterol*. 2016;22(36):8070-8077. doi:10.3748/wjg.v22.i36.8070

218. Olsson L, Lindmark G, Hammarström M-L, Hammarström S, Sitohy B. Evaluating macrophage migration inhibitory factor 1 expression as a prognostic biomarker in colon cancer. *Tumor Biol*. 2020;42(6):1010428320924524. doi:10.1177/1010428320924524

219. Jiang Z, Wang X, Li J, Yang H, Lin X. Aldolase A as a prognostic factor and mediator of progression via inducing epithelial–mesenchymal transition in gastric cancer. *J Cell Mol Med*. 2018;22(9):4377-4386. doi:10.1111/jcmm.13732

220. Zhu W, Ma L, Qian J, et al. The molecular mechanism and clinical significance of LDHA in HER2-mediated progression of gastric cancer. *Am J Transl Res*. 2018;10(7):2055-2067.

221. Ahmad SS, Glatzle J, Bajaeifer K, et al. Phosphoglycerate kinase 1 as a promoter of metastasis in colon cancer. *Int J Oncol*. 2013;43(2):586-590. doi:10.3892/ijo.2013.1971

222. Wang W, Liu Z, Chen X, et al. Downregulation of FABP5 Suppresses the Proliferation and Induces the Apoptosis of Gastric Cancer Cells Through the Hippo Signaling Pathway. *DNA Cell Biol*. Published online June 22, 2021. doi:10.1089/dna.2021.0370

223. Zaret KS, Carroll JS. Pioneer transcription factors: establishing competence for gene expression. *Genes Dev*. 2011;25(21):2227-2241. doi:10.1101/gad.176826.111

224. Koo B-K, Sasselli V, Clevers H. Retroviral Gene Expression Control in Primary Organoid Cultures. *Curr Protoc Stem Cell Biol*. 2013;27(1):5A.6.1-5A.6.8. doi:10.1002/9780470151808.sc05a06s27

225. Andersson-Rolf A, Fink J, Mustata RC, Koo B-K. A Video Protocol of Retroviral Infection in Primary Intestinal Organoid Culture. *J Vis Exp JoVE*. 2014;(90):51765. doi:10.3791/51765

226. Stoeckius M, Zheng S, Houck-Loomis B, et al. Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *Genome Biol*. 2018;19(1):224. doi:10.1186/s13059-018-1603-1

227. Arceci RJ, King AA, Simon MC, Orkin SH, Wilson DB. Mouse GATA-4: a retinoic acid-inducible GATA-binding transcription factor expressed in endodermally derived tissues and heart. *Mol Cell Biol*. 1993;13(4):2235-2246.

228. Chang C-L, Cueva-Mèndez GDL, Lao-Sirieix P, Laskey R, Fitzgerald R. Altered retinoic acid signalling and its relationship to Cdx-2 in Barrett's oesophagus and associated carcinogenesis. *Cancer Res*. 2005;65(9 Supplement):922-923.

229. Karsenty J, Helal O, de la Porte PL, et al. I-FABP expression alters the intracellular distribution of the BODIPY C16 fatty acid analog. *Mol Cell Biochem*. 2009;326(1-2):97-104. doi:10.1007/s11010-008-0004-2

<u>Curriculum Vitae</u>

Sarah Elizabeth Waye

Education

Washington University in St. Louis   **St. Louis, MO**
Ph.D. in Developmental, Regenerative & Stem Cell Biology   December 2021
Lab of Samantha Morris, Ph.D.
GPA: 4.0
- Dissertation: "Regional Reprogramming and the Small Intestine: Analysis and Modeling of Adaptive Regeneration of the Epithelium"

Georgetown University   **Washington, DC**
Bachelor of Science in Neurobiology   May 2015
Minor in Theology
GPA: 3.919, Summa Cum Laude, Phi Beta Kappa

Research Experience

Graduate Research Assistant   **June 2016-September 2021**
Washington University in St. Louis School of Medicine   St. Louis, MO
Supervised by Samantha Morris, Ph.D.
- Researching cellular reprogramming methods for the intestine by modeling and investigating adaptation of the bowel in Short Bowel Syndrome using organoids and single cell RNA sequencing

Undergraduate Research Assistant   **May 2012-May 2015**
Georgetown Lombardi Comprehensive Cancer Center   Washington, DC
Supervised by Christopher Albanese, Ph.D.
- Determining the relationship of tumor suppressor p53 and VMY-1-103 in induction of cell death in medulloblastoma cell lines

Teaching Experience

Assistant to the Course Instructor, Nucleic Acids and Protein Biosynthesis
Washington University in St. Louis

August 2017-December 2017

- Held weekly office hours and open lines of communication for students
- Wrote and graded exam questions
- Attended lectures and review sessions to assist in teaching method

Special Review Teacher, Foundations of Biology I
Georgetown University

August 2014-December 2014

- Created lesson plans for review sessions following the lecture material
- Provided twice weekly seminar style reviews for students identified by professor
- Met regularly with professor to talk about student progress

Head Teaching Assistant, Foundations of Biology I
Georgetown University

August 2013-December 2014

- Oversaw the activities of two laboratory sections a week, leading the class when the professor was not present
- Worked in tutoring center
- Graded lab reports, exams, and contributed to the constructive criticism of new teaching assistants

Teaching Assistant, Foundations of Biology I
Georgetown University

August 2012-December 2012

- Mentored a laboratory section of 12 students
- Graded exams and laboratory papers, collaborating with professors and other TAs
- Communicated with students about their progress

Professional Experience

BALSA-Kantar Associate

November 2019-February 2020

Junior/Senior Associate

- Worked directly with local consultants at Kantar STL to gain real-world experience in consulting
- Collaborated in a team of 2 to research assignments given by Kantar consultant and produce high quality deliverables in a timely manner
- Acted as Junior Consultant for 2 months and led the project as Senior for 2 months

BALSA Group

<div align="right">**August 2019-February 2021**</div>

Project Manager
- Led and organized a team of 3 consultants on a project for the BALSA Group
- Planned and created timelines of research for each consultant to achieve project goals in a timely fashion
- Maintained open lines of communication with consultants to provide feedback on deliverables and research
- Assisted consultants on writing and presentation skills to communicate clearly with specific clients

BALSA Group

<div align="right">**May 2018- August 2019**</div>

Consultant
- Researched and provided written and presentation deliverables weekly doing market research for consulting clients (3 separate projects)
- Experience in secondary research, primary research (interviews and surveys), compiling data, and presenting visually appealing summaries to clients

Awards and Honors

- "People's Choice" for Best Oral Presentation at Washington University in St. Louis Developmental Biology Retreat (2020)
- NIH T-32 Training Grant Awardee (2017-2019)
- NSF GRFP Honorable Mention (2016)
- Phi Beta Kappa Inductee (2015)
- Georgetown Neurobiology Award (2015)
- Barry M. Goldwater Scholarship (2014)
- Alpha Sigma Nu, the Jesuit Honor Society Inductee (2014)
- "Best Oral Presentation" at Georgetown Howard Hughes Research Symposium (2013)

Publications

Journal Publications
- Minn, KT., Dietmann, S., **Waye, SE.,** Morris, SA., and Solnica-Krezel, L., Gene expression dynamics underlying cell fate emergence in 2D micropatterned human embryonic stem cell gastruloids. *Stem Cell Reports*. 2021, 16:5, 1210-1227.
- (Co-First) Seiler, KM.[1], **Waye, SE.**[1], Kong, W., Kamimoto, K., Bajinting, A., Goo, WH., Onufer, EJ., Courtney, C., Guo, J., Warner, BJ., and Morris, SA. Single-Cell Analysis Reveals Regional Reprogramming during Adaptation to Massive Small Bowel Resection in Mice. *Cellular and Molecular Gastroenterology and Hepatology*. 2019, 8:3, 407-426.

- Biddy, BA., Kong, W., Kamimoto, K., Guo, C., **Waye, SE.,** Sun, T., and Morris, SA. Single-cell mapping of lineage and identity in direct reprogramming. *Nature.* 2018, 564, 219-224.
- (Co-First) **Waye, S.**[1], Naeem, A.[1], Choudhry, M.U.[1], Parasido, E., Tricoli, L., Yenugonda, V., Michaiel, J.P., Rodriguez, O., Karam, S.D., Rood, B.R., Avantaggiati, M.L., and Albanese, C. The p53 tumor suppressor protein protects against chemotherapeutic stress and apoptosis in human medulloblastoma cells. *Aging,* 2015, 7:10, 845-868.
- Ringer, L.,[1] Sirajuddin, P.,[1] Tricoli, L.,[1] **Waye, S.**, Choudhry, M. U., Parasido, E., Sivakumar, A., Heckler, M., Naeem, A., Abdelgawad, I., Liu, X., Feldman, A., Lee, R. J., Wu, C., Yenugonda, V., Kallakury, B., Dritschilo, A., Lynch, J., Schlegel, R., Rodriguez, O., Pestell, R. G., Avantaggiati, M. L., and Albanese, C. The Induction of the p53 Tumor Suppressor Protein Bridges the Apoptotic and Autophagic Signaling Pathways to Regulate Cell Death in Prostate Cancer Cells. *Oncotarget*, 2014, 5:21, 10678-10691.

Published Abstracts
- Waye, S. The cell damage and death pathways of VMY-1-103. *Mentis Vita*, 2014; 14.

Presentations

- Presented "Experimental and computational modeling of epithelial reprogramming following loss of small bowel" at Washington University Developmental Biology Retreat, St. Louis, MO, October 8, 2020.
- *Invited Speaker* Presented "Gut reactions: scRNA-seq analysis of regional reprogramming after bowel resection" at 10x Genomics User Group Meeting, St. Louis, MO, June 6, 2019.
- Presented "Gut reactions: scRNA-seq analysis of regional reprogramming after bowel resection" at Washington University Developmental Biology Department Work-in-Progress, St. Louis, MO, March 20, 2019.
- Presented "Understanding and Engineering Small Intestinal Regeneration using scRNA-seq" at Washington University DRSCB Program Recruitment, St. Louis, MO, February 1, 2019.
- Presented "Understanding and Engineering Small Intestinal Regeneration" at Washington University in St. Louis DB/C4 Research Forum, St. Louis, MO, January 31, 2018.
- Presented "Silencing of p53 enhances chemotherapeutic outcomes in human medulloblastoma through Fas signaling" at Georgetown Research Intensive Senior Experience 2014, Washington, DC, December 16, 2014.
- Presented "The DNA damage related and translational effects of compound VMY-1-103 on medulloblastoma" at the Georgetown Howard Hughes Research Symposium 2014, Washington, DC, July 22, 2014.

- Presented "The cell damage and death pathways of VMY-1-103" at the Georgetown Howard Hughes Research Symposium 2013, Washington, DC, July 28, 2013.
  - Presented "The autophagic cell death mechanism of VMY-1-103" at the Georgetown Howard Hughes Research Symposium 2012, Washington, DC, July 26, 2012.

Posters

- Experimental and computational modeling of epithelial reprogramming following loss of small bowel at EMBO | EMBL Symposium: Organoids: Modelling Organ Development and Disease in 3D Culture, Virtual, October 22, 2020.
- hiIPs Don't Lie: Improving Identity Adherence in Directly Reprogrammed Intestinal Cells at Student Invited Speaker Poster Session 2019, St. Louis, MO, September 11, 2019
- hiIPs Don't Lie: Improving Identity Adherence in Directly Reprogrammed Intestinal Cells at Washington University in St. Louis DDRCC Grant Review 2019, St. Louis, MO, June 24, 2019
- hiIPs Don't Lie: Improving Identity Adherence in Directly Reprogrammed Intestinal Cells at Developmental, Regenerative & Stem Cell Biology Retreat 2018, New Haven, MO, May 31, 2018.
- Silencing of p53 enhances chemotherapeutic outcomes in human medulloblastoma through Fas signaling at Georgetown Undergraduate Research Conference 2015, Washington, DC, April 16, 2015.

Professional Organizations and Leadership

- **BALSA Group,** project manager (2019-2021)
- **DBBS Student Advisory Committee,** board member (2018-2020)
- **DBBS Peer Mentoring Program,** co-chair (2018-2020)
- **BALSA Group,** consultant (2018-2019)
- **Phi Beta Kappa,** member (2015-Present)
- **Sigma Xi,** Associate member (2014-Present)