Amplification of Minor Pathogenic Stem Cell Variants in Crohn's Disease

by
Rahul Dev Neupane

A dissertation submitted to the Department of Biology and Biochemistry,
College of Natural Sciences and Mathematics
in partial fulfillment of the requirements for the degree of

Degree of Philosophy
in Biology

Chair of Committee: Frank McKeon

Co-Chair of Committee: Wa Xian

Committee Member: Brigitte Dauwalder

Committee Member: Preethi Gunaratne

Committee Member: Jason Hou

Committee Member: Richard Willson

University of Houston
December 2020

ACKNOWLEDGEMENTS

This work was supported by grants from the Cancer Prevention Research Institute of Texas (CPRIT; RR150104 to Wa Xian and RR1550088 to Frank McKeon), the National Institutes of Health (1R01DK115445-01A1 to Wa Xian, 1R01CA241600-01 and U24CA228550 to Frank McKeon), the US Dept. of Defense (W81XWH-17-1-0123 to Wa Xian), and the American Gastroenterology Association Research and Development Pilot Award in Technology (to Wa Xian).

I would like to thank my advisors, Dr. Wa Xian, and Dr. Frank McKeon, for their continued encouragement, support, guidance, and valuable inputs throughout this endeavor. Without their time and effort, the completion of all this work would be impossible. I would like to extend my gratitude towards my past and present committee members, Dr. Brigitte Dauwalder, Dr. Preethi Gunaratne, Dr. Jason Hou, Dr. Richard Willson, and Dr. Steven Bark, for their time, suggestions, and patience. I am immensely thankful to Shan Wang and Jingzhong Xie for all the bioinformatics related analysis throughout my project. I am grateful to Audrey-Ann Liew for all the help and support for wet laboratory-related works. I am appreciative to Wei Rao and Suchan Niroula for the helpful discussions and brainstorming for the project. I am also indebted to undergraduates Valeria Duran, Benjamin Haverty, Jeremy Siegelman, Uzma Maknojia, Roba Abousaway, Thea Pascual, and Ashley Hoffman for their contributions in histology.

ABSTRACT

Crohn's disease (CD) is a progressive inflammatory and fibrotic disorder of the intestinal tract thought to arise from defective interactions of immune cells, intestinal microbes, and intervening mucosal barriers. Here we perform novel clonogenic analyses of stem cells derived from the terminal ileum of pediatric and adult patients with CD. We show that CD stem cell libraries are dominated by two variants that display ingrained pro-inflammatory and pro-fibrotic signaling. Transplantation of these variants to immunodeficient mice triggers key features of CD including leukocyte infiltration and fibrosis. These variants, which exist at low levels in control and fetal terminal ileum, display an absolute commitment to gastric epithelial fates. Together, this work links CD to the amplification of minor stem cell variants whose conventional roles are in the response to ancient human pathogens and suggests mechanistic analogies to chronic obstructive pulmonary disease (COPD) and perhaps other chronic inflammatory conditions. If true, Crohn's disease can be mitigated by therapeutics that selectively target these variant stem cells allowing existing normal stem cells to compensate for the loss of those. We develop a high-throughput chemical screening platform to uncover small molecules that specifically target these variant stem cells. Furthermore, these small molecules, in a unique combination with promoters, protect and promote normal stem cells while eradicating variant stem cells. Lastly, the clonogenic mucosal stem cell variants tied to pathological features of the disease are selectively targeted by candidate small molecules in the xenograft model diminishing those features in parallel and leaving normal stem cells to repopulate and maintain homeostasis.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS ……ii
ABSTRACT ……iii
LIST OF TABLES ……vii
LIST OF FIGURES ……viii
LIST OF ABBREVIATIONS ……ix
LIST OF PUBLICATIONS USED IN THE DISSERTATION ……xi
1 GENERAL INTRODUCTION……1
1.1 Crohn's disease: incidence, symptoms and diagnosis ……1
1.2 Current therapies and medical management ……3
1.3 Causes ……6
1.3.1 Genetics ……6
1.3.2 Environmental factors ……8
1.3.3 Microbiota ……9
1.3.4 Innate and adaptive immunity ……9
1.3.5 Intestinal epithelial barrier ……11
1.4 Animal models ……13
1.4.1 Chemically induced models ……14
1.4.2 Immunological models ……15
1.4.3 Genetic models ……15
1.4.4 Spontaneous models ……15
2 MATERIALS AND METHODS ……17
2.1 Materials ……17
2.1.1 List of reagents and commercial kits ……17
2.1.2 List of antibodies ……17
2.1.3 List of qPCR primers ……18
2.1.4 Experimental models: cell lines ……18
2.1.5 Experimental models: organisms/strains ……18
2.1.6 Biological samples ……18

2.2 Experimental model and subject details ……19
2.2.1 Human subjects ……...19
2.2.2 Primary cell culture ……..…19
2.2.3 Animals ……..…19
2.2.4 Cell lines ……...20
2.3 Method details ……..…20
2.3.1 *In vitro* culture of human terminal ileum epithelial stem cells ……..….20
2.3.2 Stem cell differentiation ……..….21
2.3.3 Histology and immunostaining ……..….21
2.3.4 Xenografts in immunodeficient mice ……...22
2.3.5 Flow cytometry analysis ……..……23
2.3.6 RNA sample preparation ……23
2.3.7 Sequence alignment of single cell RNA sequencing ……..…23
2.3.8 Single cell RNA sequencing ……..24
2.3.9 Expression microarray data analysis ……...25
2.3.10 Expression qPCR array ……..….26
2.3.11 Statistical analysis ……..…26
2.3.12 *In vitro* culture of GFP-tagged cells in multi-well plates……...26
2.3.13 High throughput screening and imaging of cell lines ……..….27
2.3.14 High content image analysis and selection of drugs ……..…28
2.3.15 Drug validation and establishing dosage response curve ……..….29
2.3.16 *In vivo* testing of drugs in xenograft model ……...29
3 CLONING AND CHARACTERIZATION OF PATIENT-DERIVED STEM CELLS…..31
3.1 Significance of the proposed study ……..31
3.2 Research strategy ……..32
3.2.1 Cloning epithelial stem cells from Crohn's patient biopsies ……..……....32
3.2.2 Expanding single cell-derived clones ……..…..34
3.2.3 Functional characterization of intestinal stem cells ……..…...34
3.2.4 Stability of intestinal cells ……...34

3.3 Results ……35
3.3.1 Stem cell heterogeneity in Crohn's ……35
3.3.2 Crohn's variant stem cells committed to upper GI fate ……48
3.3.3 Crohn's stem cell libraries drive neutrophilic inflammation ……56
3.3.4 Variant stem cells drive fibrotic host responses ……66
3.3.5 Crohn's variant stem cells preexist in fetal terminal ileum ……73
3.4 Discussion ……84
4 DRUG DISCOVERY USING PATIENT-DERIVED STEM CELLS ……90
4.1 Significance of the proposed study ……90
4.2 Research strategy ……91
4.2.1 Phenotypic approach to drug discovery ……91
4.2.2 High throughput screening in hit discovery process ……93
4.2.3 Cell-based assay development and high content screening ……94
4.2.4 Primary and stem cell lines for high throughput screening ……95
4.2.5 Animal model for *in vivo* validation of drugs ……97
4.2.5.1 Limitations of current animal models ……97
4.2.5.2 Humanized mouse model for in vivo validation of drugs ……98
4.3 Results ……99
4.3.1 Adapting stem cells for high throughput screening ……99
4.3.2 HTS allows discovery of drugs against variant stem cells ……100
4.3.3 HTS allows discovery of promoters ……109
4.3.4 Drugs recover the phenotype in xenograft model ……110
4.4 Discussion ……114
BIBLIOGRAPHY ……124

LIST OF TABLES

3.1 Crohn's and control subjects ……...36
4.1 Target or pathway enrichment from common hits ……..104
4.2 Compounds selected for validation *in vitro* ……..108
4.3 Promoters selected for validation *in vitro* ……112

LIST OF FIGURES

3.1 Clonal analysis of terminal ileum stem cells reveals heterogenity in Crohn's ……..39
3.2 Stem cell heterogeneity by qPCR and immuofluorescence staining ……43
3.3 Stem cell heterogeneity by FACS and clonogenicity of stem cells ……...45
3.4 CLST2 and CLST3 clones are committed to upper gastrointestinal fates ……..….49
3.5 Clonal differentiation in air liquid interface (ALI) ……..……..52
3.6 Clonal differentiation in xenograft ……..……..54
3.7 CLST2 and CLST3 markers in resected tissues from three Crohn's patient ……..55
3.8 Constitutive expression of inflammatory mediators in CLST2 and CLST3 ……..58
3.9 CLST3 clones drive neutrophilic inflammation in xenografts ……..61
3.10 Pro-inflammatory activities of Crohn's libraries ……64
3.11 Inflammatory signatures of Crohn's variant stem cells overlap with GWAS ……67
3.12 Pro-fibrotic activities of CLST2 and CLST3 clones ……..70
3.13 Pro-fibrotic activities of Crohn's libraries ……...74
3.14 CLST 2 and CLST3 variant stem cells in fetal terminal ileum ……..77
3.15 Variant clones in fetal terminal ileum are similar to that in Crohn's……79
3.16 Variant stem cells in provoke pathology in lower ratios ……82
4.1 Stem cells adapted for high throughput screening ……101
4.2 Small molecules specifically target variant clones ……103
4.3 Validation of selected small molecules from HTS ……...……...106
4.4 Synergistic HTS identifies drugs that promote CLST1 ……...……...111
4.5 Validation of two best candidate drugs in xenograft mouse model ……....115

LIST OF ABBREVIATIONS

IBD	Inflammatory Bowel Disease
CD	Crohn's Disease
UC	Ulcerative Colitis
HTS	High Throughput Screening
HCA	High Content Analysis
CTE	Computed Tomography Enterography
MRE	Magnetic Resonance Enterography
5-ASA	5-Aminosalicyclic Acid
6-MP	6-mercaptopurine
MLCK	Myosin Light Chain Kinase
MMPs	Matrix Metalloproteinases
si-RNA	small-interfering Ribonucleotide Acid
GWAS	Genome-wide Association Studies
rRNA	ribosomal Ribonucleic Acid
TLR	Toll-like receptors
NOD	Nucleotide Binding Domain
NLR	Nucleotide Binding Domain (NOD) Like Receptors
iTreg	inducible regulatory T-cells
ER	Endoplasmic Reticulum
DSS	Dextran Sodium Sulfate
TNBS	Trinitrobenzenesulfonic
SCID	Severe Combined Immunodeficient

qPCR	quantitative Polymerase Chain Reaction
NSG	NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ
ALI	Air-liquid Interface
H&E	Hematoxylin and Eosin
IF	Immunofluorescence
IHC	Immunohistochemistry
FACS	Fluorescent-activated Cell Sorting
RIN	RNA Integrity Number
t-SNE	t-distributed Stochastic Neighbor Embedding
PCA	Principal Component Analysis
GFP	Green Fluorescent Protein
RFP	Red Fluorescent Protein
PBS	Phosphate Buffered Saline
CNV	Copy Number Variation
STR	Short Tandem Repeats
HPA	Human Protein Atlas
IACUC	Institutional Animal Care and Use Committee
BWH	Brigham and Women's Hospital
UCONN	University of Connecticut
BCM	Baylor College of Medicine
UNC	University of North Carolina

LIST OF PUBLICATIONS USED IN THE DISSERTATION

Duleba M, Yamamoto Y, **Neupane R**, Rao W, Xie J, Qi Y, Liew AA, Niroula S, Zhang Y, Mahalingam R, Wang S, Goller K, Ajani JA, Vincent M, Ho KY, Hou JK, Hyams JS, Sylvester FA, Crum CP, McKeon F, Xian W. (2020). Cloning of ground-state intestinal stem cells from endoscopic biopsy samples. Nature Protocol, 15(5):1612-1627.

Rao W, Wang S, Duleba M, Niroula S, Goller K, Xie J, Mahalingam R, **Neupane R**, Liew AA, Vincent M, Okuda K, O'Neal WK, Boucher RC, Dickey BF, Wechsler ME, Ibrahim O, Engelhardt JF, Mertens TCJ, Wang W, Jyothula SSK, Crum CP, Karmouty-Quintana H, Parekh KR, Metersky ML, McKeon FD, Xian W. (2020). Regenerative metaplastic clones in COPD lung drive inflammation and fibrosis. Cell, 181(4):848-864.e18.

Duleba M, Qi Y, Mahalingam R, Flynn K, Rinaldi F, Liew AA, **Neupane R**, Vincent M, Crum CP, Ho KY, Hou JK, Hyams JS, Sylvester FA, McKeon F, and Xian W. (2019). An efficient method for cloning gastrointestinal stem cells from patients via endoscopic biopsies. Gastroenterology, 156 (1):20-23.

CHAPTER 1: GENERAL INTRODUCTION

1.1 Crohn's Disease: incidence, symptoms, and diagnosis

Crohn's and ulcerative colitis are two chronic remitting and relapsing inflammatory disease of the bowel. These two parts of inflammatory bowel disease (IBD) are of unknown cause and share many clinical and therapeutic features. However, important clinical and pathophysiological differences exist between these two forms of IBD (Fiocchi, 1998). Ulcerative colitis starts from the rectum and spreads to a variable length of the large intestine while Crohn's can affect any part of the gastrointestinal tract – most commonly, terminal ileum (Both et al., 1983). The primary effect of ulcerative colitis is limited to the mucosa, but Crohn's disease is transmural in nature affecting the entire gastrointestinal tract. Finally, ulcerative colitis presents a continuous disease without the involvement of abscesses, granulomas, or lesions progressing to fistulas (Floren et al., 1987). In contrast, Crohn's disease is characterized by scarring, abscesses, and lesions that can progress to strictures or fistulas in the gastrointestinal tract requiring hospitalization and surgical intervention (Baumgart and Sandborn, 2012). The presentation of Crohn's disease is mostly regional where affected lesions are commonly separated by uninvolved regions known as 'skip areas' (Geboes, 2001). It is an obstructive condition in both adults and pediatrics, with an incidence of 3 to 20 cases per 100,000 each year in the United States (Molodecky et al., 2012). While the onset of the disease is more common in young adults, approximately a quarter of the affected patients are children (Van Limbergen et al., 2008). Patients with Crohn's are in frequent need of immunosuppression and anti-inflammatory therapy. They also have a higher chance of developing colorectal cancer (Van Limbergen et al., 2008; Sauer and Kiugathasan, 2009; Hyams, 2014; Peloquin et al., 2016) and are susceptible to *Clostridium difficile* infections, with an occasional incidence of death among patients (Saidel et al., 2011).

The disease frequently presents with abdominal pain, diarrhea, fever, anemia, weight loss, and delayed growth (in children under 12 years of age). Especially, in teenagers and young adults, the disease may have extra intestinal manifestations as well. Arthritis, cholangitis, hydronephrosis, urinary tract infections, and skin lesions are some of those (Levine and Burakoff, 2011). The patients often show signs of bowel obstruction as the bowel wall gets thickened with the involvement of the submucosa, the muscularis propria, the subserosa, and mesenteric fat. The mesenteric fat wrapping sometimes covers 50% of the circumference of the intestine (Sheehan et al., 1992). This phenomenon is specific to Crohn's disease and is called 'fat wrapping'. In other patients, inflammation can extend transmurally deep into the wall of the intestine and cause the formation of ulcers and abscesses, which is a collection of pus as a result of infection. Sometimes, fistulas can be formed as a result of this abscesses formation in the ileum connecting either two different parts of the bowel or the bowel to other organs such as the bladder or vagina (Scharl and Rogler, 2014). Stenosis or strictures are other forms of intestinal obstruction observed in Crohn's disease. Fibrosis can cause excess smooth muscle cell accumulation and collagen deposition in the bowel wall causing it to be narrow and blocking the passage of food (Rieder et al., 2013).

The duration of symptoms before being diagnosed as Crohn's is being shortened with time as the clinical symptoms are being correlated with objective findings from imaging techniques. Endoscopic and radiological techniques supported by pathology are the basis of the diagnosis of Crohn's disease (Cheifetz, 2013). One of the classic features of Crohn's – 'skip lesions', where normal-appearing mucosa and lesions with varying degrees of inflammation exist side to side, can be seen by endoscopic imaging. Similarly, endoscopes can also characterize luminal stenosis and fistula formations. Both computed tomography enterography (CTE) and magnetic resonance enterography (MRE) scans allow better visualization of

intestinal obstructions such as fistulas, and abscesses (Dambha et al., 2014; Baumgart and Sandborn, 2012). Non-invasive imaging techniques such as ultrasound have been used for initial diagnosis, disease activity evaluation, detection of stenosis, fistula, and abscesses (Panes et al., 2011). Pathology is another aspect of diagnosis of the disease which allows finding more common features of the disease such as the presence of granulomas, plasma cells, varying degrees of infiltrates from lymphocytes, detection of crypt abscesses, crypt atrophy, distortion of crypt architecture, and crypt branching (Heresbach et al., 2005). Although serology is also used as a support for diagnosis, no markers have been specific enough to establish or rule out the diagnosis of Crohn's. Some of the commonly used biomarkers are C-reactive protein, anti-*Saccharomyces cerevisiae*, antineutrophil cytoplasmic autoantibodies, and the fecal granulocyte proteins - lactoferrin and calprotectin (Peeters et al., 2001).

1.2 Current therapies and medical management

Crohn's is an aggressive disease that leads to intestinal obstruction and many associated extra-intestinal complications. Hence, current therapies are aimed towards the resolution of active inflammation followed by stable clinical and endoscopic remission of the disease. The therapeutic approach is based on the age of diagnosis, the severity of the disease, and the extent of gastrointestinal tract involvement. Initially, steroids and anti-TNFs are used as fast-acting agents for immediate symptom relief and combined with strong immune system suppressants such as methotrexate or thiopurines to achieve long-lasting maintenance (Chande et al., 2016). Corticosteroids such as prednisone have been an integral part of therapeutics for the treatment of moderate to severe Crohn's (Steinhart et al., 2000). Their combination with anti-inflammatory drugs has been the choice for treatment during the first 4-5 years of uncomplicated disease stage. In contrast to systemic steroids, topical steroid drugs such as budesonide can be targeted to deliver in specific regions of the gastrointestinal tract. A lower

dose of topical drugs has been proven to be effective in such cases compared to a higher dose of systemic drugs (Seow et al., 2008). Non-steroidal anti-inflammatory drugs such as 5-aminosalicylic acid (5-ASA) derivatives are also used as topical drugs to treat mild to moderate Crohn's. These drugs are coated with a pH-sensitive acrylic polymer and targeted to release in terminal ileum and colon at a pH of 7.0 (Gisbert et al., 2002). While corticosteroids and anti-inflammatory drugs relieve the symptoms by directly working at the site of inflammation, immunomodulatory drugs regulate or normalize the response from the immune system. Azathioprine, 6-mercaptopurine (6-MP), and cyclosporine are some examples of immunomodulators that inhibit natural killer cell and T-cell activities (Dulai et al., 2014).

Humanized monoclonal antibodies are one of the recently approved therapy for moderate to severe Crohn's disease. Vedolizumab and natalizumab are such antibodies that inhibit selective adhesion molecules. Vedolizumab targets adhesion molecule α4β7- integrin heterodimer resulting in inhibition of leukocyte migration (Sandborn et al., 2013). Similarly, natalizumab blocks α4-integrin molecules. Ustekinumab is an interleukin inhibitor blocking the p-40 subunit of IL-12 and IL-23 molecules (Sandborn et al., 2012). These are major cytokines in the pathogenesis of inflammatory bowel disease. Humanized monoclonal antibodies against matrix metalloproteinases (MMPs) are one of the emerging therapies. In experimental models of inflammations, MMP9 has been suggested as a key mediator for impairing colonic epithelial permeability and intensifying inflammation by activating myosin light chain kinase (MLCK) (Nighot et al., 2015). GS-5745, a potent and selective allosteric MMP9 inhibitor, is currently being tested in clinical trials for IBD. Moreover, small-interfering RNA (siRNA) is also one of the emerging therapies in the management of tissue fibrosis. siRNA-based silencing of CHST15, an enzyme biosynthesizing chondroitin sulfate E that promotes tissue fibrosis, in both acute and chronic dextran sodium sulfate colitis models reduced colitis activity, intestinal accumulation of

macrophages, α-SMA-positive myofibroblasts, and collagen deposition (Suzuki et al., 2017). A phase I study of 18 Crohn's patients published in 2016 showed a reduction of endoscopic inflammation as well as tissue fibrosis after administration of siRNA against CHST15 compared to placebo therapy (Suzuki et al., 2017). Lastly, although microbiome has not been identified as a specific cause in Crohn's disease, antibiotics such as ciprofloxacin and metronidazole are used in treatment. Despite the lack of consistent data in terms of effectiveness for induction and maintenance of remission by antibiotics, they might help patients with bacterial infections such as *Clostridium difficile* (Gomollon et al., 2017).

Despite all the advancements in instruments and treatment strategies, the patients with stable clinical remission are as low as 10%. Therapeutic regimens aim to maintain remission without the need to go for surgical intervention, but surgery might be required once patients develop complications such as stricture and fistula formation. The medical management of Crohn's disease has gone some significant changes over the past decades- early introduction of immunosuppressants in high-risk patients, sequential and incremental treatment strategy based on symptoms (Swaminath and Kombluth, 2007). But these treatments still rely on immunosuppressants coupled with biologics directed at TNF-alpha and other inflammatory cytokines. And while these treatments can slow the progression of the disease, they have not reduced the need for surgical intervention nor have been able to fully establish remission in patients (Baumgart and Sandborn, 2012). Surgery might be needed to remove the entire colon (colectomy), the colon and rectum (proctocolectomy), a portion of the colon (resection), or ileocolonic anastomosis (joining the end of the small intestine and first part of the colon). 50% of the patients diagnosed with Crohn's require surgery within 20 years because of intestinal complications (Cheifetz, 2013). And the remaining 50% require surgery with 10 years of the first diagnosis with recurrence being as high as 55% after 10 years. In some cases, surgery might

be needed earlier if they develop abscesses, stricture, perforations, or fistula sooner (Rieder et al., 2016). Surprisingly, surgery has not been able to induce clinical remission as well. Patients with aggressive transmural conditions tend to relapse sooner than others. Hence, many patients will require multiple surgeries over their lifetime (Frolkis et al., 2013).

1.3 Causes

About 3 to 20 per 100,000 new cases of Crohn's disease are reported each year in the United States (Molodecky et al., 2012). Although the incidence of the disease is rising in Asia and South America, it is found to be more common in the industrialized world, North America, and Western Europe (Ng et al., 2013). The prevalence of the disease is slightly higher in women and is common in Ashkenazi Jews than in any other ethnic group (Loftus, 2004). The exact etiology of Crohn's disease remains obscure, but it is believed to be caused by interactions between the immune system, microbiome, genetics, and the environment affecting the lining of the digestive tract (McGovern et al., 2015). So far, the accepted notion for Crohn's pathogenesis includes dysregulated mucosal immune responses to intestinal microbiomes and in genetically susceptible hosts, defects in innate and adaptive immunity, or defects in the intestinal barrier contributing to inflammation.

1.3.1 Genetics

About 10% to 25% of patients diagnosed with Crohn's disease have a family history or a first-degree relative with the disorder (Ng et al., 2013). Along with this, the large concordance studies performed in twins are indications of the role of genetic predisposition in Crohn's disease. Studies in twins show that 20% to 50% monozygotic pairs are concordant for the disorder compared to 10% in dizygotic pairs which is the same concordance as patients with family history. A nationwide study among twins in Germany showed concordance of 35% in

monozygotic pairs and only 3% in dizygotic pairs for the disease (Spehlmann et al., 2008). However, recent studies indicate an overestimation of concordance in monozygotic twins and the contribution of genetics in the etiology of the disease due to methodological limitations in previous studies (Halfvarson, 2011).

Nevertheless, recent genome-wide association studies (GWAS), *in silico* meta-analyses, and intestinal pathophysiology studies have made significant contributions in identifying IBD-associated gene loci. Large scale GWAS alone has identified over 200 IBD-associated genetic risk loci with stunning overlap between Crohn's, ulcerative colitis, and other immune-mediated diseases (Jostins et al., 2012; de Lange et al., 2017). So far, 71 susceptibility loci for Crohn's disease on 17 chromosomes, most significantly on chromosome 16, have been identified (Franke et al., 2010). NOD2 was one of the first genes discovered on chromosome 16 and identified as a major genetic risk factor for Crohn's disease (Hugot et al., 2001). An individual with homozygotic changes in the NOD2 gene has 20 to 40 times higher chances of developing Crohn's while a heterozygote has only 2 to 4 times higher chance (Ogura et al., 2001; Philpott et al., 2004). GWAS reveals an overlap of risk loci for Crohn's and mycobacterial infections and implicates genes of adaptive and innate immune processes in the containment of gut microbes (Hampe et al. et al., 2007; Rioux et al., 2007; Barrett et al., 2008). Interestingly, GWAS has also identified multiple genes as risk loci that have restricted expression in the intestinal mucosa. Downregulation of those genes is linked with deficiencies in anti-microbial functions of Paneth cells in Crohn's disease patients (Wehkamp et al., 2005), defective autophagy processing of microbial antigens by mucosal epithelial cells (Wehkamp et al., 2005; Hampe et al., 2007; Stappenbeck and McGovern, 2017), and altered responsiveness of mucosal immune cells (Abraham and Cho, 2009).

Despite large contributions of GWAS and meta-analyses studies in identifying risk loci for Crohn's, genetic contribution explains only 20% of the heritability of Crohn's disease. Relatively low concordance rates in monozygotic twin studies and evidence of low genetic contribution to heredity suggests large immunobiological and environmental contribution to this disease (Park et al., 2010; Halfvarson, 2011).

1.3.2 Environmental factors

Although the incidence of IBD was more common in Western countries throughout the middle of the 20th century, the number of diagnosed cases in developing countries is increasing each year suggesting this changing epidemiology is related to urbanization and industrialization (Kaplan et al., 2016; Kaplan et al., 2017). Also, evidence of increasing rate among the immigrants from low incidence regions to high incidence regions suggests an important role of the environment and lifestyle (Joossens et al., 2007). The potential risk factors for IBD associated with urbanization and change in lifestyle include a change in diet, antibiotic use, exposure to pollution, microbial exposures, and hygiene status. Cigarette smoking, both active and passive, is one of the well-studied environmental risk factors in Crohn's disease. Early tobacco use doubles the risk of developing the disorder while neither carbon monoxide nor nicotine has been directly related to the disease (Tuvlin et al., 2007; Mahid et al., 2006). An understanding of environmental influences is also critical to assess the central role of gut microbiota in the development and progression of IBD (Kostic et al., 2014). The disease often appears to be occurring after gastroenteritis and is associated with dysbiosis that results in a lower diversity of gut microbiota (García et al., 2006; Joossens et al., 2011).

1.3.3 Microbiota

Genome-wide association studies highlight risk loci for Crohn's and mycobacterial infections and reveal genes of adaptive and innate immune processes in the containment of gut microbes (Hampe et al. et al., 2007; Rioux et al., 2007). 16S ribosomal RNA (rRNA) gene sequencing suggests thousands of mostly anaerobic commensal species colonize the human gut that belongs to four major phyla (Bacteroidetes, Firmicutes, Actinobacteria, and Proteobacteria). The microbiota varies greatly along the length of the gastrointestinal tract within an individual and varies among different individuals dictated by genetics, dietary, or drug-induced factors (Muegge et al., 2011; Dethlefsen et al., 2011). The reduced diversity especially with Firmicutes and Bacteroidetes phyla is seen in patients with Crohn's disease and is related to an elevated risk of postsurgical recurrence of ileal Crohn's disease (Qin et al., 2010; Frank et al., 2007; Sokol et al., 2008). Further, metagenomics from fecal samples revealed that by the end of the first year after birth the microbial ecosystem converged towards that of adults (Palmer et al., 2007). This highlights the importance of avoiding early life influences that could cause dysbiosis and potentially cause increased risk to IBD. Despite the risks, research in mice with mutations in human susceptibility loci confirms that reduced diversity alone is not enough to trigger disease-related phenotypes (Bloom et al., 2011).

1.3.4 Innate and adaptive immunity

Toll-like receptors (TLR) and nucleotide-binding domain (NOD) like receptors (NLR) are among pattern recognition receptors in dendritic cells that interpret evolutionary conserved microbial and viral associated molecular patterns to direct other immune cells towards tolerance or immunity. Innate immunity response is rapid but limited. Overexpression of TLR2 (Hart et al., 2005) and TLR4 (Baumgart et al., 2009) along with exaggerated lipopolysaccharides response (Baumgart et al., 2009) have been reported in Cohn's disease. In contrast, induced deletion of

TLR5 drove intestinal inflammation in murine models (Vijay-Kumar et al., 2007). In other research, TLR induced signaling through MyD88 prevented inflammation in animals suggesting the dual role of TLR receptors (Rakoff-Nahoum et al., 2004). Similarly, polymorphisms in NOD-like receptors are strongly correlated with Crohn's disease. Homozygous mutations in the NOD2 gene risk 20 to 40 times higher chances of developing Crohn's in a patient while heterozygous mutations have only 2 to 4 times higher chance (Ogura et al., 2001; Philpott et al., 2004). Polymorphisms in NOD2 are associated with weaker cytokine response and impaired autophagy (Cooney et al., 2010) due to which adaptive T-cell response is not induced properly.

Leukocytes are rapidly recruited and inappropriately retained in mucosa mediated by selectins, integrins and their ligands from the immunoglobulin superfamily (such as ICAM-1, MAdCAM-1), chemokines (such as CCL20) and their receptors (chemokine receptor 9), and fibronectin in Crohn's disease. Adaptive immunity is considered to retain and propagate intestinal inflammation but probably not start it. The two main components of T-cell mediated immunity**,** effector T-cells and inducible regulatory T-cells (iTreg), originate from the same precursor cells but their differentiation is inversely regulated to maintain a balance between immunity and unrestrained inflammation (Zhou at al., 2008). Crohn's disease is characterized by the imbalance of these two T-cells reported by serological studies and supported by GWAS. GWAS link gene loci associated in Treg (such as IL10, IL2RA, SMAD3) and effector T-cells (Th1 and Th17) differentiation (such as CPEB4) to the disorder (Franke et al., 2010). For example, gene mutations reported in IL10R*,* involved in T-cell regulation, have been linked to the early onset of the disease possibly because of dysregulated T-cell balance (Glocker et al., 2009).

1.3.5 Intestinal epithelial barrier

GWAS has identified multiple genes as risk loci with restricted expression in intestinal epithelial cells such as related to cell adhesion (CDH1, LAMB1), cell polarity (PARD3), tight junction assembly and regulation (GNAI2, MAGI2, MYO9B, PTPN2), epithelial differentiation (HNF4A), mucin regulation (ECM1, MUC3A, MUC1, MUC19), membrane transport (ITLN1), epithelial restitution (PTGER4), microbial sensing (CARD15), and membrane receptor kinase regulation (ERRFI1) (Declan F. McCole, 2014; Franke et al., 2010). Later, using an ex vivo approach, it was found that permeability is increased in both ulcerated and non-ulcerated epithelia of Crohn's patients (Pearson et al., 1982; Hollander et al., 1988; Ukabam et al., 1983). Subsequently, more studies revealed that permeability was a result of alteration in the expression of tight junction proteins such as claudins (Zeissig et al., 2007, Prasad et al., 2005). Such findings even led to the use of permeability as a sensitive prognostic indicator of recurrence in clinical remission (Wyatt et al., 1993; D'Incà et al., 1999). More studies have linked dysfunctional epithelial barrier to IBD since then. Paneth cells are specialized epithelial cells that defend the mucosa by secretion of antimicrobial peptides such as defensins as well as regulate commensal microbiota (Salzman at al., 2010). Deficiencies in anti-microbial functions of Paneth cells (Wehkamp et al., 2005), decreased expression of antimicrobial peptide granules (Cadwell et al., 2008), and defective autophagy processing of microbial antigens by Paneth cells (Hampe et al., 2007; Stappenbeck and McGovern, 2017) are reported in patients with the disorder.

Decreased expression of mucins, secreted by goblet cells of epithelia which act as a protective layer, is often correlated with inflamed regions in the terminal ileum of Crohn's patients (Buisine et al., 1999). But the decreased expression of mucins and antimicrobial peptides have been previously linked with endoplasmic reticulum (ER) stress in secretory goblet

and Paneth cells possibly because of existing inflammation (Kaser et al., 2010). In an ER-stress response study, deletion of XBP1 protein (a key component of ER-stress response) in epithelia of mice led to the absence of Paneth cells as well as a reduced number of goblet cells (Kaser et al., 2008). Those mice also showed spontaneous small intestinal inflammation and heightened sensitivity towards induced colitis. In contrast to this, a study in AGR2-/- mice showed diminished mucin and abnormal Paneth cell expansion preceding the inflammation and showed increased susceptibility to induced colitis (Park et al., 2009; Zhao et al., 2010). Anterior gradient homolog 2 (AGR2) is expressed by epithelial secretory cells and it has been shown to play a critical role in intestinal mucus production such as MUC2. AGR2-/- mice lacked intestinal mucus and MUC2 protein, showed abnormal expansion and localization of Paneth cells, showed increased expression of proinflammatory cytokines, and showed increased susceptibility to DSS-induced colitis. A minimal level of ER-stress response was reported and since abnormal Paneth cell expansion and localization preceded the inflammation it could not be an inflammatory response as thought before. Despite the controversy of whether diminished mucin and abnormal Paneth cells are a result of inflammation ongoing or are a cause to begin inflammation, the role of the epithelial barrier is important in IBD. Consistent with this, a study in GPA33-/- (tight junction protein) mice also showed increased permeability upon induced colitis as well as elevated levels of immune activity in colonic mucosa (Williams et al., 2015).

A recent study with a single-cell RNA sequencing profile of whole intestinal cells from Crohn's patients showed enrichment of inflammation-associated genes (such as HLA-DRA, HLA-C, CD74, LCN2, HLA-DRB1) including chemokines (such as CCL20, CXCL1) in a group of undifferentiated epithelial stem cell-like populations (Parikh et al., 2019). Similarly, another study showed enriched pathways from RNA-sequencing of epithelial cells from pediatric Crohn's patients are mostly among immune response (such as antigen processing, chemokine

receptors, MHC II antigen presentation, interleukin signaling, etc.) (Howell et al., 2018). In addition to the possibility of a dysfunctional epithelial barrier being a cause for inflammation, these studies suggest that epithelial cells might even directly dysregulate the immune system which can manifest IBD such as ulcerative colitis and Crohn's disease.

1.4 Animal models

The current approach to therapy in Crohn's disease is largely towards alleviating symptoms using immunomodulators, immunosuppressive, and corticosteroids. The etiology of this disease remains obscure. Hence the treatment strategies remain largely nonspecific. In such cases, the significance of animal models becomes important as they can improve our understanding of pathogenic mechanisms underlying this disease. The use of animals allows investigators to manipulate genes of interest possibly involved in pathogenesis and study-specific physiological events occurring before the onset of the disease. Moreover, various approaches like drug trials can be safely investigated in animal models which gives a practical advantage compared to patient investigations. Several mouse models of intestinal inflammation have provided significant information about mucosal immunity, intestinal homeostasis, and even aided in the development of potential therapies for the past two decades. One of the biggest contributions of animal models has been the development of anti-cytokine therapies for Crohn's disease patients (Neurath, M.F., 2014). The use of anti-IL-12, anti-IL23 (Kashani and Schwartz, 2019), anti-IL-18 (Maerten et al., 2004), T helper 1 (Th1) cell blockers (Imam et al., 2018), anti-IFNγ (Ito et al., 2006), and anti-TNF (Adegbola et al., 2018) treatments in Crohn's disease are the contributions of animal models of intestinal inflammation. The mouse has been the most widely used species to establish a variety of models of inflammation including chemically induced, immune-mediated, and genetically manipulated.

1.4.1 Chemically induced models

Dextran sodium sulfate (DSS), a negatively charged polysaccharide, induced colitis is possibly the most widely used mouse model of intestinal inflammation which closely resembles human ulcerative colitis (Okayashu et al., 1990). Administration of DSS causes severe colitis evidenced by damage of epithelial cells, infiltrations with neutrophils, and ulcerations in pathology, causes weight loss, and bloody diarrhea. While the exact mechanism by which DSS exerts toxicity to epithelial cells is unknown, it compromises the epithelial barrier exposing submucosal immune cells and underlying tissues to luminal microbiome and antigens. This result is prompt and intense inflammatory response is limited to the large intestine (Okayashu et al., 1990). Both acute and chronic inflammation can be achieved in a reproducible manner by simply modifying the concentration and duration of the treatment which makes it popular in IBD research (Wirtz et al., 2007). Moreover, accumulating evidence of colitis in severe combined immunodeficiency (SCID) and Rag-/- mice reveals that adaptive immunity might not be required for this model (Dieleman et al., 1994). Thus, DSS has been very widely used in studying innate immunity mechanisms in the development of intestinal inflammation and reestablishment of barrier integrity through tissue repair.

Trinitrobenzenesulfonic (TNBS) acid acts as a hapten to bind colonic or microbiome proteins of high molecular weight to elicit immunological response against those proteins (Morris et al., 1989). It is administered intrarectally with ethanol which disrupts the barrier to allow TNBS to access intestinal epithelial cells as well as penetrate the bowel wall. TNBS-induced colitis exerts cell-mediated immune response especially T helper 1 (Th1) inflammation which causes CD4 T cells infiltration and secretion of proinflammatory cytokines (Antoniou et al., 2016). Most importantly, fibrosis in lamina propria is observed with this model which makes it useful in studying the mechanism underlying it. Similarly, oxazolone is another hapten administered

intrarectally which leads to elevated production of various cytokines in the distal colon of mice (Kiesler et al., 2015).

1.4.2 Immunological models

T cells or bone marrow precursor cells adoptively transferred into immunodeficient mice constitute immunologically mediated models. Transfer of $CD4^+$ T cells that express a high level of *CD45RB* into congenic severe combined immunodeficient (SCID) mice caused a wasting disease with significant lymphoid cell accumulation in intestinal lamina propria (Morrissey et al., 1993). Also, T cells selected in an aberrant thymic microenvironment and transferred into mice showed severe colitis in the colon (Hollander et al., 1995). Similarly, the transfer of $CD8^+$ T cells in mice showed inflammation in the small intestine (Steinhoff et al., 1999). Such models are critical in understanding how pathogenic and regulatory T cells control intestinal inflammation.

1.4.3 Genetic models

Gene knockout models such as IL-10 (Kuhn et al., 1993), IL-2 (Sadlack et al., 1993), TCRα/β (Mombaerts et al., 1993), and Gi2-α (Rudolph, 1995) knockout mice and transgenic models such as E-cadherin transgenic mouse have helped to identify key regulators in gut inflammation. For example, IL-10 deficient mice showed arrested growth, anemia, chronic enterocolitis, extensive mucosal hyperplasia, and an aberrant expression of MHC class II molecules (Kuhn et al., 1993). Such models have helped to reveal important immune-related molecules in chronic intestinal inflammation. Similarly, knockouts of genes specific to intestinal epithelial such as GPA33-/- (Williams et al., 2015) and AGR2-/- (Park et al., 2009; Zhao et al., 2010) have revealed the role of epithelial barrier integrity in mucosal immunity.

1.4.4 Spontaneous models

Spontaneous models have also gained traction in recent years as inflammation occurs in a way similar to human disease without any extra manipulations. C3H/HeJBir model is one such example where spontaneous and chronic inflammation is observed in the colonic and cecal regions in mice at a young age (Sundberg et al., 1994).

CHAPTER 2: MATERIALS AND METHODS

2.1 Materials

2.1.1 List of reagents and commercial kits

REAGENTS/ KITS	SOURCE	IDENTIFIER
Trizol Reagent	Thermo Fisher	Cat# 15596018
Feeder removal beads	Miltenyi Biotec	Cat# 130-095-531
Power SYBR Green PCR master mix	Thermo Fisher	Cat# 4368706
TrypLE Express Enzyme (1X), phenol red	Thermo Fisher	Cat# 12605036
Corning Matrigel GFR Membrane Matrix	Fisher Scienitific	Cat# CB-40230
Collagenase Type IV	Life Technologies	Cat# 17104-019
Fixation/Permeabilization Solution Kit (Cytofix/Cytoperm)	BD Biosciences	Cat# 554714
Maxima First Strand cDNA Synthesis Kit for RT-qPCR, with dsDNase	Thermo Fisher	Cat# K1671
DNeasy Blood & Tissue Kit (250)	QIAGEN	Cat# 69506
RNeasy Mini Kit (250)	QIAGEN	Cat# 74106
GeneChip Human Exon 1.0 ST Array	Affymetrix	Cat# 900650

2.1.2 List of antibodies

ANTIBODIES	SOURCE	IDENTIFIER
Goat polyclonal E-Cadherin	R&D Systems	RRID: AB_355504
Rabbit monoclonal human SOX9	Abcam	RRID:AB_2728660
Mouse monoclonal human CEACAM5	R&D systems	RRID:AB_1207848
Rabbit polyclonal human VSIG1	Sigma-Aldrich	RRID:AB_10670795
Mouse monoclonal human PSCA	Santa Cruz Biotechnology	RRID:AB_1128761
Mouse monoclonal human MUC2	Santa Cruz Biotechnology	RRID:AB_2815005
Rabbit polyclonal human CHGA	Abcam	RRID:AB_301704
Rabbit polyclonal KI67	Abcam	RRID:AB_302459
Rabbit polyclonal human DEFA6	Sigma-Aldrich	RRID:AB_1847595
Rabbit polyclonal human GPA33	Abcam	RRID:AB_10867420
Rabbit polyclonal human CLDN18	Sigma-Aldrich	RRID:AB_2669866
Rabbit polyclonal human LCN2	Abcam	RRID:AB_956105
Rabbit polyclonal human MUC5AC	Abcam	RRID:AB_11143557
Mouse monoclonal STEM121	Takara Bio	RRID:AB_2801314
Mouse monoclonal CCL20	Thermo Fisher	RRID:AB_2608262
Rabbit polyclonal human TNFRSF1A	Sigma-Aldrich	RRID:AB_1846232
Rat monoclonal mouse CD45	Thermo Fisher	RRID: AB_657749
Rat monoclonal mouse Ly6G	R&D systems	RRID: AB_2232806
Mouse monoclonal α-SMA	Abcam	RRID: AB_262054
Rabbit polyclonal FN1	Abcam	RRID:AB_2262874

2.1.3 List of qPCR primers

GENE	FORWARD SEQUENCE	REVERSE SEQUENCE
Human CEACAM5	5'- CTTCATTTCAGGAAGACTGAC	5'- TTAGTAGAGATGGGGTTTCAC
Human CDX1	5'- GGAGAAGGAGTTTCATTACAG	5'-TGCTGTTTCTTCTTGTTCAC
Human HOXA11	5'- CTCCCATTGAATCTCCTTTG	5'- TCTTTATTTTCCTTGTGCCC
Human NOX1	5'- CCGGTCATTCTTTATATCTGTG	5'-CAACCTTGGTAATCACAACC
Human HOXA13	5'- ACGCAGTTACTGAAATCTTG	5'- AAGGACAAGCAGATGTTTAC
Human GPA33	5'- GGTTACTACATCTGTACCTCC	5'- TGATGCCAATGATAATGAGG
Human VSIG1	5'- TTCTTCACATCCAGAAGTTG	5'- TGCTTTTGCCTTATTCCTTG
Human CLDN18	5'- CATTGTCTCAGGTCTTTGTG	5'- TGTGGACATCCAGAAGTTAG
Human CD74	5'- GAGTCACTGGAACTGGAG	5'- CATGGGATGAGGTACAGG
Human LCN2	5'- GGAAAAAGAAGTGTGACTACTG	5'- GTAACTCTTAATGTTGCCCAG
Human PSCA	5'- AAGTGGACTGAGTAGAACTG	5'- GTGTTTATTAAGGGCCTACG
Human TFF2	5'- GAAGACTGCCATTACTAAGAG	5'- AGTTTCTTCTTTGGTTTCGG
Human MUC5AC	5'- AATGGTGGAGATTTTGACAC	5'-TTCTTGTTCAGGCAAATCAG
Human DPCR1	5'- AGGTGCTACTACATTCCAAG	5'- TACTATAAACAAGGCCTGGAG
Human EMB	5'- ACTTCATGGGAAAAACAAGC	5'- CCTTTACACTCCCATTACTAC
Human SPINK1	5'- TTCTGAAGAGACGTGGTAAG	5'- TGTTACCAGATAGACTCAACA
Human PRSS1	5'- GATTCATGTCAGGGTGATTC	5'- GTAGACCTTGGTGTAGACTC
Human GAPDH	5'- GTCTCCTCTGACTTCAACAGCG	5'- ACCACCCTGTTGCTGTAGCCA
Human TNFRSF1A	5'- CCCCTGGTCATTTTCTTTG	5'- ATTTCCCACAAACAATGGAG
Human CCL20	5'- AAGTTGTCTGTGTGCGCAAATC	5'- CCATTCCAGAAAAGCCACAGT
Mouse GAPDH	5'- CATGGCCTTCCGTGTTCCTA	5'- TGTCATCATACTTGGCAGGTTT
Mouse α-SMA	5'- GTCCCAGACATCAGGGAGTAA	5'- TCGGATACTTCAGCGTCAGGA
Mouse CD45	5'- ATGGTCCTCTGAATAAAGCCCA	5'- TCAGCACTATTGGTAGGCTCC

2.1.4 Experimental models: cell lines

CELL LINES	SOURCE	IDENTIFIER
GP2-293	Clontech	Cat# 631458

2.1.5 Experimental models: organisms/strains

ORGANISMS/STRAINS	SOURCE	IDENTIFIER
Mouse: NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ	The Jackson Laboratory	Stock No:005557 | NSG

2.1.6 Biological samples

BIOLOGICAL SAMPLES	SOURCE	IDENTIFIER
Human biopsies from Crohn's and control patients	UCONN Health; UNC; BCM	See Table 1 for a list of patients included in this study
Human resected terminal ileum from Crohn's patients	BWH	

2.2 Experimental model and subject details

2.2.1 Human subjects

Terminal ileum endoscopic biopsies were obtained from pediatric or adult Crohn's patients, functional controls lacking mucosal inflammation, and fetal demise cases under informed parental consent and institutional review board approval at the Connecticut Children's Medical Center, Hartford, CT, USA, the University of North Carolina, Chapel Hill, Chapel Hill, NC, USA, the Brigham and Women's Hospital, Boston, MA, USA, and Baylor College of Medicine, Houston, TX, USA. Patient-derived epithelial cell lines were derived from the biopsies as described in the method below. Patient data is listed in Table 1.

2.2.2 Primary cell culture

The primary cells used in this study were derived from human terminal ileum biopsies or fetal demise intestine. All cells were cultured in the culture medium described below. The cells were tested as mycoplasma-free. Short tandem repeats (STR) profiling data from clones and patient blood were used as a genotyping marker to ensure the authentication of each cell line.

2.2.3 Animals

All animal experiments were approved by the Animal Care Committee at the University of Houston (IACUC 16-002). Intestine stem cell variants were subcutaneously transplanted to NSG (NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ) mice. Animal studies were conducted under maximum containment in an animal biosafety level 2 facility. NSG mice were immune-compromised, drug or test naive, and housed in a sterile condition. Both male and female mice (6-8 weeks old) were utilized in these studies. At 1-2 weeks post-transplant, mice were humanely euthanized according to IACUC-approved criteria.

2.2.4 Cell lines

Human GP2-293 cell line was obtained from Clontech. GP2-293 cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with penicillin, streptomycin, and glutamine. All cell lines were incubated at 37 °C with 7.5% CO_2.

2.3 Method details

2.3.1 *In vitro* culture of human terminal ileum epithelial stem cells

Mucosal stem cell libraries (Wang et al., 2015; Duleba et al., 2020) were generated from 1 mm endoscopic biopsies which were collected into cold F12 media (Gibco, USA) with 5% fetal bovine serum (Hyclone, USA), and then were minced by sterile scalpel into 0.2-0.5 mm^3 fragments. The minced tissue was digested in 2 mg/ml collagenase type IV (Gibco, USA) at 37 °C for 30-60 minutes with agitation. Dissociated cells were passed through a 70 µm Nylon mesh (Falcon, USA) to remove masses and then were washed four times in cold F12 media, and seeded onto a feeder layer of lethally irradiated 3T3-J2 cells in c-FAD media containing 125 ng/mL R-Spondin1 (R&D systems, USA), 1 µM Jagged-1 (AnaSpec Inc, USA), 100 ng/ml Human Noggin (Peprotech, USA), 2.5 µM Rock-inhibitor (Calbiochem, USA), 2 µM SB431542 (Cayman Chemical, USA), 10 mM nicotinamide (Sigma-Aldrich, USA). Cells were cultured at 37 °C in a 7.5% CO_2 incubator. The culture media was changed every two days. Colonies were digested by 0.25% trypsin-EDTA solution (Gibco, USA) for 5-8 min and passaged every 7 to 10 days. Colonies were trypsinized by TrypLE Express solution (Gibco, USA) for 8-15 min at 37 °C and cell suspensions were passed through 30 µm filters (Miltenyi Biotec, Germany). Approximately 20,000 epithelial cells were seeded to each well of a 6-well plate. Cloning cylinder (Pyrex, USA) and high vacuum grease (Dow Corning, USA) were used to select single

colonies for pedigrees. Gene expression analyses were performed on cells derived from passage 4-10 (P4-P10) cultures.

2.3.2 Stem cell differentiation

Air-liquid interface (ALI) culture of terminal ileum epithelial cells was performed as described (Wang et al., 2015). Briefly, Transwell inserts (Corning Incorporated, USA) were coated with 20% Matrigel (BD biosciences, USA) and incubated at 37 °C for 30 min to polymerize. 200,000 irradiated 3T3-J2 cells were seeded to each *Transwell* insert and incubated at 37 °C, 7.5% CO_2 incubator overnight. QuadroMACS Starting Kit (LS) (Miltenyi Biotec, Germany) was used to purify the stem cells by the removal of feeder cells. 200,000-300,000 stem cells were seeded into each Transwell insert and cultured with stem cell media. At confluency (3-7 days), the apical media on the inserts was removed through careful pipetting, and the cultures were continued in differentiation media (stem cell media without nicotinamide) for an additional 6-12 days prior to harvesting. The differentiation media was changed every one or two days.

2.3.3 Histology and immunostaining

Histology, hematoxylin and eosin (H&E) staining, immunohistochemistry, and immunofluorescence were performed using standard techniques. For immunofluorescence and immunohistochemistry, 4% paraformaldehyde-fixed, paraffin-embedded tissue slides were subjected to antigen retrieval in citrate buffer (pH 6.0, Sigma-Aldrich, USA) at 120 °C for 20 min, and a blocking procedure was performed with 5% bovine serum albumin (BSA, Sigma-Aldrich, USA) and 0.05% Triton X-100 (Sigma-Aldrich, USA) in DPBS(-) (Gibco, USA) at room temperature for 1 hr. The sources of primary antibodies used in this study, including anti-mucin 2 (sc-515032; Santa Cruz Biotechnology, USA), -Ki67 (550609; BD Biosciences, USA), -chromogranin A (ab15160; Abcam, UK), -alpha defensin 6 (HPA019462; Sigma-Aldrich, USA), -

E-cadherin (AF648; R&D Systems, USA), –GPA33 (ab108938; Abcam, UK), -Ceacam5 (MAB41281; Novus Biologicals, USA), -PSCA (sc-80654; Santa-Cruz Biotechnology, USA), -Lipocalin2 (ab41105; Abcam, UK), -VSIG1 (HPA036310; Sigma-Aldrich, USA), -CLDN18 (HPA018446; Sigma-Aldrich, USA), -MUC5AC (ab78660; Abcam, UK), -SOX9 (ab185966; Abcam, UK), -CCL20 (MA523843; Thermofisher Scientific, USA), -TNFRSF1A (HPA004102; Sigma-Aldrich, USA), -Alpha Smooth Muscle Actin (ab7817; Abcam, UK), -Fibronectin1 (ab2413; Abcam, UK), -STEM121 (Y40410; Clonetech Laboratories, USA), -CD45 (14-0451-85; Thermo Fisher, USA) and -LY6G (MAB1037; R&D systems, USA) are listed (table S3). Secondary antibodies used here are Alexa Fluor-488 or Alexa Fluor-594 Donkey anti-goat/mouse/rabbit IgG antibody (Thermo Fisher, USA). All images were captured by using the Inverted Eclipse Ti-Series (Nikon, Japan) microscope with Lumencor SOLA light engine and Andor Technology Clara Interline CCD camera and NIS-Elements Advanced Research v.4.13 software (Nikon, Japan) or LSM 780 confocal microscope (Carl Zeiss, Germany) with LSM software. Bright field cell culture images were obtained on an Eclipse TS100 microscope (Nikon, Japan) with Digital Sight DSFi1camera (Nikon, Japan) and NIS-Elements F3.0 software (Nikon, Japan).

2.3.4 Xenografts in immunodeficient mice

Two to three million epithelial cells were harvested by trypsinization, mixed with 50% Matrigel (Becton Dickinson, Palo Alto) to a volume of 100 ul, and injected subcutaneously in NSG (NODscid IL2ranull) (Shultz et al., 2012) mice (Jackson Laboratories, Bar Harbor) and harvested one or two weeks later.

2.3.5 Flow cytometry analysis

Clonogenic cell libraries from patients with or without Crohn's were trypsinized and harvested as a single cell suspension. Feeders were removed as mentioned above and approximately 300,000 epithelial cells were fixed and permeabilized by using Fixation/Permeabilization Solution Kit (BD biosciences, USA, cat. 554714). After a blocking procedure with Permeabilization solution at 4 °C for 30 min, cells were incubated with primary and Alexa Fluor 488 Secondary antibodies (Thermo Fisher, USA) for 1hr at 4 °C, with five washing events between each step. Primary antibodies used in these experiments include mouse monoclonal anti-Ceacam5 antibody (MAB41281; Novus Biologicals, USA), rabbit polyclonal VSIG1 antibody (HPA036310; Sigma-Aldrich, USA), and mouse monoclonal anti-human PSCA antibody (sc-80654; Santa-Cruz Biotechnology, USA). Samples were collected and analyzed on a Sony SH800S Cell Sorter (Sony Biotechnology, USA).

2.3.6 RNA sample preparation

For stem cell colonies, RNA was isolated using the PicoPure RNA Isolation Kit (Life Technologies, USA). For ALI structure, RNA was isolated using the Trizol RNA Isolation Kit (Life Technologies, USA). RNA quality (RNA integrity number, RIN) was measured by Agilent 2100 Bioanalyzer and Agilent RNA 6000 Nano Kit (Agilent Technologies, USA). RNAs having a RIN > 8 were used for microarray analysis.

2.3.7 Sequence alignment of single-cell RNA sequencing

The single-cell mRNA sequencing (scRNA-seq) libraries were established using the 10X Genomics Chromium system (Single Cell 3' Reagent Kit v2). The scRNA-seq libraries were sequenced on the Illumina HiSeq X Ten with 10K cells for Crohn's case and fetal TI case. For normal cases, the scRNA-seq library was sequenced on the Illumina NextSeq 500 with 2K cells.

Demultiplexing, alignment, and UMI-collapsing were performed using the Cellranger toolkit (version 2.1.0, 10X Genomics) (Ferguson and Chen, 2020). The raw paired-end reads were trimmed to 26 bps for Read1 and 98 bps for Read2. The trimmed reads were mapped to both the human genome (hg19) and the mouse genome (mm10). The reads uniquely mapped to the human genome were used for downstream analysis.

2.3.8 Single-cell RNA sequencing

The scRNA-seq data analyses were performed using the Seurat package (version 2.3.4; Satija et al., 2015). We kept the genes with expression in at least three cells, and excluded cells expressing less than 200 genes. We identify the cell with SOX9 high expression as stem cells and excluded the cells with high mitochondrial percentage or with an outlier level of UMI content. The normalization was performed using the global-scaling normalization method, which normalizes the gene expression measurements for each cell by the total expression, and then multiplies by 10,000, and finally log-transforms the result. The variable genes were identified using a function to calculate average expression and dispersion for each gene, divides these genes into bins, and then calculates a z-score for dispersion within each bin ("x.low.cutoff = 0.0125", "x.high.cutoff = 3", and "y.cutoff = 0.5"). We scaled the data to regress out the variation of mitochondrial gene expression.

We performed PCA based on the scaled data to identify significant principal components (PCs). We selected the PCs with p-values less than 0.01 as input to perform clustering analysis and visualization by t-SNE. We detected the marker genes in each cell subpopulation using two methods of Wilcoxon rank-sum test and DESeq2. For the Wilcoxon rank-sum test, we used the default parameter. For DEseq2, we kept the marker genes with the average log-fold change above 0.1 and adjust the p-value fewer than 0.05.

Contaminating 3T3-J2 fibroblast cells were identified by murine reads. In addition, the cells in the S stage of the cell cycle were identified based on the marker gene of SLBP (Nestorowa et al., 2016). The cells in the G2 or M stage of the cell cycle were identified based on the marker genes of UBE2C, AURKA, CENPA, CDC20, HMGB2, CKS2, and CKS1B. The cells in the G0 stage of the cell cycle were identified based on the marker genes of G0S2. In addition, the ambiguous cells with few marker genes were also removed, which could correspond to sequencing low-quality cells. Finally, we integrated the clean data of normal and Crohn's cases to perform clustering analysis and visualization by t-SNE.

2.3.9 Expression microarray data analysis

Total RNAs obtained from immature colonies and ALI-differentiated epithelia were used for microarray preparation with WT Pico RNA Amplification System V2 for amplification of DNA and Encore Biotin Module for fragmentation and biotin labeling (NuGEN Technologies, USA). All samples were prepared according to the manufacturer's instructions and hybridized onto GeneChip Human Exon 1.0 ST array (Affymetrix, USA). GeneChip operating software was used to process Cel files and calculate probe intensity values. To validate sample quality, quality checks were conducted using Affymetrix Expression Console software. The intensity values were log2-transformed and imported into the Partek Genomics Suite 6.6 (Partek Incorporated, USA) (Wodehouse et al., 2019). Exons were summarized to genes and a 1-way ANOVA was performed to identify differentially expressed genes. The heatmaps with hierarchical clustering analysis of the global gene expression pattern in different regions were performed using the pheatmap package (Li et al., 2018) (https://cran.rproject.org/web/package/pheatmap/index.html) in R (version 3.5.1). The pathway enrichment analysis was performed using Enrichr (Chen et al., 2013) based on WikiPathways (Slenter et al, 2018) database, and tissue enrichment analysis performed using ARCHS4 Tissues (Lachmann et al., 2018). The network analysis was

constructed using ClueGO (v 2.5.4) (Bindea et al., 2009) and CluePedia (v 1.5.4) (Bindea et al., 2013) plug-ins of Cytoscape (Shannon et al., 2003), based on the KEGG (Kanehisa et al, 2000) and Reactome database (Wu and Haw, 2017).

2.3.10 Expression qPCR array

For clonal analysis from a pool of stem cells derived from each patient, an array of genes with significance in Crohn's disease were used to run a quantitative real-time PCR on 275 clones from 21 patients. The genes include CDX1, HOXA11, NOX1, HOXA13, GPA33, CEACAM5, VSIG1, CLDN18, CD74, LCN2, PSCA, TFF2, MUC5AC, DPCR1, EMB, SPINK1, and PRSS1. The genes listed were selected based on the microarray (5-fold, $p<0.05$). cDNA was generated using the High Fidelity cDNA synthesis kit (Applied Biosystems, USA). qPCR was performed using SYBR green dye (Applied Biosystems, USA) on QuantStudio 6 Flex Real-Time PCR system (Applied Biosystems, USA) using primers designed to amplify genes of interest. Expression data were normalized to GAPDH.

2.3.11 Statistical analysis

Unpaired two-tailed student's t-test was used to determine the statistical significance between the two groups. Statistical analyses were performed using R (version 3.5.1). The ''n'' numbers for each experiment are provided in the text and figures. $P < 0.05$ was considered statistically significant. Asterisks denote corresponding statistical significance *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$ and ****$p < 0.0001$.

2.3.12 *In vitro* culture of GFP-tagged cells in multi-well plates

Clones representative of three clusters were selected based on expression qPCR array, FACS profiling, ALI, and xenograft characterization from each possible Crohn's patient. Along with

these, clones from control patients were also selected. Retrovirus carrying green fluorescent protein (GFP) reporter was infected for genetic labeling in the selected clones. The pseudotyping vector pVSV-G and pMX-IRES-eGFP plasmid were transfected into GP2 retroviral packaging cell line following the manufacturer's guidelines (jetPRIME, Polyplus Transfection). The virus was harvested and concentrated 48 hours later. Retroviral infection was performed when the stem cell colonies were small, and the infected cells were sorted after 48 hours to get a pure population of GFP-positive cell lines. The GFP-positive cell lines were cultured in 384 multi-well plates with a feeder layer of lethally irradiated 3T3-J2 cells (Greiner Bio-One, USA) using the methodology described above.

2.3.13 High throughput screening and imaging of cell lines

The GFP-tagged cell lines were seeded on multiple 384 well plates (varying according to the number of compounds in a library) with a feeder layer of irradiated 3T3-J2 fibroblast cells. The stem cells were allowed to grow until they divided to become 4-5 cells within a colony. Afterward, they were transported for treatment with four selected chemical libraries (1 μM) to High Throughput Research and Screening Center at the Institute of Biosciences and Technology (Houston, Texas), Texas A&M University. The selected libraries included Custom Clinical Library, Prestwick Chemical Library, Selleck Bioactive Chemical Library, and UT Austin Kinase Inhibitors Library. Positive and negative control lanes were allocated within each plate. A highly potent drug was used as a positive control and negative control was just DMSO since most of the drugs are dissolved in DMSO. The cells were allowed to grow for 5 days at 37 °C, 7.5% CO_2 incubator after treatment. After the cells in the negative control lanes were at good confluency, the cells were prepared for imaging. Each multi-well plates were washed with Phosphate Buffered Saline (Gibco, USA) and fixed with 4% paraformaldehyde at room temperature for 25 minutes. Paraformaldehyde was replaced with Phosphate Buffered Saline

(PBS). The plates were imaged using Inverted Eclipse Ti-Series (Nikon, Japan) microscope with Lumencor SOLA light engine, paired with High Content Microscope system (Nikon, Japan), Andor Technology Clara Interline CCD camera, NIS-Elements Advanced Research v.4.13 software (Nikon, Japan), and NIS-Elements HC software (Nikon, Japan). High Content Analysis (HCA) system built on NIS-Elements platform streamlined with automated well plate acquisition and multiple-well plate job run was used for high throughput imaging of phase contrast as well as FITC channel.

2.3.14 High content image analysis and selection of drugs

NIS-Elements High Content Analysis (HCA) system was used for image data management of multiple-well plate job runs. The cells labeled with Green Fluorescent Protein (GFP) in each well of multi-well plates were imaged with features of each stem cell colonies. The changes in cell phenotype compared to untreated control lanes were measured based on the fluorescent signal threshold using automated image analysis. The features like area, colony number were exported from the automated analysis. Treated wells were compared to untreated wells based on GFP-signal area and the number of stem cell colonies. The treated well was normalized with the untreated well based on the area which was represented in terms of survival rate and was compared between control and variant clones. Z-score (Zhang et al., 1999; Iversen et al., 2006) was calculated based on the difference between positive and negative control and was used as criteria for assessing the quality of runs. Only plates with Z'-factor >0.6 were used. The compounds with a coefficient of variation between plate duplicates larger than 20% were ignored. B-score (Malo et al., 2006) was calculated via R package platetools v0.0.2 (https://github.com/swarchal/platetools) to control the edge and positional bias to help infer the potential hits of inhibitors. Only the compounds with a cutoff of B-score < -2 were considered as potential inhibitors. The selection of drugs was made based on the maximum differences in

survival rate between patient-matched cell lines (cut-off set at 20%), their targets, structural spectrum, pathways related to them, and their possible relation implicated in Crohn's disease. If there were not patient-matched pedigrees, the median value of survival rates in a contained group was used as a representative value for that group of pedigrees.

2.3.15 Drug validation and establishing dosage response curve

The selected drugs were further validated in the cell lines used for high throughput screening. 96 multi-well plates (Greiner Bio-One, USA) were used to generate dose-response curves for each selected hit. Fresh samples of compounds were used for this purpose and half-maximal inhibitory concentration was generated using the curve. With a negative control setup for each drug, the cells were allowed to grow for 6 days after treatment and then replated to observe how drugs affected the clonogenicity after treatment. The dosage-response curves of survival rate for a certain compound were calculated by fitting a three-parameter log-logistic dose-response model to the survival rate data using R package *drc* (Ritz et al., 2015) v3.0.1. R packages *lmtest* (Zeileis and Hothorn, 2002) v0.9-36 and *sandwich* (Zeileis, 2004) v2.4-0 to obtain robust standard errors to address the fact that some variance heterogeneity is present. The effective dose (ED50) value was estimated by module ED in the R package *drc*.

2.3.16 *In vivo* testing of drugs in the xenograft model

Two to three million control and Crohn's patient-derived epithelial cells were harvested by trypsinization, mixed with 50% Matrigel (Becton Dickinson, Palo Alto) to a volume of 100 μl, and injected subcutaneously in NSG (NODscid IL2ranull) mice (Jackson Laboratories, Bar Harbor). After 2 to 4 days of xenograft injected, mice were administered with 80 mg/kg of Tanespimycin (dissolved in corn oil, I.P) five times in 7 days or 25 mg/kg of Gedatolisib (dissolved in 5% D-glucose, 0.3% Lactic acid, pH 3.5, I.P) three times in 7 days. Control mice were administered

with corn oil and DMSO only for the same number of days. Mice were sacrificed after seven days of treatment and xenografts were collected from them for histology and RNA extraction.

CHAPTER 3: CLONING AND CHARACTERIZATION OF PATIENT-DERIVED STEM CELLS

3.1 Significance of the proposed study

Crohn's is considered a multifactorial disease with no definite etiology. Despite advancements in treatment strategies, the patients with stable clinical remission are as low as 10%. Therapeutic regimens aim to maintain remission without the need to go for surgical intervention, but surgery might be required once patients develop complications such as stricture and fistula formation. 50% of the patients diagnosed with Crohn's require surgery within 20 years because of intestinal complications (Baumgart and Sandborn, 2012). And the remaining 50% require surgery within 10 years of the first diagnosis with recurrence still being as high as 55% after 10 years (Peyrin-Biroulet et al., 2010). Genome-wide association studies (GWAS) and meta-analyses studies have revealed 71 risk loci for Crohn's alone (Franke et al., 2010). But genetic contribution explains only 20% of the heritability of Crohn's disease (Ng et al., 2013). Relatively low concordance rates in monozygotic twin studies and evidence of low genetic contribution to heredity suggests large immunobiological and environmental contribution to this disease (Park et al., 2010; Halfvarson, 2011).

GWAS has identified multiple genes as risk loci with restricted expression in intestinal epithelial cells such as related to cell adhesion, tight junction assembly, mucin regulation, and many more (Declan F. McCole, 2014; Franke et al., 2010). This observation is further supported by studies in a large number of IBD patients (Buisine et al., 1999) as well as in several genetic knockout animal models (Park et al., 2009; Williams et al., 2015). Both studies along with GWAS highlight the possibility of the dysfunctional epithelial barrier being a cause for initiation and progression of the disease. Despite a good correlation between the disorder and barrier defects, it is unclear whether these defects are primary events driving the inflammation or

simply secondary consequences of the inflammatory state of this disease. It is also unclear how epithelial barrier dysfunction might explain the alternate regional presentations of Crohn's (Farmer et al., 1975; Cleynen et al., 2016), it's "skip-lesion" patterning, or the high rates of disease recurrence following ileocolonic resection (Baumgart and Sandborn, 2012; Regueiro et al., 2016). While various chemically-induced, genetically- engineered and immune-moderated model systems of intestinal inflammation have been valuable in depicting specific aspects of the disease, no single model has been able to fully capture the complexity of IBD (Pizarro et al., 2003). Hence, our understanding of the cellular and molecular mechanisms underlying this chronic disease is limited. Despite the obvious role of the mucosal barrier in Crohn's, there have been limited studies of patient-specific stem cells due to the lack of a stable cloning system for intestinal mucosal stem cells. The present work aims to address the role of the intestinal barrier in Crohn's disease through clonal analysis of epithelial stem cells derived from endoscopic biopsies of Crohn's and control patients.

3.2 Research strategy

3.2.1 Cloning epithelial stem cells from Crohn's patients' biopsies

Currently, embryonic stem cells (ESCs) and induced pluripotent stem cells (iPSCs) lead stem cell research, therapeutic cloning, and regenerative medicine (Karagiannis et al., 2019). These remarkable cells together constitute pluripotent stem cells that can become any cell type of the body. Despite their tremendous advantages, iPSCs, and ECSs demand taxing and inefficient processes for commitment to desired lineages (Medvedev et al., 2010; Sun et al., 2014), present risk of teratoma formation, and show lack of regenerative potential in derived lineages (Muller and Dzierzak, 1993; Amabile et al., 2013; Suzuki et al., 2013). In contrast, adult or somatic stem cells show indefinite self-renewal and generate all the cell types of the tissue

from which they originate. They work to replace dead cells and maintain homeostasis in an organ. Their unique potential to regenerate tissues may fuel cell-based drug discovery and lead to therapies for diseases that have resisted traditional therapeutic approaches. Green and colleagues established a system to grow epidermal stem cells *in vitro* in an immature state (Rheinwald and Green, 1975). The technology was dependent on irradiated 3T3-J2 fibroblast cells used as feeder cells to serially cultivate and differentiate epidermal stem cells into a stratified squamous epithelium. Adding on that, stem cells from other stratified epithelia, such as thymus, cornea, and lung were subsequently developed (Senoo et al., 2007; Rama et al., 2010; Kumar et al., 2011). These *in vitro* generated stratified epithelia have been used for autologous stem cell transplantation in patients with skin burns, damaged cornea resulting from accidents, and genetic conditions (Green, 2008; Rama et al., 2010; Hirsch et al., 2017). However, columnar epithelial stem cells, such as those from the intestine, liver, and pancreas, resist cloning with the technology pioneered by Green and colleagues.

Later, Wang and colleagues established a platform to isolate and clone epithelial stem cells from the fetal gastrointestinal tract by formulating a media with components supporting stem cell renewal (Wang et al., 2015). These stem cells were successfully propagated and differentiated into cell types of intestinal epithelia in an air-liquid interface (ALI) cultures. Duleba and colleagues employed the same platform to clone epithelial stem cells from a typical 1 mm-diameter endoscopic biopsies from the human gastrointestinal tract (Duleba et al., 2019; Duleba et al., 2020). The ground-state intestinal stem cells (gISCs) were cloned in an undifferentiated manner and showed unlimited proliferative potential with high clonogenicity. The stem cells despite extensive *in vitro* passaging maintained their genomic integrity, and fate commitment upon differentiation. Such technology opens doors to the potential of somatic stem cells in disease modeling, therapeutic cloning, and regenerative medicine.

3.2.2 Expanding single cell-derived clones

From a typical 1 mm-diameter endoscopic biopsy that harbors 300,000-900,000 epithelial cells, 100-200 epithelial stem cell colonies appear 7-10 days post-seeding in a monolayer of irradiated 3T3-J2 fibroblast cells (Duleba et al., 2020). The gISCs show above 70% clonogenicity with the appearance of compact colonies from seeding single cells in each well of a 384 well plate previously seeded with irradiated 3T3-J2 fibroblast cells. These single cell-derived stem cell colonies appear 7-14 days post-seeding and propagate to greater than a billion cells within 8 weeks. Stem cell libraries, representative of the biopsy, as well as single cell-derived clones from those libraries, representative of a fraction of the biopsy, can be expanded for downstream applications.

3.2.3 Functional characterization of intestinal stem cells

The gISCs also show multipotency upon an air-liquid interface (ALI) differentiation *in vitro* (Duleba et al., 2020). Both stem cell libraries and single cell-derived stem cells differentiate into polarized columnar epithelial marked by the presence of *MUC2* positive goblet cells, *DEFA6* positive Paneth cells, *CHGA* positive endocrine cells, and *VILLIN* positive enterocytes. These four cell types constitute the major cell types of intestinal epithelia. The ALI cultures from gISCs of different regions of the gastrointestinal tract yield a three-dimensional epithelium whose cross-sectional histology is similar to the *in situ* of the analogous regions (Wang et al., 2015; Duleba et al., 2020).

3.2.4 Stability of intestinal stem cells

The single cell-derived stem cells show extraordinary genomic stability despite long-term culture *in vitro* (Wang et al., 2015; Duleba et al., 2020). The somatic allelic copy number variation (CNV) using the best practices pipeline of GATK v.4.0.4 did not show any

chromosomal aneuploidies during the long-term passaging of the clones at passage 5 (50 days in culture), passage 10 (100 days in culture), and passage 20 (200 days in culture). The long-term stability together with an unlimited proliferative potential of a single cell-derived clone allows to study and untangle heterogeneity within a patient biopsy.

3.3 Results

3.3.1 Stem cell heterogeneity in Crohn's

Using the technology described in methods (Wang et al., 2015; Duleba et al., 2019; Duleba et al., 2020), we processed 1 millimeter (mm) endoscopic biopsies taken from terminal ileum of 38 Crohn's patients and 11 control patients (Table 3.1). From each patient, 100-300 epithelial stem cell colonies which hereby are referred to as "libraries" were derived. These libraries were propagated in sterile culture conditions for at least a week before analysis. Clonal lines were generated from single cells of these libraries by expanding colony formed from seeding single cell in each well of a 384-well plate (Fig. 3.1A). The clonal lines were grown separately from the libraries they were derived from and were treated as individual lines.

Single-cell RNA sequencing (scRNAseq; Satija et al., 2015) of one Crohn's patient (SPN-29) and one control patient (SPN-19) stem cell library revealed that Crohn's patient harbor multiple clusters of cells in contrast to control. t-SNE (t-Distributed Stochastic Neighbor Embedding; van der Maaten and Hinton, 2008) analysis of the whole genome expression from ten thousand stem cells showed that libraries from the control case are dominated by a single cluster of cells (Cluster 1, or CLST1) while that of Crohn's has a majority of two additional clusters (Cluster 2, or CLST2 and Cluster 3, or CLST3) on top of CLST1 that dominates the control case (Fig. 3.1B). CLST1 expressed normal intestinal epithelial markers such as

Table 3.1. Crohn's and control subjects

Study Pat.	Age	Sex	Diagnosis	Terminal Ileum Report	Current Therapy	Remission	Hospital
SPN-16	10	F	Two juvenile polyps, no adenomatous changes	Normal Ileum	NA	NA	UCONN
SPN-19	16	F	Normal	Normal Ileum	NA	NA	UCONN
SPN-48	9	F	Normal	Normal Ileum	NA	NA	UNC
SPN-04	39	M	Normal	Normal Ileum	NA	NA	BCM
SPN-05	35	M	Normal	Normal Ileum	NA	NA	BCM
SPN-06	66	F	Normal	Normal Ileum	NA	NA	BCM
SPN-07	58	F	Normal	Normal Ileum	NA	NA	BCM
SPN-11	63	F	Normal	Normal Ileum	NA	NA	BCM
SPN-12	67	M	Normal	Normal Ileum	NA	NA	BCM
SPN-13	55	F	Normal	Normal Ileum	NA	NA	BCM
SPN-14	67	M	Normal	Normal Ileum	NA	NA	BCM
SPN-01	41	M	Crohn's	Normal Ileum	5-ASA, Azathioprine	Yes	BCM
SPN-02	41	F	Crohn's	Stricturing Ileum	Vedolizumab	Yes	BCM
SPN-03	43	M	Crohn's	Stricturing Ileum	Budesonide, 6-MP	Unknown	BCM
SPN-08	38	F	Crohn's	Macroscopic disease in Ileum	Azathioprine, Infliximab	Yes	BCM
SPN-09	30	M	Crohn's	Macroscopic disease in Ileum	Budesonide, Adalimumab	No	BCM
SPN-10	67	M	Crohn's	NSNF in Ileum	Infliximab	No	BCM
SPN-15	14	F	Crohn's	Normal Ileum	None	Yes	UCONN
SPN-17	22	M	Crohn's	Mild active inflammation	Methotrexate	Yes	UCONN
SPN-18	19	M	Crohn's	N/A	N/A	No	UCONN
SPN-20	12	M	Crohn's	N/A	N/A	No	UCONN
SPN-21	20	M	Crohn's	Macroscopic disease in TI, had previous resection TI/Caecum	Humira, Methotrexate	Yes	UCONN
SPN-22	21	N/A	Crohn's	Normal Ileum	Infliximab	Yes	UCONN
SPN-23	21	N/A	Crohn's	Normal Ileum	Infliximab	Yes	UCONN
SPN-24	17	N/A	Crohn's	Normal Ileum, had resection of TI/Caecum	Infliximab	Yes	UCONN
SPN-25	12	N/A	Crohn's	Normal Ileum	Methotrexate	Yes	UCONN

Table 3.1. (cont.) Crohn's and control subjects

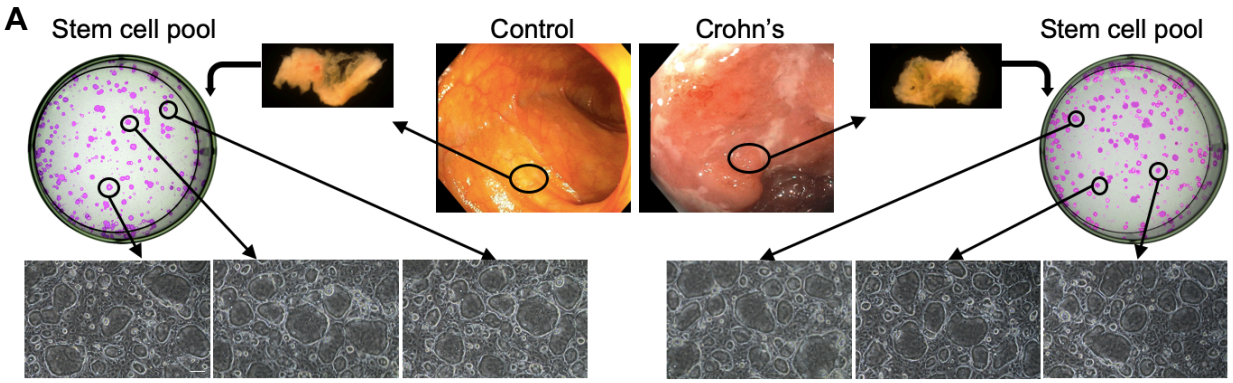
SPN-26	11	N/A	Crohn's	Normal Ileum, Macroscopic disease in TC	Infliximab, Methotrexate	No	UCONN
SPN-27	11	N/A	Crohn's	Normal Ileum	Infliximab, Methotrexate	Yes	UCONN
SPN-28	12	N/A	Crohn's	Normal Ileum	Methotrexate	Yes	UCONN
SPN-29	14	N/A	Crohn's	Normal Ileum	Methotrexate	Yes	UCONN
SPN-30	12	N/A	Crohn's	Microscopic disease in Ileum	Methotrexate	Yes	UCONN
SPN-31	20	N/A	Crohn's	Normal Ileum	Infliximab	Yes	UCONN
SPN-32	17	N/A	Crohn's	Normal Ileum	Infliximab, Methotrexate	Yes	UCONN
SPN-33	14	N/A	Crohn's	Mild increase in inflammatory cells in Ileum, single multinucleated giant cell in pathology	Infliximab, Methotrexate	Yes	UCONN
SPN-34	17	N/A	Crohn's	Mild ileitis, superficial ulceration at ileocolonic junction	Infliximab	Yes	UCONN
SPN-35	21	N/A	Crohn's	Normal Ileum	Humira	Yes	UCONN
SPN-36	15	N/A	Crohn's	Microscopic disease in Ileum	Humira	Yes	UCONN
SPN-37	14	N/A	Crohn's	Normal Ileum	Infliximab, Methotrexate	Yes	UCONN
SPN-38	10	N/A	Crohn's	Normal Ileum, Microscopic disease in Des Colon	Infliximab, Methotrexate	No	UCONN
SPN-39	20	N/A	Crohn's	Normal Ileum	Pentasa	Yes	UCONN
SPN-40	17	N/A	Crohn's	Normal Ileum	Infliximab	Yes	UCONN
SPN-41	19	N/A	Crohn's	Microscopic disease in Ileum	Stelara	No	UCONN
SPN-42	17	N/A	Crohn's	Normal Ileum	Humira	Yes	UCONN
SPN-43	21	N/A	Crohn's	Normal Ileum	N/A	Yes	UCONN
SPN-44	22	N/A	Crohn's	Normal Ileum	Remicade	Yes	UCONN
SPN-45	18	N/A	Crohn's	Macroscopic disease in Ileum	Remicade	Yes	UCONN
SPN-46	12	N/A	Crohn's	Normal Ileum	Infliximab	Yes	UCONN
SPN-47	13	N/A	Crohn's	Normal Ileum	Methotrexate	Yes	UCONN
SPN-49	N/A	N/A	Crohn's	Microscopic disease in Ileum	None	No	UNC

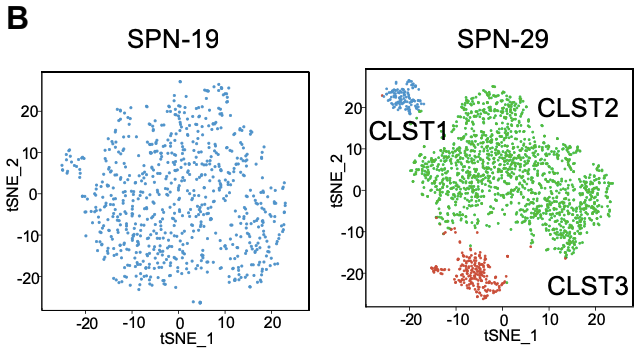
*N/A – Not Available, NA – Not Applicable
*UCONN – University of Connecticut, BCM – Baylor College of Medicine,
UNC – University of North Carolina

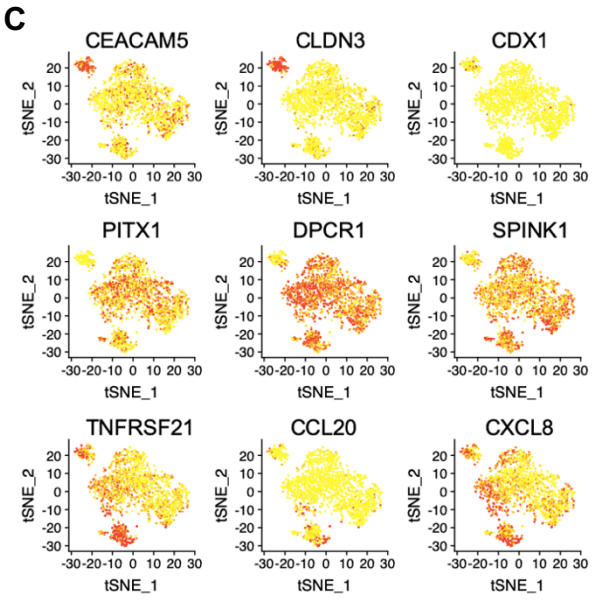
CEACAM5, CLDN3, CDX1, and HOXB7 while CLST2 and CLST3 combinedly expressed VSIG1, DPCR1, SPINK1, and CLDN18. Further CLST2 was distinguished from CLST3 by expression of PITX1 and CLST3 from CLST2 by expression of TNFRSF21, CLL20, and CXCL8 expression (Fig. 3.1C). Upon integration of two single-cell RNA sequencing profiles from control and Crohn's cases, CLST1 from Crohn's merged well with the population of cells that dominated the control (Fig. 3.1D). The same normal mucosal marker (CDX1) was evident on the cells from control and the same markers (PITX1, TNFRSF21) were evident on the two clusters from Crohn's (Fig. 3.1E).

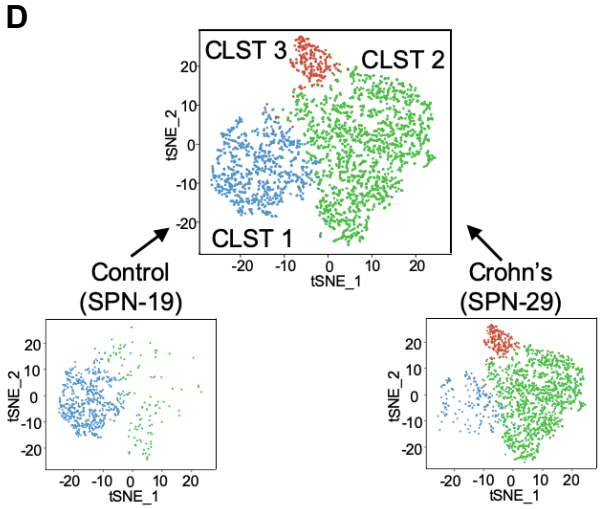
To further examine the heterogeneity observed in terminal ileum stem cells of one Crohn's patient, we extended our study to 11 control and Crohn's patient libraries. We randomly sampled 36 single cell-derived clones grown in a 384 well plate from these cases and subjected them to whole-genome microarray transcriptomic analysis. Principal Component Analysis (PCA) of whole-genome expression profiles of these clones showed separation into three clusters. The clones from control cases fell into a single cluster CLST1 while that of Crohn's patients fell into at least two clusters (Fig. 3.1F). More than 800 differentially expressed genes (<1.5-fold, $p<0.05$) were evident among these clusters in a heat map analysis (Fig. 3.1G). Consistent with the single-cell RNA sequencing markers for these clusters, the intensity values from the expression data showed that the CLST1 clones had high expression of normal intestinal mucosal markers such as HOXA11, HOXA13, CDX1, GPA33, NOX1, and CEACAM5. In contrast to CLST1 clones, the two variant clones combinedly showed high expression of EMB, CLDN18, VSIG1, SPINK1, DPCR1, and PRSS1. CLST2 was distinguished from CLST3 by expression of CD74 and LCN2, and CLST3 clones by high expression of PSCA, TFF2, and MUC5AC (Fig. 3.1H).

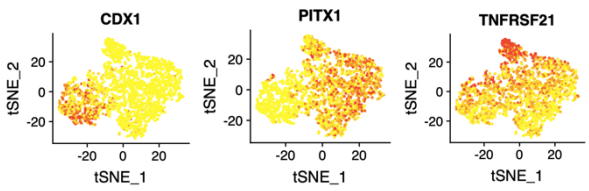
Fig. 3.1. Clonal analysis of terminal ileum stem cells reveals heterogeneity in Crohn's **A**. Schematic workflow of generating "libraries" of single cell-derived colonies and subsequently of clonal stem cell lines from 1mm endoscopic biopsies of control and Crohn's terminal ileum. White light imaging of control and Crohn's terminal ileum, the site of and resulting biopsy, as well as Rhodamine red-stained plates bearing the 100-300 colonies derived from a typical biopsy. Individual colonies are sampled from the pool and grown as separate lines. Scale bar, 100 μm. **B.** Unmerged tSNE profiles of scRNA-seq data of cells from control (SPN-19) and Crohn's (SPN-29) libraries. **C.** Mapping of cluster-specific expression marker genes onto tSNE profile of Crohn's library from case SNP-29. CEACAM5, CLDN3, and CDX1 mark CLST1; PITX1 marks CLST2; DPCR1 and SPINK1 mark CLST2 and 3, and CLST3 shows the highest expression of TNFRSF21, CCL20, and CXCL8. **D**. Aggregate tSNE profiles of single-cell RNAseq data of control (SPN-19) and Crohn's (SPN-29) stem cell libraries to reveal three primary clusters. **E**. Mapping of cluster-specific markers onto an integrated tSNE profile assembled from control (SPN-19) and CD (SPN-29) data. **F.** Principal component analysis of whole-genome expression data of sampled clones corresponding to *D*. **G.** Heatmap of differentially expressed (1.5x, $p<0.05$) genes from whole-genome data of CLST1, CLST2, and CLST3 clones. **H.** Heatmap of selected, differentially expressed genes from whole-genome expression data of sampled clones from multiple patients.


A
Stem cell pool
Control
Crohn's
Stem cell pool


B
SPN-19
SPN-29
CLST1
CLST2
CLST3
tSNE_1
tSNE_2


C
CEACAM5
CLDN3
CDX1
PITX1
DPCR1
SPINK1
TNFRSF21
CCL20
CXCL8
tSNE_1
tSNE_2


D
CLST 3
CLST 2
CLST 1
Control
(SPN-19)
Crohn's
(SPN-29)
tSNE_1
tSNE_2


E
CDX1
PITX1
TNFRSF21
tSNE_1
tSNE_2

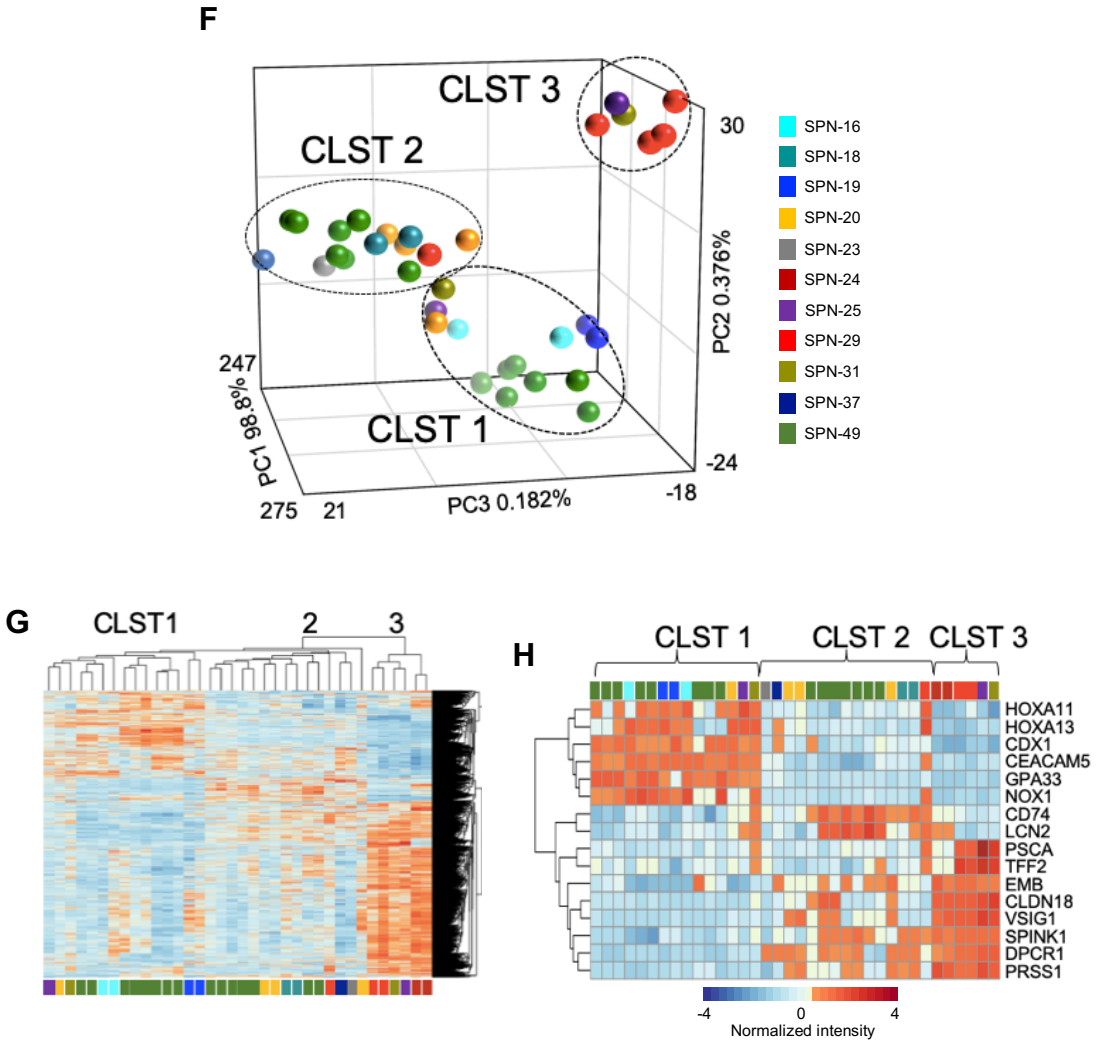
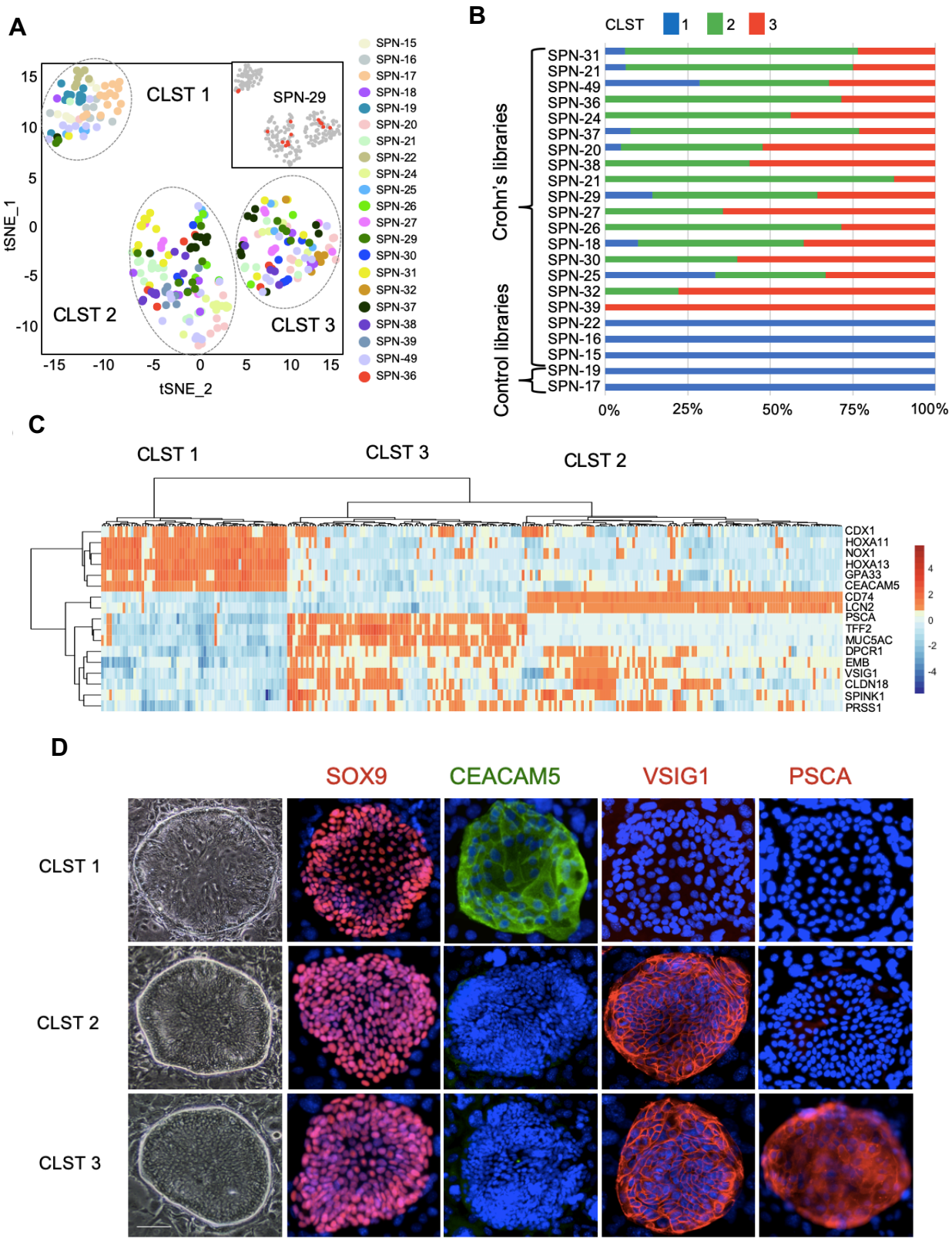

F
CLST 3
CLST 2
CLST 1
30
-24
PC2 0.376%
247
PC1 98.8%
275
21
PC3 0.182%
-18
SPN-16
SPN-18
SPN-19
SPN-20
SPN-23
SPN-24
SPN-25
SPN-29
SPN-31
SPN-37
SPN-49
G
CLST1
2
3
H
CLST 1
CLST 2
CLST 3
HOXA11
HOXA13
CDX1
CEACAM5
GPA33
NOX1
CD74
LCN2
PSCA
TFF2
EMB
CLDN18
VSIG1
SPINK1
DPCR1
PRSS1
-4
0
4
Normalized intensity

Fig. 3.1. (cont.)

Using these 17 markers to distinguish between CLST 1, 2, and 3, the gene expression study was further expanded to an additional 7 to 28 clones from each stem cell library (a total of 275 clones) from 21 control and Crohn's patients. The single cell-derived clones were assessed by RT-qPCR for the expression of these markers. These clones also separated into three clusters when expression data was visualized by tSNE analysis (Fig. 3.2A). The inset shows that Crohn's patient SPN-29 studied for sc-RNA seq yielded clones that distributed into three clusters. Similarly, seven other patients were distributed into three clusters, eight patients into two clusters, and the remaining in a single cluster. Clones from two control patients were only limited to CLST1 as expected whereas clones from 16 out of 19 Crohn's patients were mostly dominated by CLST2 and CLST3 (Fig. 3.2B). Three Crohn's patients only showed clones only in CLST1 suggesting they might be like controls. The separation of these clones into different clusters was because of differences in gene expression which is shown by heat map analysis (Fig. 3.2C). The three representative clusters showed consistent expression of 17 genes as seen in whole-genome expression data (Fig. 3.2C, Fig. 3.1H).

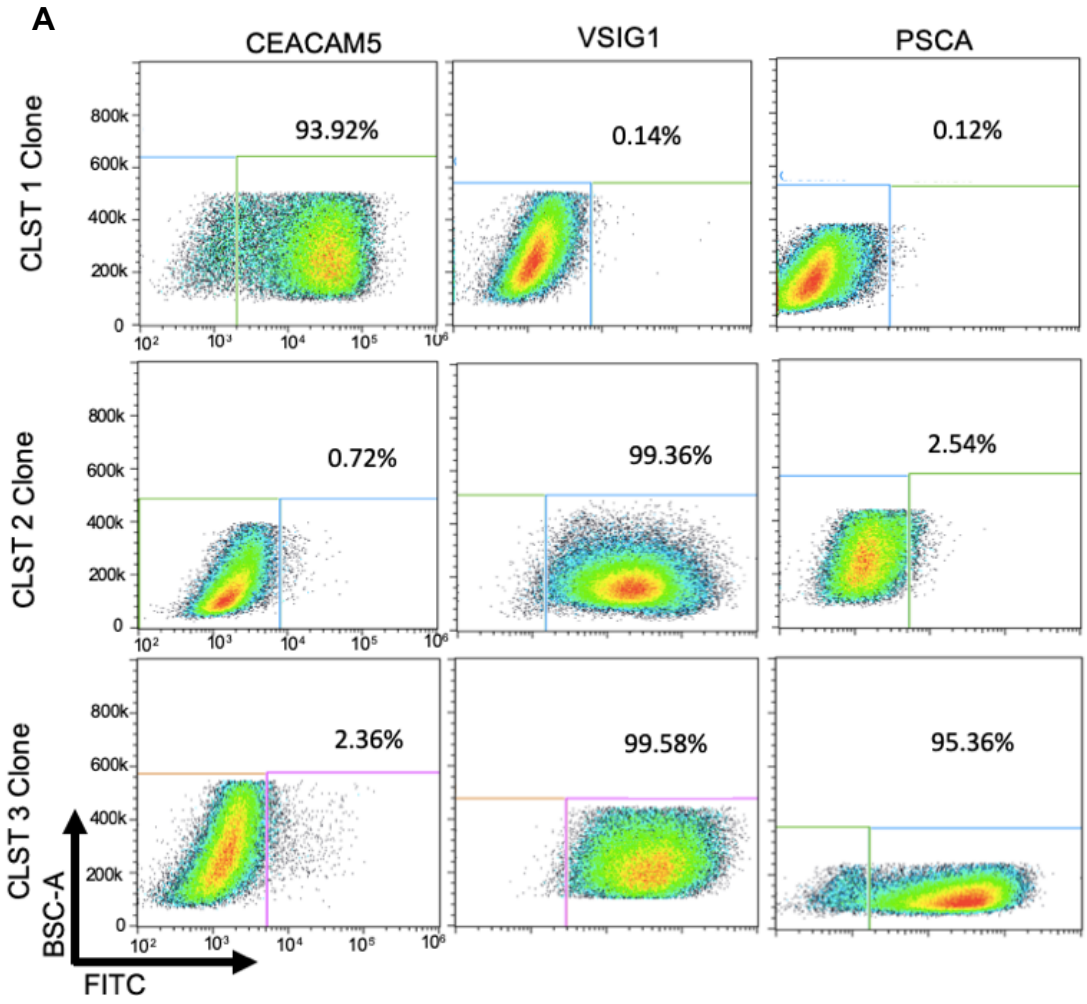
Stem cell colonies of representative CLST 1, 2, and 3 from Crohn's patient SPN-29 showed expression of SOX9, a marker of intestinal epithelial stem cells (Roche et al., 2015; Wang et al., 2015) in immunofluorescence using antibodies. Similar to gene expression via transcriptomic analysis, CLST1 clones stained positive for CEACAM5, both variants stained positive for VSIG1 and CLST3 alone stained positive for PSCA (Fig. 3.2D). Upon seeing consistent and stable protein expression of cell junctional markers CEACAM5 (for CLST 1), VSIG1 (for CLST 2 + 3), and PSCA (for CLST 3), we tested whether these can be used in fluorescence-activated cell sorting (FACS) for quantification. Patient-matched clones from SPN-29 stained positively (>90%) for CEACAM5, VSIG1, and PSCA for their respective clusters and negatively in other clusters (Fig. 3.3A). Similarly, clones from five more Crohn's patients and

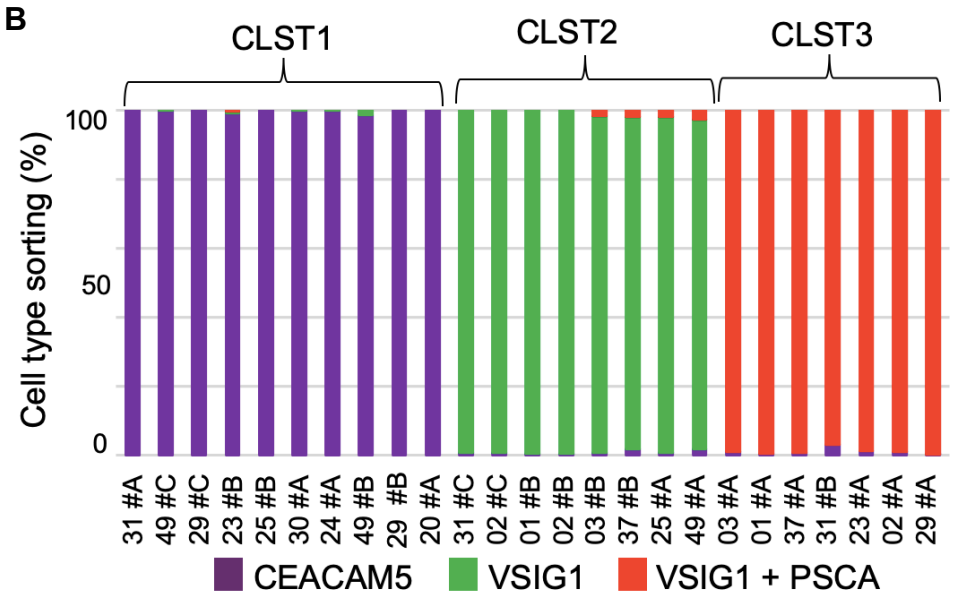
Fig. 3.2. Stem cell heterogeneity by qPCR and immunofluorescence staining
A. tSNE analysis of the 275 clones based on the expression data of 16 markers. Inset highlights the distribution of clones sampled from the libraries of a single Crohn's patient (SPN-29) relative to the overall clusters identified. **B.** Histogram depicting the distribution of clone types resulting from a random sampling of 7-28 clones from control and Crohn's stem cell libraries determined by marker analysis. **C.** Expression heatmap of 16 marker genes assessed by quantitative RT-PCR across 275 clones sampled from 2 control and 19 Crohn's stem cell libraries. **D.** Phase contrast and immunofluorescence imaging of colonies from cloned CLST1, 2, and 3 lines from a single Crohn's library using antibodies to SOX9, CEACAM5, VSIG1, and PSCA. Scale bar, 100 μm.


A
CLST 1
SPN-29
CLST 2
CLST 3
tSNE_1
tSNE_2
SPN-15
SPN-16
SPN-17
SPN-18
SPN-19
SPN-20
SPN-21
SPN-22
SPN-24
SPN-25
SPN-26
SPN-27
SPN-29
SPN-30
SPN-31
SPN-32
SPN-37
SPN-38
SPN-39
SPN-49
SPN-36
B
CLST 1 2 3
Crohn's libraries
Control libraries
SPN-31
SPN-21
SPN-49
SPN-36
SPN-24
SPN-37
SPN-20
SPN-38
SPN-21
SPN-29
SPN-27
SPN-26
SPN-18
SPN-30
SPN-25
SPN-32
SPN-39
SPN-22
SPN-16
SPN-15
SPN-19
SPN-17
0%
25%
50%
75%
100%
C
CLST 1
CLST 3
CLST 2
CDX1
HOXA11
NOX1
HOXA13
GPA33
CEACAM5
CD74
LCN2
PSCA
TFF2
MUC5AC
DPCR1
EMB
VSIG1
CLDN18
SPINK1
PRSS1
D
SOX9
CEACAM5
VSIG1
PSCA
CLST 1
CLST 2
CLST 3

three control patients were also stained for those markers to check consistency across clones from multiple patients (Fig. 3.3B). Then, we took advantage of these markers to quantify relative proportions of CLST 1, 2, and 3 across 11 control and 39 Crohn's patient stem cell libraries using FACS (Fig. 3.3C). The variant clones (CLST 2 and 3) represented approximately 64.6+/-32.6% of all clones in 32 out of 39 Crohn's cases, compared to 1.2+/-1.4% in 11 control cases suggesting the existence of these variants is a distinguishable feature in Crohn's libraries ($P=2.5\ e^{-14}$) (Fig. 3.3D, Fig. 3.3E). Seven of Crohn's cases stained positively only to CEACAM5 (<95%), which is a CLST1 marker suggesting that these cases might be similar to controls. Each of the clones derived from these patient libraries could be propagated continuously for a minimum of one year during which they maintained a clonogenicity of greater than 50%, a feature of mucosal stem cells (Wang et al., 2015; Duleba et al., 2020; Fig. 3.3F).

Fig. 3.3. Stem cell library heterogeneity by FACS and clonogenicity of stem cells
A. Fluorescence-activated cell sorting (FACS) profiles of individual examples of CLST1, CLST2, and CLST3 clones stained by antibodies to the CLST1-specific marker CEACAM5, the CLST2+CLST3 marker VSIG1, and the CLST3-specific marker PSCA. **B.** Histogram of FACS data from CLST 1, 2, and 3 clones across patients with red (PSCA+) superimposed on the green (VSIG1; CLST2+CLST3) signal. **C.** FACS profiling of control and Crohn's stem cell libraries using antibodies to CEACAM5, VSIG1, and PSCA. **D.** Histogram of FACS data from 11 control libraries and 38 Crohn's libraries with red (PSCA+) superimposed on the green (VSIG1; CLST2+CLST3) signal. **E.** Box plot of FACS profiles of 11 control and 38 Crohn's libraries comparing the proportion of variant clones (CLST2+CLST3) to normal terminal ileum stem cells (CLST1). Medians, Q1 and Q3, and p values are indicated. **F.** Clonogenicity of CLST1, CLST2, and CLST3 stem cell variants at passage 5 and passage 25 (P5 and P25).


A
CEACAM5
VSIG1
PSCA
CLST 1 Clone
CLST 2 Clone
CLST 3 Clone
93.92%
0.14%
0.12%
0.72%
99.36%
2.54%
2.36%
99.58%
95.36%
800k
600k
400k
200k
0
BSC-A
FITC


B
CLST1
CLST2
CLST3
Cell type sorting (%)
100
50
0
31 #A
49 #C
29 #C
23 #B
25 #B
30 #A
24 #A
49 #B
29 #B
20 #A
31 #C
02 #C
01 #B
02 #B
03 #B
37 #B
25 #A
49 #A
03 #A
01 #A
37 #A
31 #B
23 #A
02 #A
29 #A
CEACAM5
VSIG1
VSIG1 + PSCA

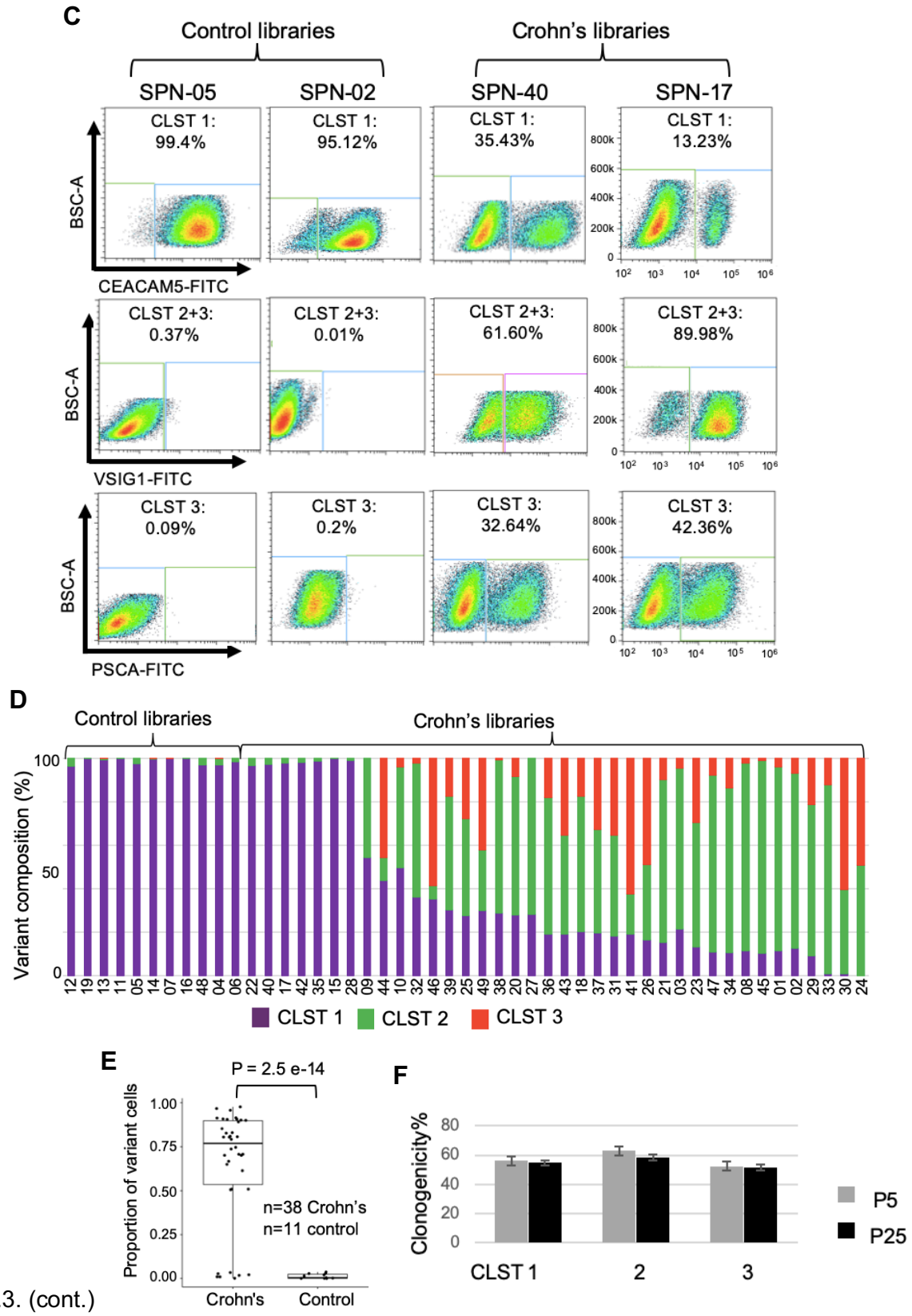
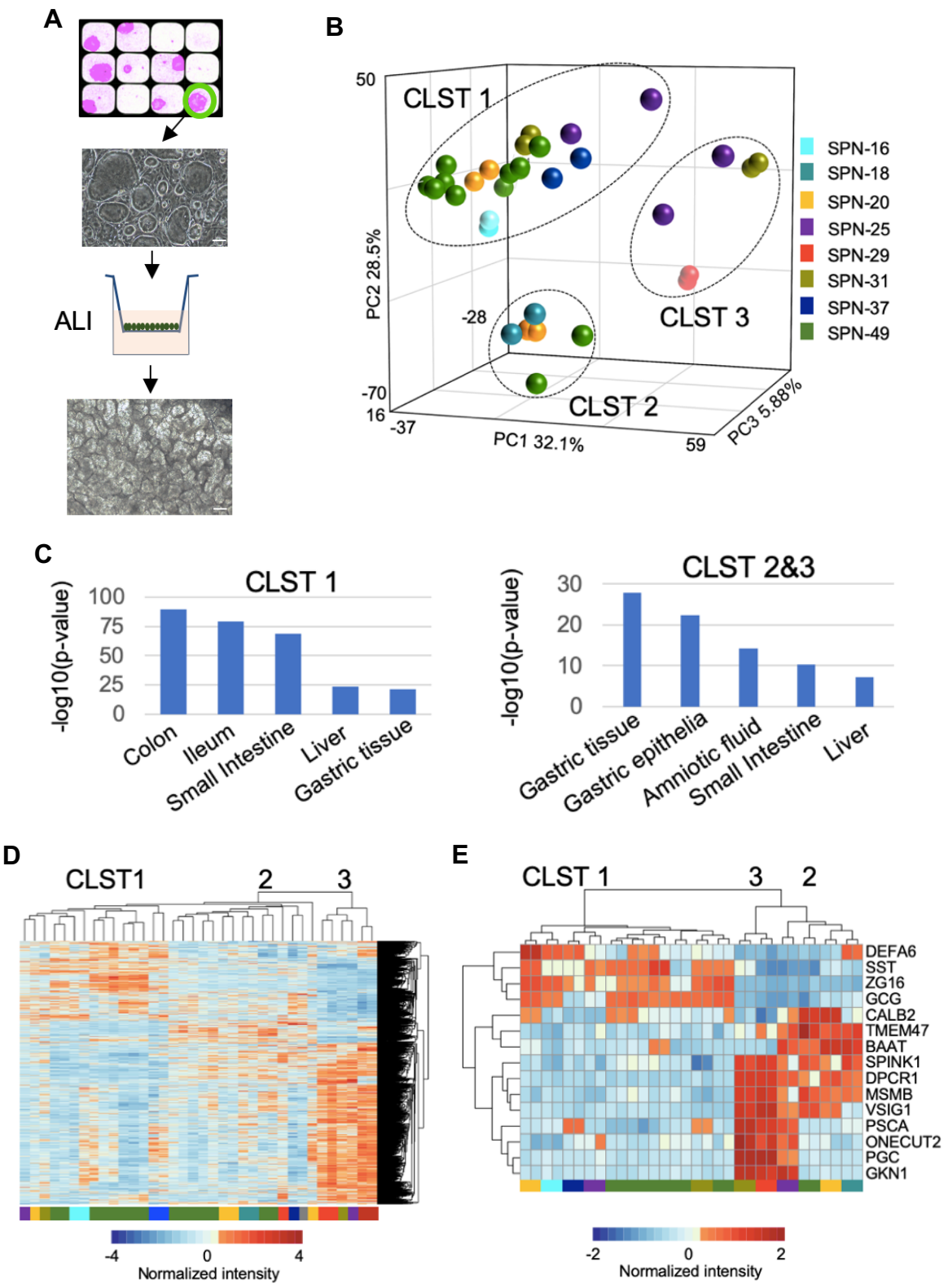

C
Control libraries
Crohn's libraries
SPN-05
SPN-02
SPN-40
SPN-17
CLST 1: 99.4%
CLST 1: 95.12%
CLST 1: 35.43%
CLST 1: 13.23%
BSC-A
CEACAM5-FITC
CLST 2+3: 0.37%
CLST 2+3: 0.01%
CLST 2+3: 61.60%
CLST 2+3: 89.98%
VSIG1-FITC
CLST 3: 0.09%
CLST 3: 0.2%
CLST 3: 32.64%
CLST 3: 42.36%
PSCA-FITC
D
Control libraries
Crohn's libraries
Variant composition (%)
CLST 1
CLST 2
CLST 3
E
P = 2.5 e-14
Proportion of variant cells
n=38 Crohn's
n=11 control
Crohn's
Control
F
Clonogenicity%
CLST 1
2
3
P5
P25

Fig 3.3. (cont.)

3.3.2 Crohn's variant stem cells committed to upper GI fate

Another feature of mucosal stem cells is their ability to differentiate into multiple cell lineages. In order to examine their fate upon differentiation, we made use of the air-liquid interface (ALI) differentiation system (Wang et al., 2015; Duleba et al., 2020) (Fig. 3.4A). Principal Component Analysis (PCA) (>1.5 fold, $p<0.05$) of whole-genome microarray expression data on ALI-differentiated CLST 1, 2, and 3 variants from seven Crohn's patient and CLST1 from single control patient showed that the respective clones still occupied three distinct space suggesting that they are committed to different epithelia (Fig. 3.4B). Further, this observation in PCA was supported by tissue enrichment analysis on expression data from these differentiated clones. Upregulated genes (>1.5-fold, $p<0.05$) from each cluster were compared across publicly available Human Protein Atlas (HPA) dataset (Uhlén et al., 2015) for all the epithelial tissues in gastrointestinal and other regions. Consistent with results from sc-RNA seq, expression data, and immunofluorescence staining on CLST1 stem cells from both control and Crohn's patients, tissue enrichment showed that CLST1 epithelia were closest to the colon, ileum, and small intestine. But CLST2 and CLST3 epithelia were closest to gastric tissue and gastric epithelia (Fig. 3.4C). The separation in PCA and differences in epithelia commitment was because of nearly 1500 differentially expressed genes (CLST1 vs CLST2 and CLST3) and 500 differentially expressed genes (CLST2 vs CLST3) among the variants (Fig. 3.4D). Heatmap analysis showed markers of the normal small intestine and colonic epithelia in ALI-differentiated CLST1 clones (GCG, ZG16, SST, DEFA6, and TMEM47) relatively high compared to the other two clusters, whereas differentiated CLST2 and CLST3 clones showed higher expression of gastric epithelial markers (BAAT, CALB2, and MSMB, VSIG1, DPCR1, SPINK1, ONECUT2, PSCA, PGC, andGKN1). CLST3 showed higher expression of markers such as PSCA, PGC, and GKN1 compared to CLST2 (Fig. 3.4E).

Fig. 3.4. CLST 2 and CLST3 clones are committed to upper gastrointestinal tract fates
A. Schematic for *in vitro* differentiation of clones starting from libraries seeded to 384-well plates, sampling, expansion, and exposure to air-liquid interface in *Transwell* plates. **B.** Principal component analysis from whole-genome expression data (1.5x, $p<0.05$) of ALI differentiated CLST1, CLST2, and CLST3 clones from indicated patient libraries. **C.** *Top*, histogram depicting the top 5 enriched tissues ($p<0.01$) determined by ARCHS4 Tissues of differentially expressed genes in ALI-differentiated CLST1 clones. *Bottom*, histogram depicting the five most enriched tissues ($p<0.01$) determined by ARCHS4 Tissues of differentially expressed genes in ALI-differentiated CLST2 and 3 clones. **D.** Heatmap of differentially expressed (1.5x, $p<0.05$) genes from whole-genome data of CLST1, CLST2, and CLST3 ALI-differentiated clones. **E.** Expression heatmap of selected marker genes from whole-genome data of ALI differentiated clones, including those of goblet cells (ZG16), Paneth cells (DEFA6), endocrine cells (SST, GCG), and junctional markers (VSIG1).


A
ALI
B
CLST 1
CLST 2
CLST 3
PC1 32.1%
PC2 28.5%
PC3 5.88%
SPN-16
SPN-18
SPN-20
SPN-25
SPN-29
SPN-31
SPN-37
SPN-49
C
CLST 1
-log10(p-value)
Colon
Ileum
Small Intestine
Liver
Gastric tissue
CLST 2&3
Gastric tissue
Gastric epithelia
Amniotic fluid
Small Intestine
Liver
D
CLST1
Normalized intensity
E
CLST 1
DEFA6
SST
ZG16
GCG
CALB2
TMEM47
BAAT
SPINK1
DPCR1
MSMB
VSIG1
PSCA
ONECUT2
PGC
GKN1
Normalized intensity

Cross-sections of ALI-differentiated epithelia of each clone respective of three clusters from five Crohn's patients including SPN-29 and respective of a single cluster from three controls including SPN-19 stained to various small intestine, colonic and gastric markers in immunofluorescence confirmed our observation from tissue enrichment and heatmap analysis (Fig. 3.5). CLST1 clones from control and Crohn's patients differentiated into normal intestinal epithelia that were marked by MUC2 expressing goblet cells, CHGA expressing enteroendocrine cells, and DEFA6 expressing Paneth cells. CLST1 differentiated epithelia were also marked by the continuous expression of tight junction marker GPA33 and cell adhesion marker CEACAM5 typical of small intestine and colon. In contrast, CLST2 and CLST3 were devoid of goblet cells, enteroendocrine cells, or Paneth cells, and also showed discontinuous expression or absence of junctional marker GPA33 and cell adhesion marker CEACAM5. Instead, CLST2 and CLST3 combinedly expressed gastric junctional adhesion markers VSIG1 and CLDN18, and gastric pseudo-goblet cells marker MUC5AC while CLST1 epithelia stained negative for these. Expression of MUC5AC expressing pseudo-goblet cells were more evident in CLST3 compared to CLST2. Further, only CLST2 stained positively for LCN2, a transporter protein, and only CLST3 stained positive for PSCA, a gastric cell membrane glycoprotein (Fig. 3.5).

In order to assess if the *in vitro* stem cell phenotype and ALI-differentiated fate of these clusters were influenced by the culture system, we employed a humanized xenograft mouse model to inject CLST 1, 2, and 3 stem cells subcutaneously to immunodeficient (*NODscid IL2ra*^{null} [NSG]; Shultz at al., 2005; Ito et al., 2012; Shultz et al., 2012) mice (Fig. 3.6A). Cross-sectional staining of harvested xenografts to junction specific marker E-cadherin in immunofluorescence showed that the epithelial stem cells from each cluster differentiated into polarized columnar epithelia. The E-cadherin-positive epithelia were also stained positive for

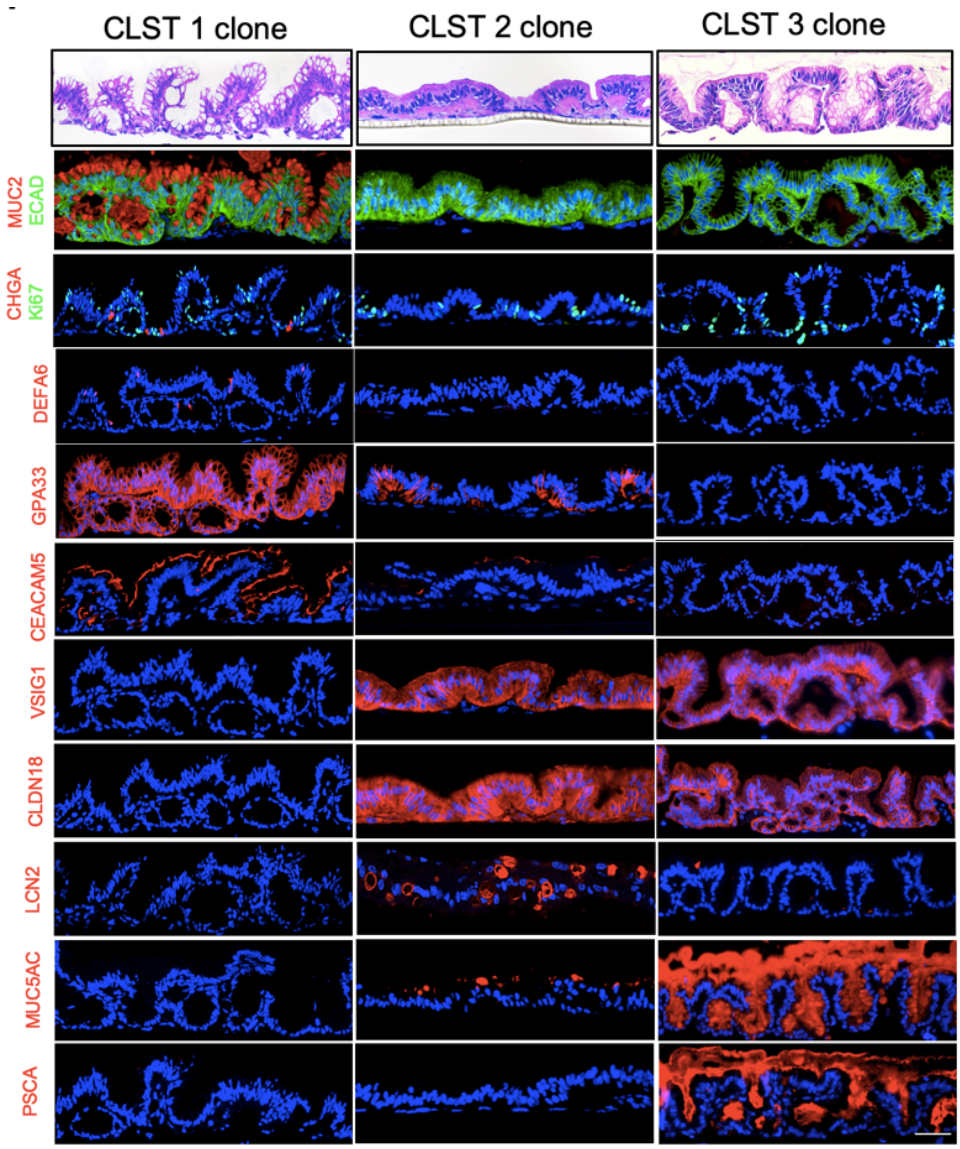


Fig. 3.5. Clonal differentiation in an air-liquid interface (ALI)
ALI-differentiated CLST1 cells form epithelia dominated by MUC2+ goblet cells, CHGA+ endocrine cells, DEFA6+ Paneth cells, and CEACAM5+ epithelium. CLST2 cells form a VSIG1+, CLDN18+, and LCN2+ epithelia lacking goblet cells. CLST3 cells form epithelia lacking goblet cells but express high levels of MUC5AC, VSIG1, and PSCA. Additional staining by KI67 (proliferation marker) shows cells proliferating in all three clusters. Scale bar, 100 μm.

human-specific marker STEM121 (Kelly et al., 2004) suggesting they were unlikely to be of host origin. Immunofluorescence staining of cross-sections of three variants from five Crohn's patient including SPN-29 and single variant from three control patients including SPN-19 showed consistent results to the ALI-differentiation of these clones. CLST1 differentiated with the presence of MUC2, continuous GPA33, and apical CEACAM5 expression while CLST2 and CLST3 were devoid of MUC2 and had discontinuous expression or absence of GPA33 and CEACAM5. In addition, CLST2 and CLST3 expressed gastric markers VSIG1, CLDN18, and MUC5AC combinedly. Gastric pseudo-goblet cells expressing MUC5AC were still more evident in CLST3 compared to CLST2. Only CLST2 was marked by LCN2 expression while CLST3 had no expression (Fig. 3.6A).

In addition, we showed the stability of fate commitment of three variants from Crohn's patient SPN-29 and one variant from control patient SPN-19 despite *in vitro* passaging for 20 passages over 200 days. The differentiation of clones was analyzed by ALI *in vitro* with markers such as MUC2, GPA33, CHGA, DEFA6, MUC5AC, VSIG1, and CLDN18 as well as a xenograft *in vivo* with markers MUC2 and MUC5AC (Fig. 3.6B). We saw no change in marker expression in differentiated clones even after extensive propagation *in vitro*.

Finally, we examined histological sections of resected terminal ileum of six Crohn's patients to validate our *in vitro* cloning and differentiation observations. The antibodies specific to marker genes of CLST1 (CEACAM5, GPA33, and MUC2), CLST2 (VSIG1 and LCN2), and CLST3 (VSIG1 and MUC5AC) were used in immunofluorescence to assess this. The histological cross-sections from all six patients showed regions of normal epithelia (GPA33+, MUC2+, CEACAM5+) co-existed with regions of gastric metaplasia (VSIG1+, LCN2+, MUC5AC+) (Fig. 3.7). These histological observations draw a connection between the CLST2

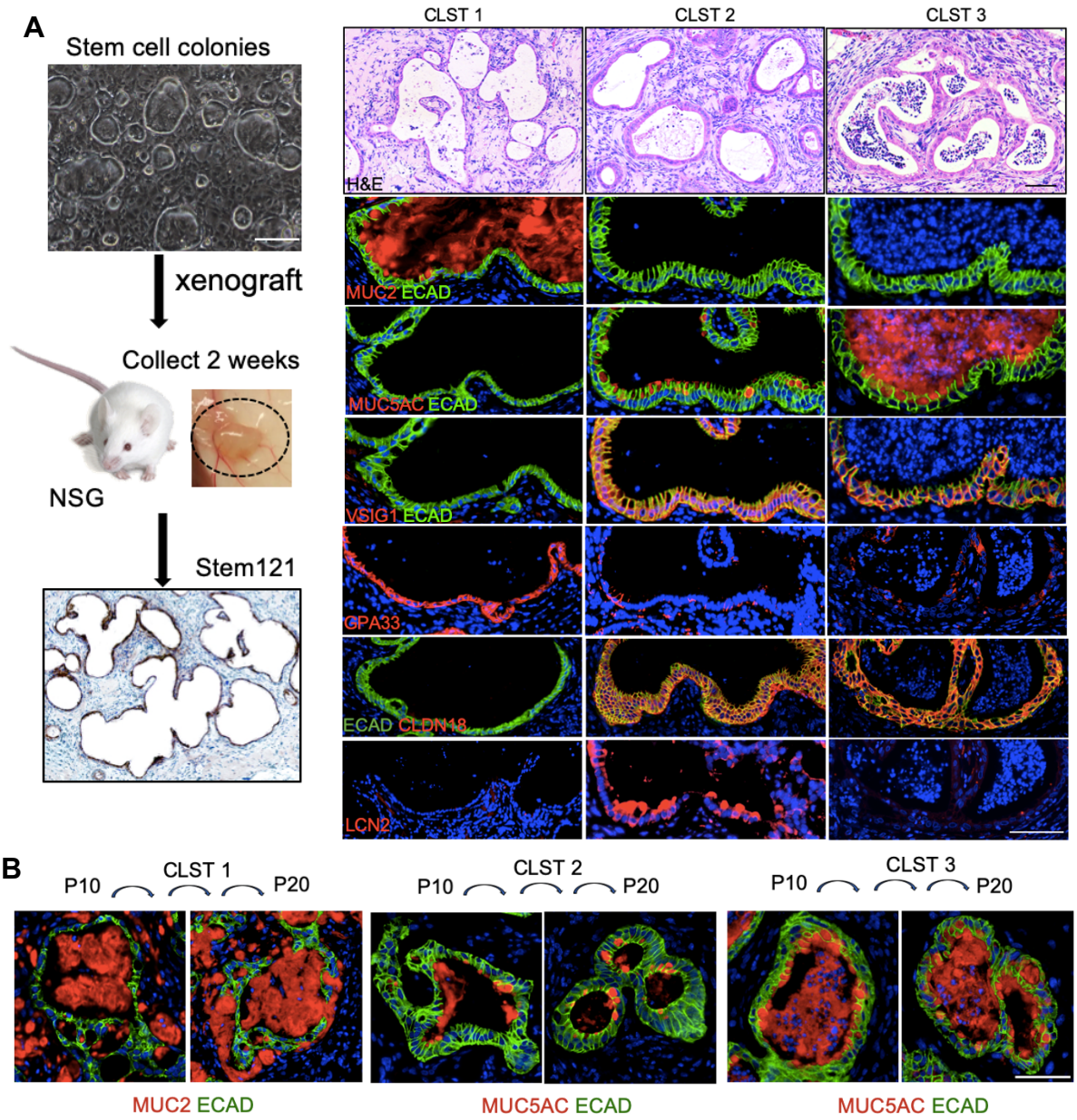


Fig. 3.6. Clonal differentiation in xenograft
A. *Left*, Schematic of subcutaneous transplantation of expanded clones into immunodeficient NSG mice, the generation of a xenograft nodule, and the xenograft histology showing the formation of epithelial cysts that are positive for the human-specific STEM-121 monoclonal antibody. *Right*, Histology sections of xenograft nodules formed two weeks after transplanting two million stem cells expanded from CLST1, CLST2, or CLST3 clones stained with H&E or by immunofluorescence with antibodies to MUC2, MUC5AC, VSIG1, GPA33, CLDN18, LCN2, and ECAD. Scale bar, 100 μm. **B.** Stability of differentiation phenotype of CLST1, CLST2, and CLST3 clones at *in vitro* passage 10 (approximately 100 days) and at passage 20 (approximately 200 days) assessed by *in vivo* subcutaneous transplantation to NSG mice and immunofluorescence with antibodies to ECAD, MUC2, and MUC5AC. Scale bar, 100 μm.

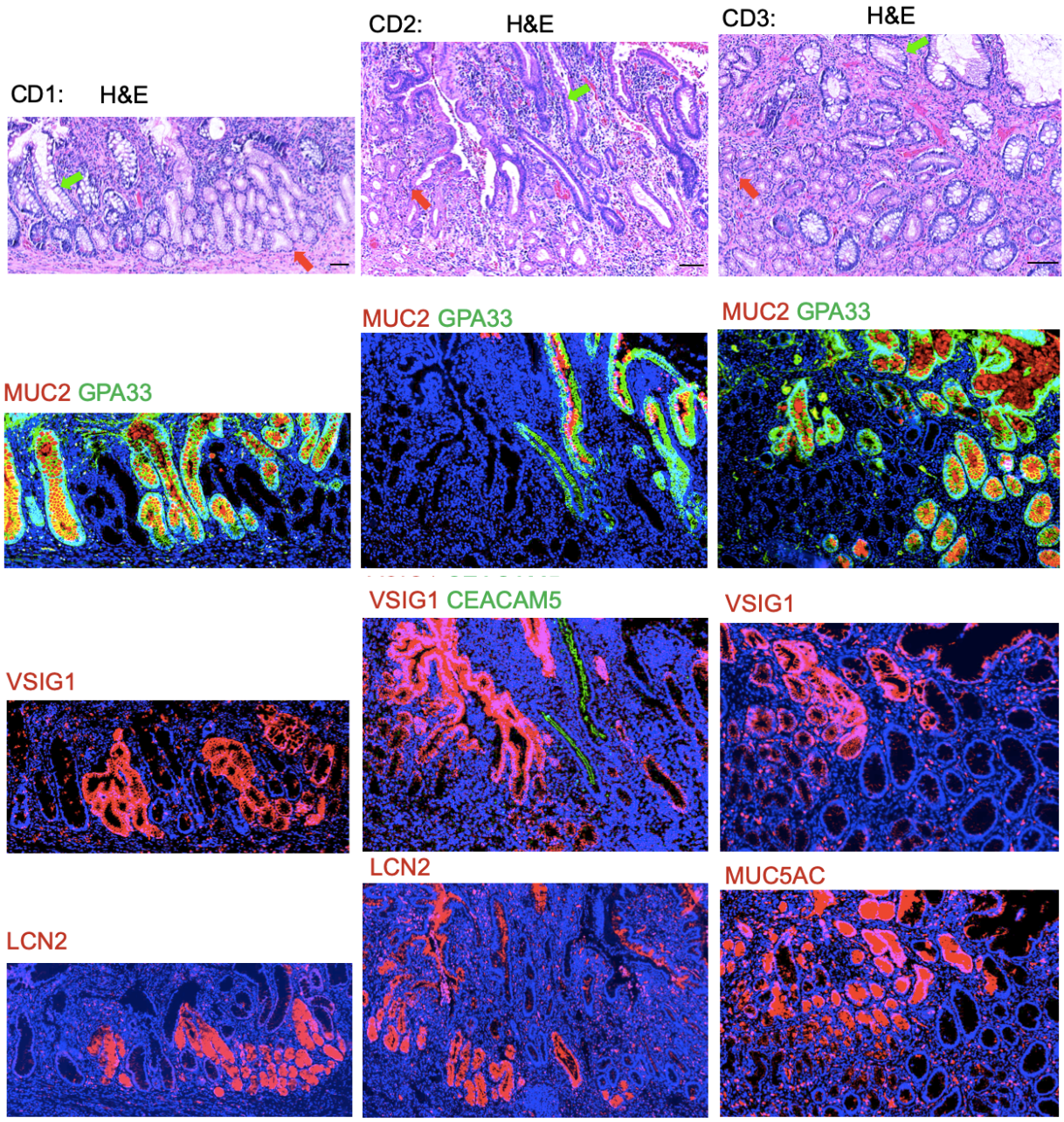


Fig 3.7. CLST 2 and CLST3 markers in three Crohn's patient resected tissue
H&E and Immunofluorescence of histological sections of resected terminal ileum of three Crohn's cases with antibodies to CLST1 markers MUC2 and GPA33, CLST2 and CLST3 markers VSIG1, CLST2 marker LCN2, and the CLST3 marker MUC5AC. Green arrows showing normal mucosa and red arrows showing metaplastic mucosa. Scale bar, 100 μm.

and CLST3 clones that dominate Crohn's biopsies with the well-established findings of gastric metaplasia in affected mucosa of Crohn's patients (Liber, 1951; Lee, 1964; Kariv et al., 2010; Agarwal et al., 2013; Thorsvik et al., 2019).

3.3.3 Crohn's stem cell libraries drive neutrophilic inflammation

Given the existence of at least two variants in 82% of the Crohn's patient studied, we next asked if these CLST2 and CLST3 had any primary role in the manifestation of inflammation which is a major problem in inflammatory bowel disease (IBD). The question remained if they simply aided secondarily in the progression of inflammation because of loss of goblet cells, enteroendocrine cells, Paneth cells, and defect in tight junction protein GPA33 in differentiated epithelia of CLST2 and CLST3, all of which act as the primary defense against pathogens in intestine.

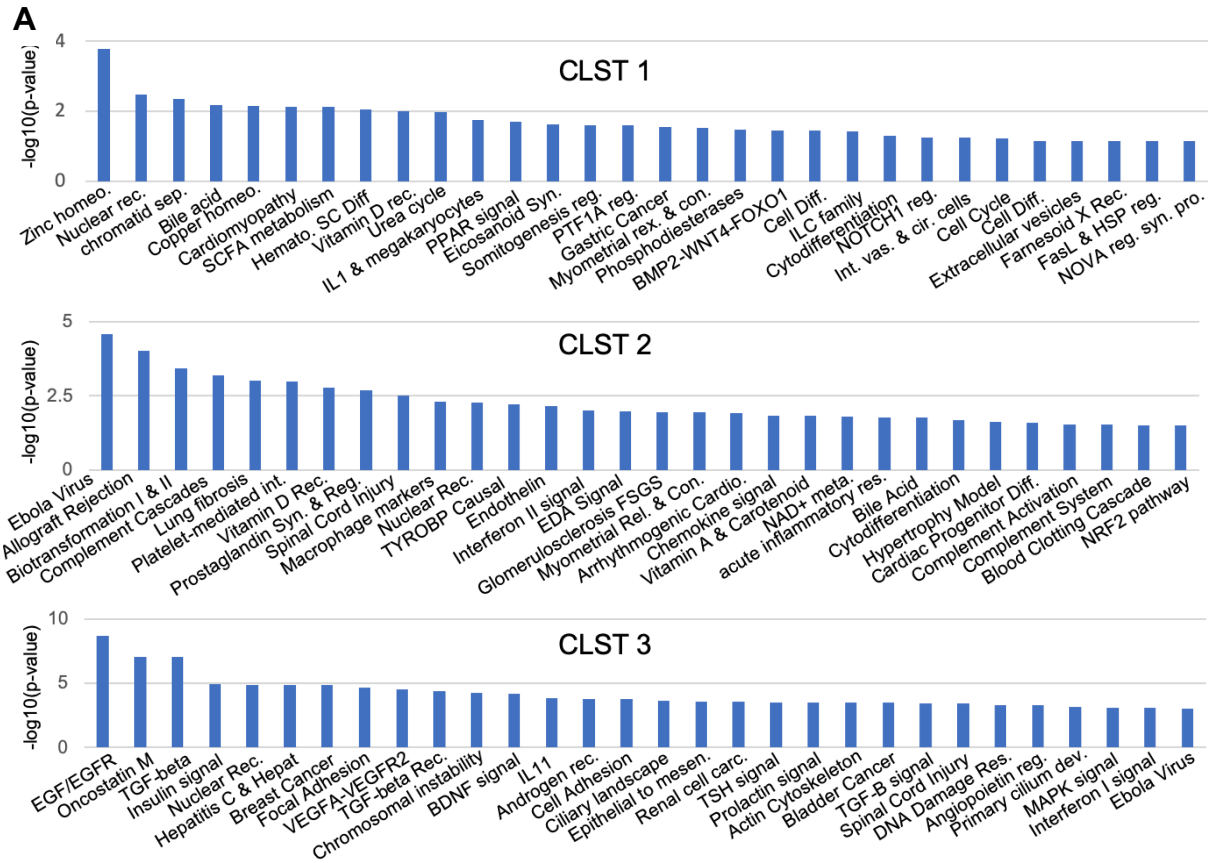
We looked at whole-genome expression profiles of the variant clones across Crohn's patients and controls to answer this question. The gene set enrichment analysis (Chen et al., 2013, Slenter et al., 2018) of differentially expressed gene sets (>1.5 fold, $p<0.05$) between CLST1 and CLST2 and between CLST1 and CLST3 revealed that the top 30 enriched pathways among these three clusters were distinct with each other (Fig. 3.8A). The top five enriched pathways of CLST1 clones were related to metabolisms such as *Zinc Homeostasis*, *Nuclear Receptor Signaling*, *Bile Acid Transport*, and *Copper Homeostasis*. In contrast, the enriched pathways of CLST2 clones were related to inflammation and fibrosis signaling such as *Ebola Virus Responses*, *Allograft Rejection*, *Complement Cascades*, and *Lung Fibrosis*. CLST3 clones also showed similar enriched pathways related to inflammation such as *EGFR, Oncostatin M, TGF-beta*, and the response to *Hepatitis C and Hepatitis B*. The differences seen in enriched pathways were explained by nearly 200 differentially expressed genes related to

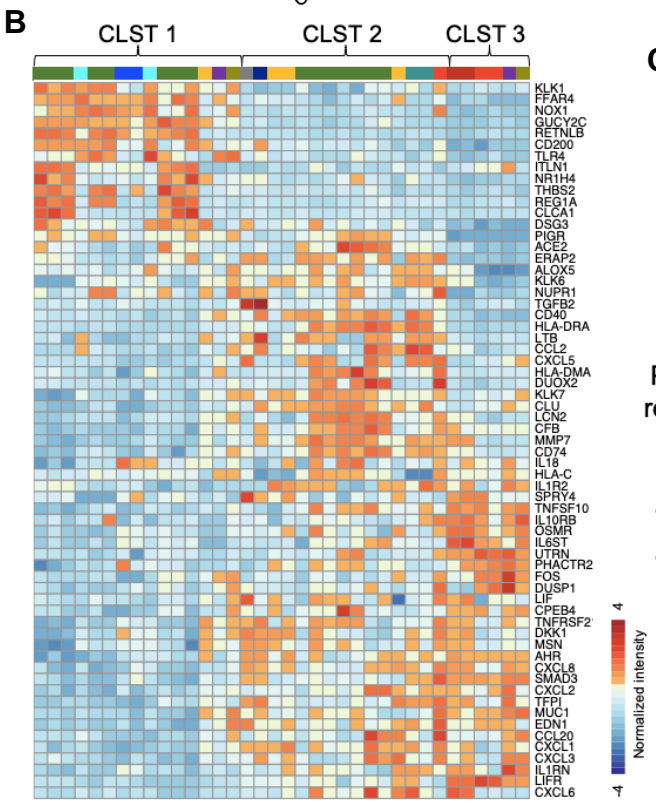
inflammation, inflammatory bowel disease, and fibrosis between CLST1 and CLST2 and between CLST1 and CLST3 clones. A heatmap of selected genes relevant to inflammatory bowel disease and Crohn's disease showed higher expression of genes such as NOX1, GUCY2C, RETNLB, CD200, TLR4, KLK1, FFAR4, NR1H4, THBS2, REG1A, CLCA1, and ITLN1 in CLST1 clones compared to the other two clusters (Fig. 3.8B).

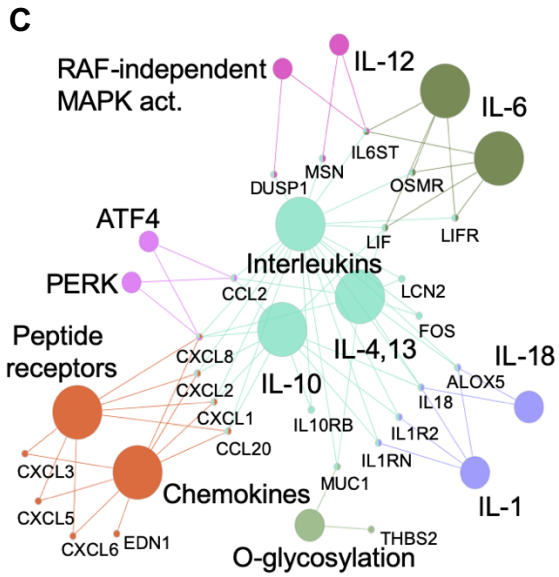
CLST2 and CLST3 clones shared several inflammatory genes, specifically chemokines and their receptors, such as CXCL1, CXCL2, CXCL3, CXCL8, AHR, DDK1, TNFRSF21, TNFSF10, and CCL20. Only CLST2 clones showed expression of some chemokines like CCL2 and CXCL5, genes related to antigen presentation like CD74, HLA-DMA, and HLA-DRA, and other genes like LCN2, CD40, TGFB2, and MMP7. Moreover, CLST3 clones showed specific expression of Oncostatin M pathway genes (OSMR, IL6ST, LIFR; West et al., 2017), interferon signaling (e.g. INFAR1, INFAR2, IL10RB, INFGR1; Schneider et al., 2014), angiogenesis (VEGFA, HIF1A), and the thiopurine target RAC1 (Poppe et al., 2006) (Fig. 3.8B). Next, network analysis (Bindea et al., 2009, Bindea et al., 2013, Shannon et al., 2003) was performed on CLST2 and CLST3 inflammation associated genes which showed overall interactions between inflammatory signaling pathway genes to include those dominated by interleukins 1, 6, 8, 12, and 18, as well as by chemokines such as CXCL1-6 and CCL20 among others (Fig. 3.8C).

Since the gene expression profiles of CLST2 and CLST3 showed shared and unique inflammatory gene signature, we made use of the humanized xenograft mouse model to subcutaneously inject the patient-matched stem cells from Crohn's patient SPN-29 in NSG mice to see if these two variants drive immune response in a host. Surprisingly, we found that only CLST3 clones xenografts drove extra- and intra-luminal cellularization of which intra-luminal cells appeared to be leukocytes by hematoxylin-eosin (H&E) staining and later confirmed by

Fig. 3.8. Constitutive expression of inflammatory mediators in CLST2 and CLST3
A. Wiki pathway enrichment analysis applied to all genes differentially expressed between CLST1 and CLST2 and CLST1 and CLST3 stem cells. **B.** Expression heatmap of a partial list of inflammatory genes differentially represented among CLST1, CLST2, and CLST3. **C.** Network analysis constructed from differentially expressed inflammatory genes in *B*.


A
CLST 1
-log10(p-value)
Zinc homeo.
Nuclear rec.
chromatid sep.
Bile acid
Copper homeo.
Cardiomyopathy
SCFA metabolism
Hemato. SC Diff
Vitamin D rec.
Urea cycle
IL1 & megakaryocytes
PPAR signal
Eicosanoid Syn.
Somitogenesis reg.
PTF1A reg.
Gastric Cancer
Myometrial rex. & con.
Phosphodiesterases
BMP2-WNT4-FOXO1
Cell Diff.
ILC family
Cytodifferentiation
NOTCH1 reg.
Int. vas. & cir. cells
Cell Cycle
Cell Diff.
Extracellular vesicles
Farnesoid X Rec.
FasL & HSP reg.
NOVA reg. syn. pro.
CLST 2
-log10(p-value)
Ebola Virus
Allograft Rejection
Biotransformation I & II
Complement Cascades
Lung fibrosis
Platelet-mediated int.
Vitamin D Rec.
Prostaglandin Syn. & Reg.
Spinal Cord Injury
Macrophage markers
Nuclear Rec.
TYROBP Causal
Endothelin
Interferon II signal
EDA Signal
Glomerulosclerosis FSGS
Myometrial Rel. & Con.
Arrhythmogenic Cardio.
Chemokine signal
Vitamin A & Carotenoid
NAD+ meta.
acute inflammatory res.
Bile Acid
Cytodifferentiation
Hypertrophy Model
Cardiac Progenitor Diff.
Complement Activation
Complement System
Blood Clotting Cascade
NRF2 pathway
CLST 3
-log10(p-value)
EGF/EGFR
Oncostatin M
TGF-beta
Insulin signal
Nuclear Rec.
Hepatitis C & Hepat
Breast Cancer
Focal Adhesion
VEGFA-VEGFR2
TGF-beta Rec.
Chromosomal instability
BDNF signal
IL11
Androgen rec.
Cell Adhesion
Ciliary landscape
Epithelial to mesen.
Renal cell carc.
TSH signal
Prolactin signal
Actin Cytoskeleton
Bladder Cancer
TGF-B signal
Spinal Cord Injury
DNA Damage Res.
Angiopoietin reg.
Primary cilium dev.
MAPK signal
Interferon I signal
Ebola Virus

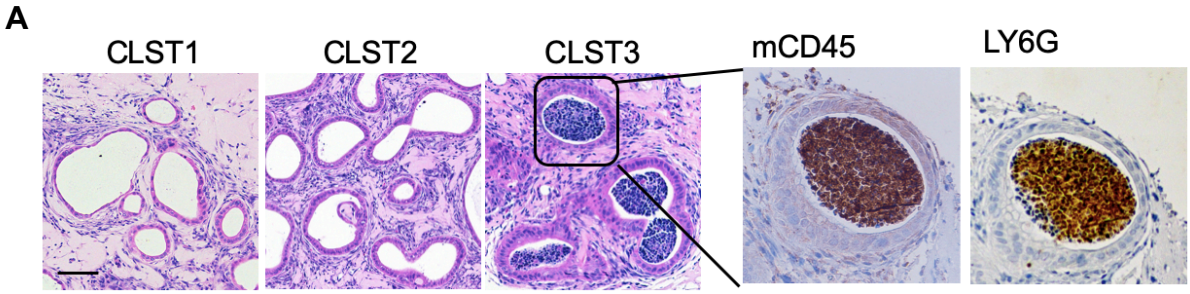

B
CLST 1
CLST 2
CLST 3
KLK1
FFAR4
NOX1
GUCY2C
RETNLB
CD200
TLR4
ITLN1
NR1H4
THBS2
REG1A
CLCA1
DSG3
PIGR
ACE2
ERAP2
ALOX5
KLK6
NUPR1
TGFB2
CD40
HLA-DRA
LTB
CCL2
CXCL5
HLA-DMA
DUOX2
KLK7
CLU
LCN2
CFB
MMP7
CD74
IL18
HLA-C
IL1R2
SPRY4
TNFSF10
IL10RB
OSMR
IL6ST
UTRN
PHACTR2
FOS
DUSP1
LIF
CPEB4
TNFRSF2
DKK1
MSN
AHR
CXCL8
SMAD3
CXCL2
TFPI
MUC1
EDN1
CCL20
CXCL1
CXCL3
IL1RN
LIFR
CXCL6
4
-4
Normalized intensity

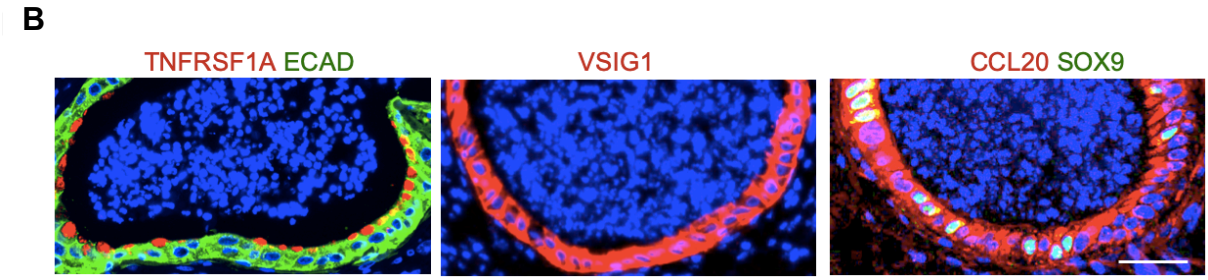

C
RAF-independent
MAPK act.
IL-12
IL-6
IL6ST
MSN
DUSP1
OSMR
ATF4
LIF
LIFR
Interleukins
PERK
CCL2
LCN2
FOS
Peptide
receptors
CXCL8
IL-4,13
IL-18
IL-10
ALOX5
CXCL2
IL18
CXCL1
IL10RB
IL1R2
CCL20
IL1RN
CXCL3
Chemokines
MUC1
IL-1
CXCL5
CXCL6
EDN1
THBS2
O-glycosylation

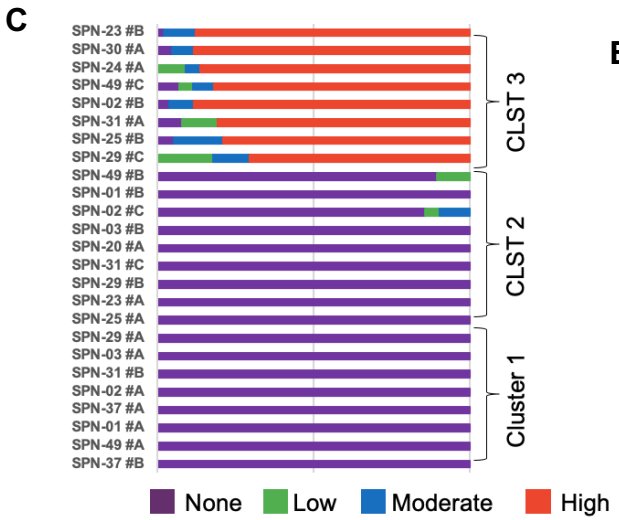
reaction to hematopoietic lineage marker CD45 and neutrophil marker LY6G (Rose et al., 2012) antibodies in immunohistochemistry staining (Fig. 3.9A). CLST1 clones (from Crohn's patient SPN-29 and control patient SPN-19) induced no intra- or extra-luminal cellularization while CLST2 (from SPN-29) induced extra-luminal cellularization which was negatively stained for CD45 and LY6G (Fig. 3.9A). Some of the genes expressed in CLST3 clone stem cells, including TNFRSF1A and CCL20, also showed expression in the epithelia produced in xenografts of cloned CLST3 cells (Fig. 3.9B). We expanded the xenografts to seven more clones each representing CLST 1, 2, and 3 to gain more confidence in the result we observed from xenograft of Crohn's patient SPN-29 clones. Cross-section staining of those clones xenograft in hematoxylin-eosin (H&E) and in immunohistochemistry to CD45 and LY6G antibodies still showed only CLST3 clones xenograft yielded neutrophilic intra-luminal accumulation while CLST2 showed extra-luminal cellularization that was not leukocytes. The result is shown in the histogram constructed by quantifying neutrophilic accumulation (Brazil and Parkos, 2018) in the lumen of epithelia formed in xenografts based on morphometric scoring standards (Fig. 3.9C). Moreover, the neutrophil immune response by the host to CLST3 xenografts and differential gene expression in CLST1, 2, and 3 clones remained stable even after continuous *in vitro* passaging over 250 days (Fig. 3.9 D, E). Hematoxylin-Eosin (H&E) staining of CLST3 clone xenograft from Crohn's patient SPN-29 showed equally strong neutrophil accumulation in lumen after extensive 20 passaging (Fig. 3.9D). Similarly, histogram showed the relative expression of three markers VSIG1, TNFRSF1A, and CCL20 in RT-qPCR remained unchanged in CLST1, 2, and 3 despite extensive passaging (Fig. 3.9E).

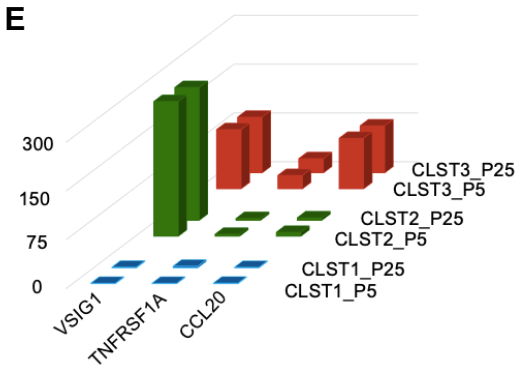
Next, we asked if neutrophilic accumulation was a feature of Crohn's patients' stem cell libraries given that the feature was specific to CLST3 clones and the proportion of CLST3

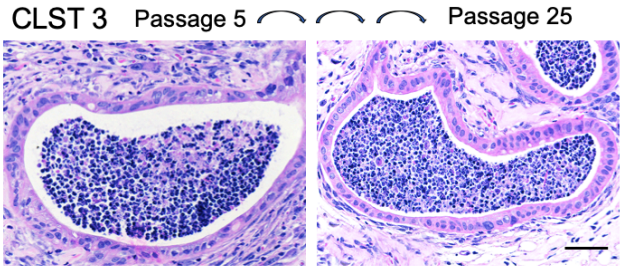
Fig. 3.9. CLST3 clones drive neutrophilic inflammation in xenografts
A. Histology of xenografts resulting from transplants of individual CLST1, CLST2, or CLST3 clones showing that only CLST3 xenografts trigger infiltration by neutrophils. Scale bar, 100 μm. **B.** Expression of discrete inflammatory markers, including TNFRST1A and CCL20, in xenograft epithelia generated from CLST3 clones. **C.** Assessment of the degree of neutrophil infiltration in xenografts resulting from transplantation of CLST1, CLST2, and CLST3 clones. **D.** Stable neutrophil inflammation from CLST3 clones propagated to passage 5 (P5) and passage 25 (P25) *in vitro*. Scale bar, 100 μm. **E.** Histogram of VSIG1, TNFRSF1A, and CCL20 gene expression in clonal representatives of Clusters 1-3 at *in vitro* passage 5 and passage 25.


A
CLST1
CLST2
CLST3
mCD45
LY6G


B
TNFRSF1A ECAD
VSIG1
CCL20 SOX9


C
SPN-23 #B
SPN-30 #A
SPN-24 #A
SPN-49 #C
SPN-02 #B
SPN-31 #A
SPN-25 #B
SPN-29 #C
SPN-49 #B
SPN-01 #B
SPN-02 #C
SPN-03 #B
SPN-20 #A
SPN-31 #C
SPN-29 #B
SPN-23 #A
SPN-25 #A
SPN-29 #A
SPN-03 #A
SPN-31 #B
SPN-02 #A
SPN-37 #A
SPN-01 #A
SPN-49 #A
SPN-37 #B
CLST 3
CLST 2
Cluster 1
None
Low
Moderate
High

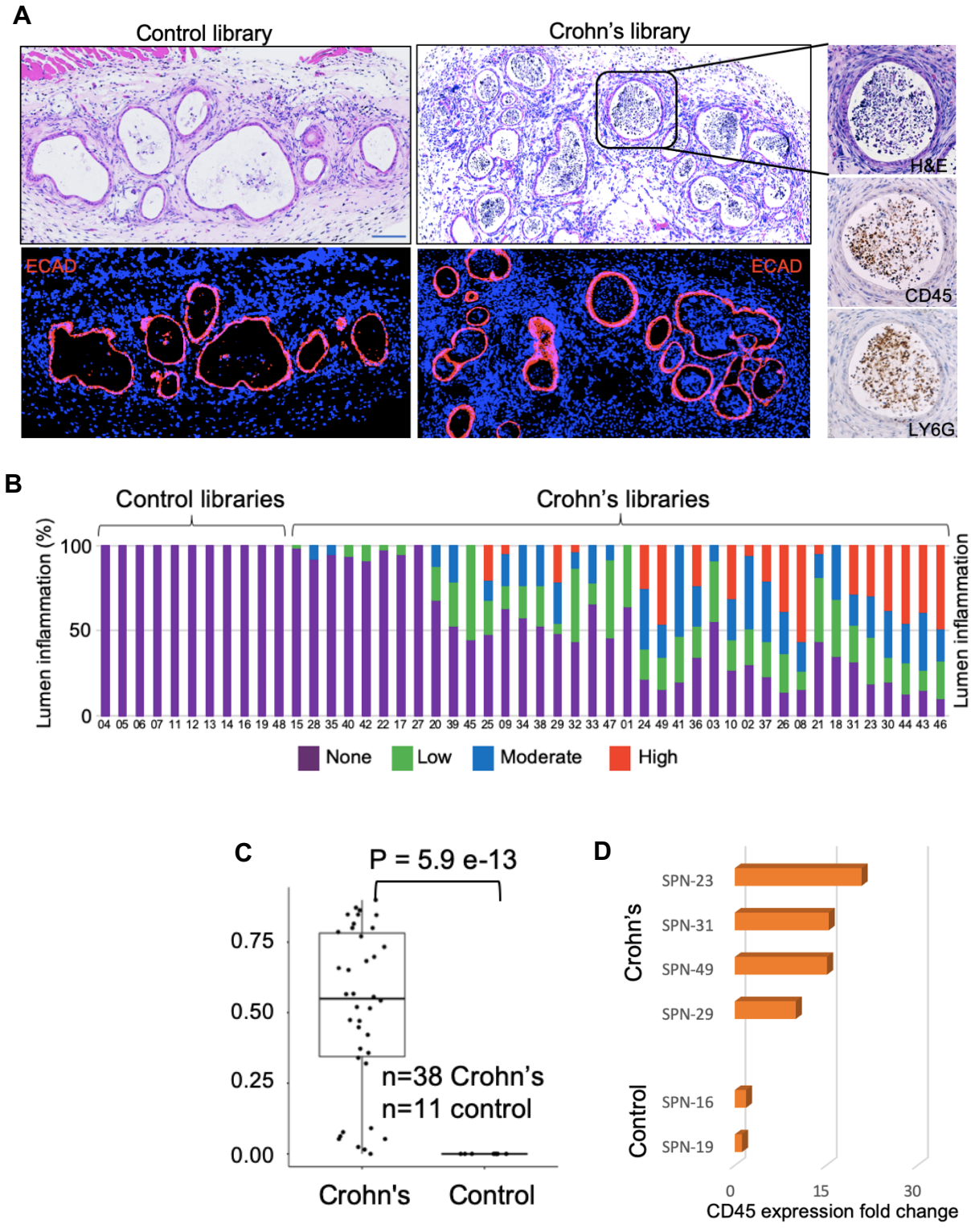

E
300
150
75
0
VSIG1
TNFRSF1A
CCL20
CLST3_P25
CLST3_P5
CLST2_P25
CLST2_P5
CLST1_P25
CLST1_P5


D
CLST 3
Passage 5
Passage 25

clones (by PSCA) varied among libraries (0.5% to 63%) (Fig. 3.3D). Upon xenograft of 38 Crohn's patients and 11 control patient stem cell libraries in NSG mice, we observed that 51% of Crohn's patients including SPN-29 were distinguished by high intra-luminal infiltration of neutrophils and 25% by moderate infiltration (Fig. 3.10A, B, D). The remaining 24% had either low levels of infiltration or no infiltration at all. All the control libraries' xenografts including SPN-19 showed no signs of neutrophil accumulation. The entire 49 patients' libraries xenograft result is shown in a histogram based on quantification of neutrophilic accumulation by morphometric standards (Fig. 3.10B). Box plot analysis of the fraction of lumen infiltration (high + moderate + low) in Crohn's patients vs control patients showed that infiltration by neutrophils is a major distinguishable feature of Crohn's library xenografts ($P= 5.9\ e^{-13}$) (Fig. 3.10C). The difference in the degree of neutrophilic infiltration (high or moderate or low or none) across Crohn's patients was justified by the proportion of the CLST3 stem cells in that particular patient library (Fig. 3.11A). The fraction of high and moderate lumen infiltration in xenografts of Crohn's patient libraries strongly correlated (coefficient of correlation=0.82, $P= 9.7\ e^{-13}$) to the proportion of PSCA positive stem cells (CLST3) within that library suggesting a higher degree of infiltration was due to higher proportion of CLST3 stem cells and vice-versa (Fig. 3.11A).

Genome-wide association studies of Crohn's patients have been enormously successful in identifying risk loci and candidate genes that have shaped our understanding of this disease (Rioux et al., 2007; Hampe et al., 2007; Barrett et al, 2008; Franke et al., 2010). As the CLST2 and CLST3 stem cells derived from Crohn's terminal ileum showed pro-fibrotic and pro-inflammatory gene profiles and corresponding activities host responses upon transplantation to mice, we asked if the differentially expressed genes in CLST2 and CLST3 clones include genes implicated by GWAS in the risk for Crohn's. Using gene expression criteria (1.5-fold, $p<0.05$), we identified 75 genes differentially expressed in CLST2 and CLST3 stem cells previously

Fig. 3.10. Pro-inflammatory activities of Crohn's libraries
A. Histology of xenografts of control and Crohn's libraries stained with H&E or immunofluorescence with antibodies to ECAD and counterstained with DAPI for nuclei. Inset focuses on a single epithelial cyst stained with H&E, and antibodies to the hematopoietic marker CD45 and the neutrophil marker LY6G. Scale bars, 100 μm. **B.** Histogram of lumen inflammation scored by low, moderate, and high lumen standards across 11 control and 38 Crohn's patient-specific libraries. **C.** Box plots of total lumen inflammation (low + moderate + high) across 11 controls and 38 Crohn's libraries with indicated medians, Q1s, Q3s, and p-value. **D.** Histogram of CD45 expression fold change between control and Crohn's library xenografts.


A
Control library
Crohn's library
H&E
ECAD
ECAD
CD45
LY6G
B
Control libraries
Crohn's libraries
Lumen inflammation (%)
100
50
0
04 05 06 07 11 12 13 14 16 19 48 15 28 35 40 42 22 17 27 20 39 45 25 09 34 38 29 32 33 47 01 24 49 41 36 03 10 02 37 26 08 21 18 31 23 30 44 43 46
Lumen inflammation
None
Low
Moderate
High
C
P = 5.9 e-13
0.75
0.50
0.25
0.00
n=38 Crohn's
n=11 control
Crohn's
Control
D
Crohn's
SPN-23
SPN-31
SPN-49
SPN-29
Control
SPN-16
SPN-19
0
15
30
CD45 expression fold change

implicated in Crohn's by three distinct GWAS analyses (Hampe et al., 2007; Rioux et al., 2007; Liu et al., 2015) and by computational analyses of linkage disequilibrium (LD) blocks (e.g. GRAIL, Raychaudhuri et al., 2009); Fig. 3.11B). A majority of the matches we identified involve single genes among several within LD blocks (e.g. NUPR1 at rs26526), and several highlights multiple genes coordinately upregulated within a single LD block containing related genes such as CXCL1, 2, 3, and 5 at rs2472649 (Jostins et al., 2012).

3.3.4 Variant stem cells drive fibrotic host responses

The top five enriched pathways for variant clones from nine Crohn's patients included *Lung Fibrosis* and *TGF-beta* pathways both of which are implicated in fibrosis (Araya et al., 2010) (Fig. 3.8A). Fibrosis is a major complication of Crohn's disease causing bowel obstruction and immobility because of which surgical intervention is required (Burke et al., 2007; Baumgart and Sandborn, 2012; D'Haens et al., 2019; Yoo et al., 2020). Further dissecting gene expression profiles of variant clones for fibrosis-related genes, we found at least 50 differentially expressed genes among CLST 1, 2, and 3 from nine Crohn's patients including SPN-29 and two control patients including SPN-19 (Fig. 3.12A). The heatmap shows that genes such as CCL24, CDH11, IL1B, F2R, TLR4, RERTNLB, TPSG1, and RGS16 were highly expressed in CLST1 clones compared to the other two variants. CLST2 and CLST3 clones shared some of the genes which were relatively high in these clusters compared to CLST1 clones. The genes include PLAU, TGFA, PDGFC, PLA2G10, TIMP1, NAPEPLD, AHR, and SMAD3. Similarly, CLST2 clones alone had higher expression of several genes such as CCR1, BAX, CFTR, NT5E, ALOX5, CFLAR, MMP7, C3, CD74, HLA-DRA, and TGFB2. Moreover, CLST3 clones had higher expression of interferon and interleukin receptor genes (such as IFNAR1, IFNAR2, IFNGR1, IL1RN, and IL6ST), the thiopurine target RAC1, and other genes such as FAM11B, FGFR2, PIK3CB, NFKBIA, CRBN, GCLC, GOLM1, MMP12, CAV1, TBGB1, PPARG, EGFR1,

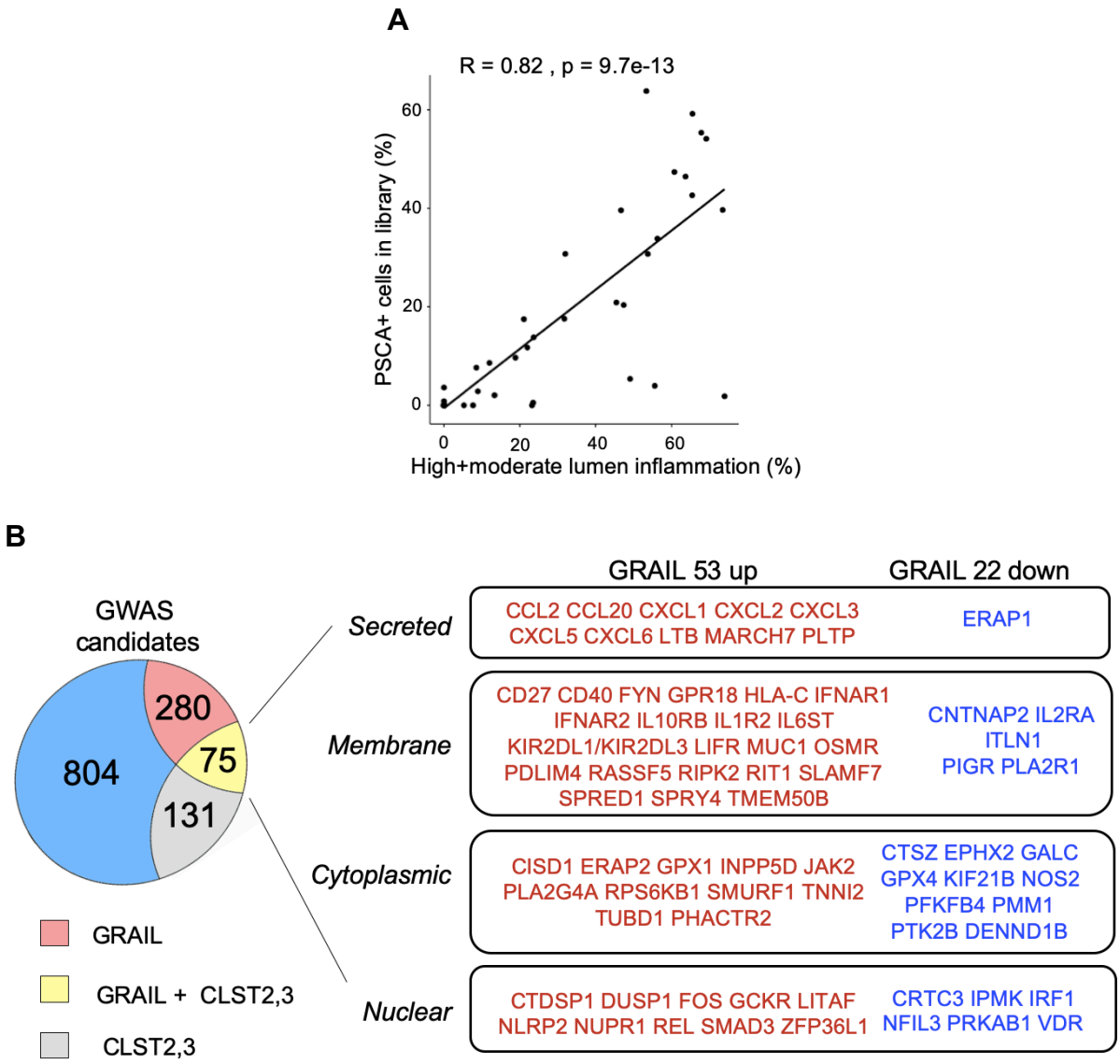


Fig. 3.11. Inflammatory signatures of Crohn's variant stem cells overlap with GWAS
A. Graph showing the correlation between lumen with high and moderate inflammation in library xenografts and the proportion of PSCA+ cells in the respective libraries. R=correlation coefficient. **B.** Overlap between differentially expressed genes in CLST2 and CLST3 (versus CLST1) clones and the 1,290 genes within linkage disequilibrium blocks implicated by three separate GWAS studies. Of the 206 CLST2 and CLST3 genes overlapping with the LD blocks, 75 overlaps with those in the LD block implicated by the GRAIL algorithm including 53 differentially over-expressed and 22 genes differentially under-expressed.

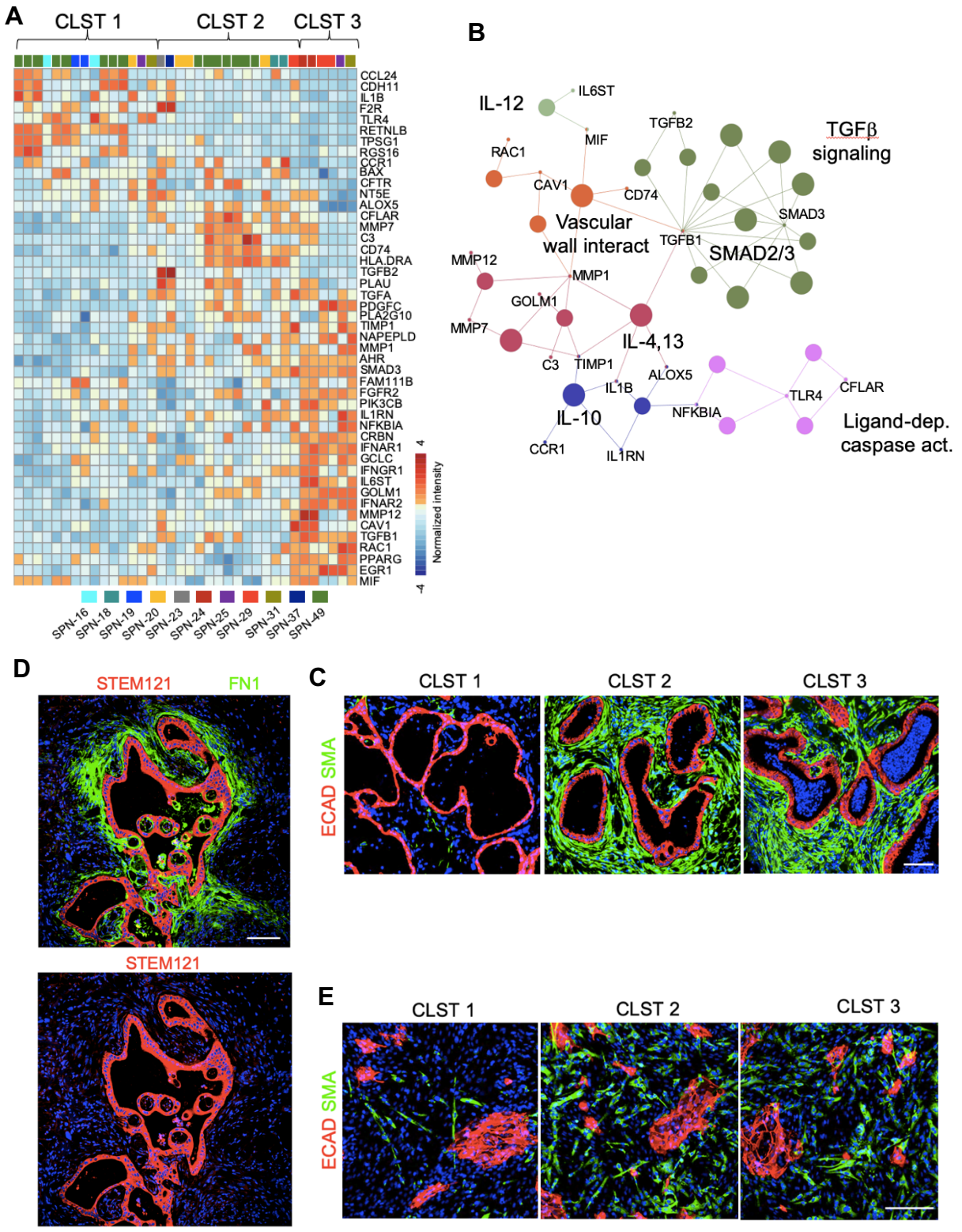
and MIF. The network analysis was performed on CLST2 and CLST3 fibrosis-associated genes which showed overall interactions between fibrosis signaling pathway genes to include those dominated by interleukins 4, 6, 10, 12, and 13, as well as by TGF-b and SMAD2/3 among others (Fig. 3.12B).

Given the pro-fibrotic gene signatures from both CLST2 and CLST3 compared to CLST1, we tested if these variant clones from Crohn's patient SPN-29 are driving fibrosis as host response from xenografts injected in NSG mice. Alpha-smooth muscle actin (α-SMA)-expressing myofibroblasts have been linked to fibrosis in Crohn's and other chronic inflammatory conditions through their contractile properties and their production of extracellular matrix including collagen fibrils (Araya et al., 2010; Li and Kuemmerle, 2020). The extra-luminal cells in both CLST2 and CLST3 xenografts (from Crohn's patient SPN-29) cross-sections stained strongly reacted to α-SMA antibody in immunofluorescence while CLST1 xenograft (from Crohn's patient SPN-29 and control patient SPN-19) stained mostly negative (Fig. 3.12C). We note here that the myofibroblasts in these xenografts, which express both α-SMA and fibronectin (FN1), do not react to anti-human STEM-121 antibodies (Fig. 3.12D) and therefore are of host origin. The fibrotic host response *in vivo* was also evident when we cocultured CLST 1, 2, and 3 clones from Crohn's patient SPN-29 with human fibroblast *in vitro*. Both CLST2 and CLST3 clones induced differentiation of α-SMA positive myofibroblasts after three days of coculture further confirming *in vivo* results (Fig. 3.12E). We expanded the immunoreactivity assay to xenografts from seven more clones each representing CLST 1, 2, and 3 to gain more confidence in the result we observed from xenograft of Crohn's patient SPN-29 clones. Cross-section staining of those clones' xenograft in immunofluorescence to α-SMA and Fibronectin1 (FN1) antibodies still showed both CLST2 and CLST3 xenograft with the extra-luminal accumulation of myofibroblasts while CLST1 showed no extra-luminal accumulation. The result

is shown in the histogram constructed by quantifying α-SMA positive myofibroblasts accumulation (Brazil and Parkos, 2018) sub-mucosal of epithelia formed in xenografts based on morphometric scoring standards (Fig. 3.12F). Moreover, α-SMA myofibroblast accumulation response by the host to CLST2 and CLST3 xenografts remained stable even after continuous *in vitro* passaging over 250 days. Immunofluorescence staining of CLST2 and CLST3 clone xenografts from Crohn's patient SPN-29 showed equally strong extra-luminal myofibroblast accumulation after extensive 20 passaging (Fig. 3.12G).

Next, we asked if myofibroblast accumulation was a feature of Crohn's patients' stem cell libraries given that the feature was specific to both CLST2 and CLST3 clones and the proportion of CLST2 and CLST3 clones (by VSIG1) varied among libraries (64.6%+/-32.6%) (Fig. 3.3D). The xenograft of a single Crohn's patient SPN-30 and control patient SPN-19 libraries in NSG mice showed extensive accumulation of myofibroblasts in xenograft of Crohn's library compared to control library (Fig. 3.13A). Further, this observation was supported by the high expression of SMA in xenografts from four Crohn's libraries assessed by qPCR (Fig. 3.13B). Upon xenograft of 38 Crohn's patients' and 11 control patients' stem cell libraries in NSG mice, we observed that 82% of Crohn's patients including SPN-29 were distinguished by the high extra-luminal accumulation of myofibroblasts. The remaining 18% had either low levels of accumulation or no accumulation at all. All the control libraries' xenografts including SPN-19 showed no signs of myofibroblasts accumulation. The entire 49 patients' libraries xenograft result is shown in a histogram based on quantification of myofibroblast accumulation by morphometric standards (Fig. 3.13C). Box plot analysis of the fraction of strong myofibroblast accumulation in Crohn's patients vs control patients showed that fibrosis is a major distinguishable feature of Crohn's library xenografts ($P< 1.3\ e^{-14}$) (Fig. 3.13D). The difference in the degree of myofibroblast accumulation (strong or moderate or none) across Crohn's patients

Fig. 3.12. Pro-fibrotic activities of CLST2 and CLST3 clones
A. Expression heatmap of fibrosis-related genes that are differentially expressed in whole-genome expression dataset of CLST1, CLST2, and CLST3 clones. **B.** Network analysis using differentially expressed fibrosis-related genes in *A*. **C.** Immunofluorescence labeling of ECAD and α-SMA on histological sections of xenograft nodules resulting from transplants of CLST1, CLST2, and CLST3. Scale bar, 100 μm. **D.** Histological sections of Crohn's library xenograft stained with the anti-human Stem121 antibody and antibodies to fibronectin (FN1), a marker of myofibroblasts. Scale bar, 100 μm. **E.** Immunofluorescence labeling of ECAD and α-SMA on CLST1, CLST2, and CLST3 clones co-cultured with human fibroblast *in vitro*. Scale bar, 100 μm. **F.** Quantification of the extent of submucosal myofibroblast accumulation in xenograft nodules resulting from transplants of discrete clones of CLST1, CLST2, and CLST3. **G.** Immunofluorescence with antibodies to ECAD (red) and α-SMA (green) on histological sections of xenografts derived from CLST2 clones at P5 and P25 (left) or CLST3 clones at P5 and P25 showing epigenetic stability of the profibrotic phenotype. Scale bar, 100 μm.


A
CLST 1
CLST 2
CLST 3
CCL24
CDH11
IL1B
F2R
TLR4
RETNLB
TPSG1
RGS16
CCR1
BAX
CFTR
NT5E
ALOX5
CFLAR
MMP7
C3
CD74
HLA.DRA
TGFB2
PLAU
TGFA
PDGFC
PLA2G10
TIMP1
NAPEPLD
MMP1
AHR
SMAD3
FAM111B
FGFR2
PIK3CB
IL1RN
NFKBIA
CRBN
IFNAR1
GCLC
IFNGR1
IL6ST
GOLM1
IFNAR2
MMP12
CAV1
TGFB1
RAC1
PPARG
EGR1
MIF
4
-4
Normalized intensity
SPN-16
SPN-18
SPN-19
SPN-20
SPN-23
SPN-24
SPN-25
SPN-29
SPN-31
SPN-37
SPN-49
B
IL-12
IL6ST
MIF
TGFB2
TGFβ
signaling
RAC1
CAV1
CD74
SMAD3
Vascular
wall interact
TGFB1
SMAD2/3
MMP12
MMP1
GOLM1
MMP7
IL-4,13
C3
TIMP1
ALOX5
IL1B
TLR4
CFLAR
IL-10
NFKBIA
Ligand-dep.
caspase act.
CCR1
IL1RN
D
STEM121
FN1
STEM121
C
CLST 1
CLST 2
CLST 3
ECAD SMA
E
CLST 1
CLST 2
CLST 3
ECAD SMA

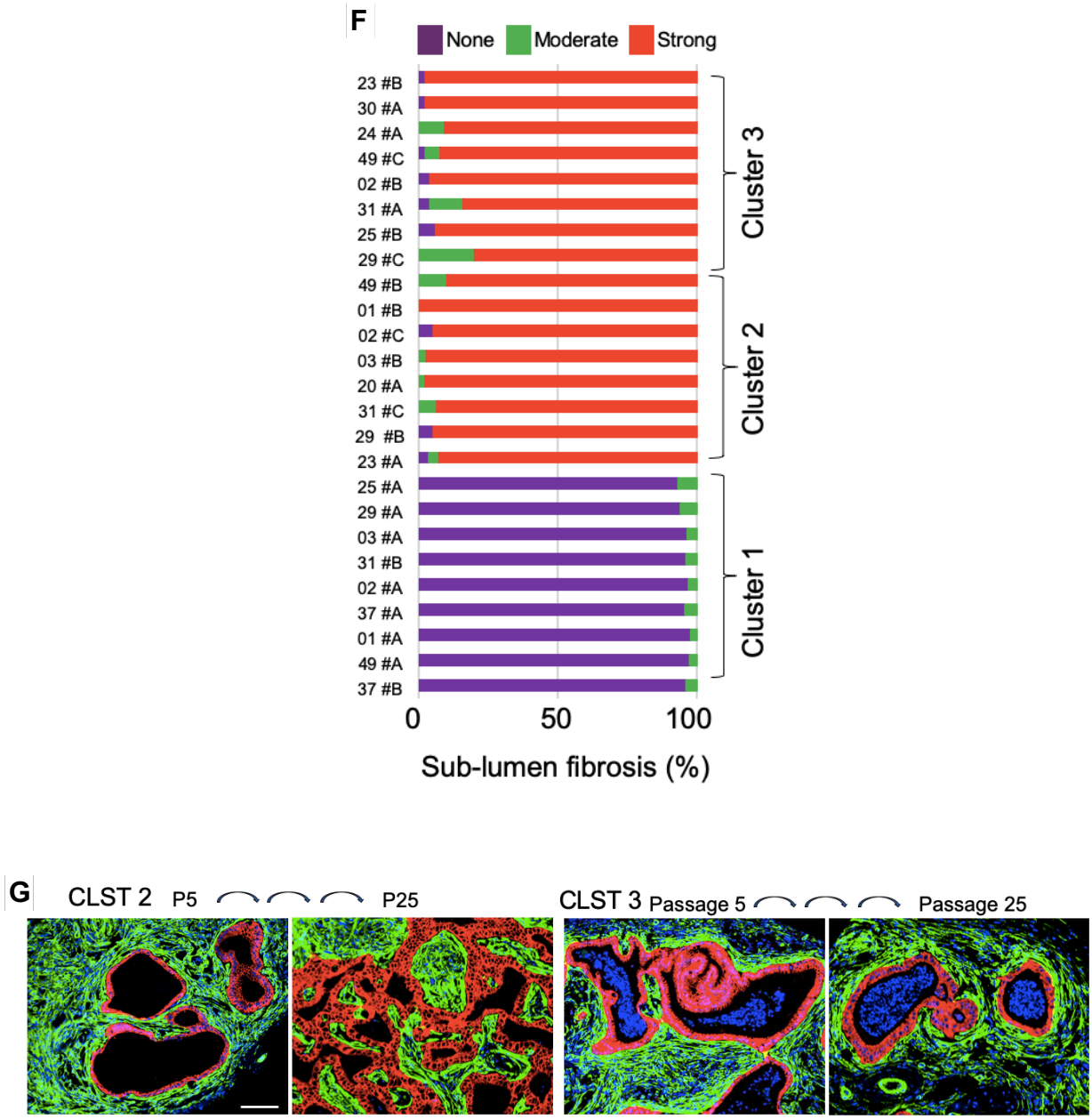

F
None
Moderate
Strong
23 #B
30 #A
24 #A
49 #C
02 #B
31 #A
25 #B
29 #C
49 #B
01 #B
02 #C
03 #B
20 #A
31 #C
29 #B
23 #A
25 #A
29 #A
03 #A
31 #B
02 #A
37 #A
01 #A
49 #A
37 #B
Cluster 3
Cluster 2
Cluster 1
0
50
100
Sub-lumen fibrosis (%)
G
CLST 2
P5
P25
CLST 3
Passage 5
Passage 25

Fig. 3.12 (cont.)

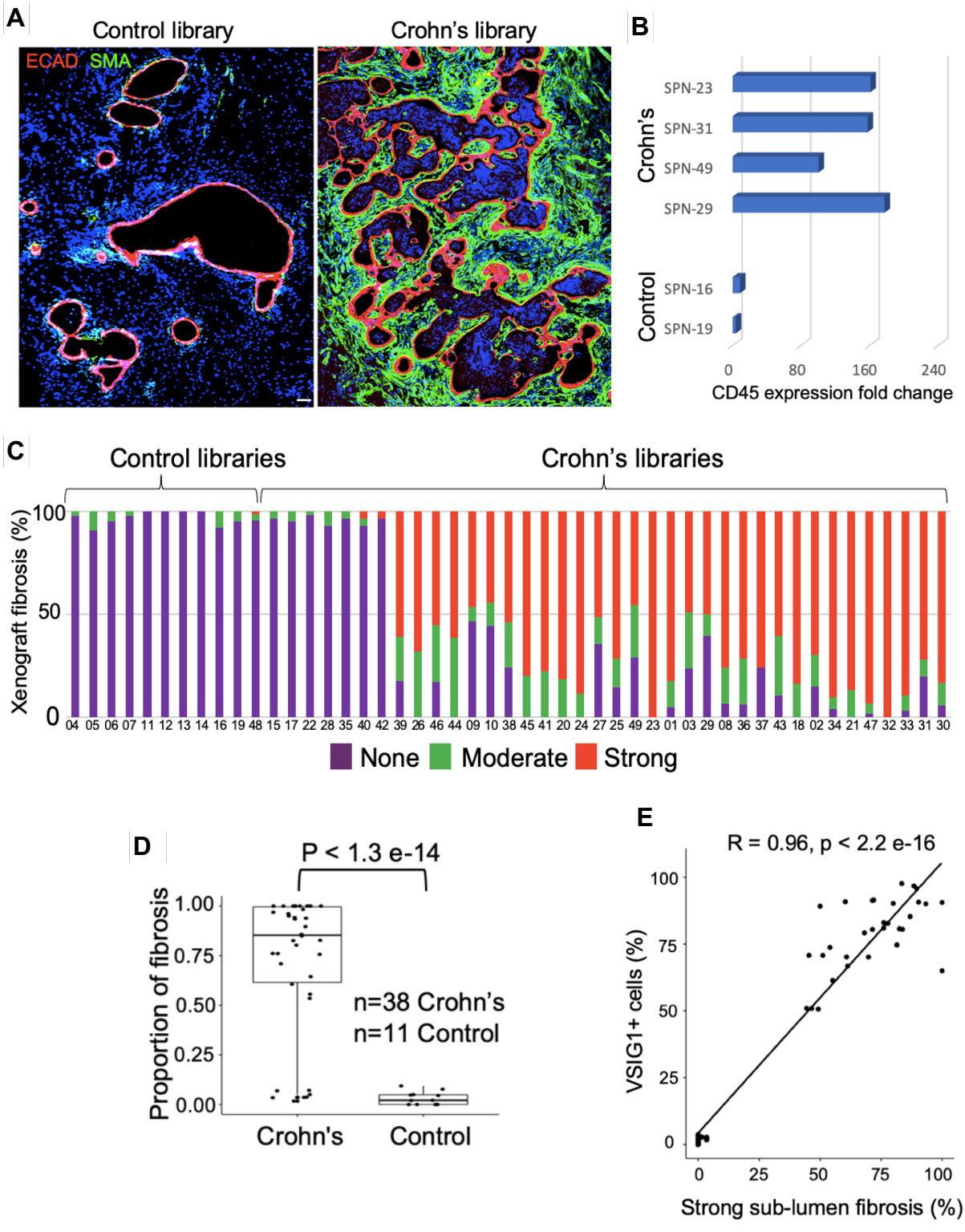
was justified by the proportion of the CLST2 and CLST3 stem cells in that particular patient library. The fraction of strong myofibroblast accumulation in xenografts of Crohn's patient libraries strongly correlated (coefficient of correlation=0.96, $P< 2.2\ e^{-16}$) to the proportion of VSIG1 positive stem cells (CLST2+3) within that library suggesting a higher degree of accumulation was due to a higher proportion of CLST2 and CLST3 stem cells and vice-versa (Fig. 3.13E).

3.3.5 Crohn's variant stem cells preexist in fetal terminal ileum

The majority of the Crohn's patients studied (82%) were dominated by CLST2 and CLST3 variants and this was a distinguishable feature of Crohn's ($P=2.5\ e^{-14}$). We also strongly correlated different proportions of these variants to the extent of inflammation ($P=9.7\ e^{-13}$) and fibrosis ($P<2.2\ e^{-16}$), two major complications of Crohn's (Fig. 3.11A, Fig. 3.13E). Given that the variants found in Crohn's patients are likely to be the drivers of this disease, we now wanted to understand the origin of these. Understanding the origin is important in assessing the roles of these variants in the progression of the disease and perhaps in designing new therapeutics.

Although CLST2 and CLST3 stem cells dominated Crohn's terminal ileum libraries, they were also evident in control case SPN-19 at very minor ratios revealed by single-cell RNA-sequencing (Fig. 3.1D). The variant stem cells found in more control cases (1.2%+/-1.3%) at much lower ratios compared to CLST1 suggested that they might be reflections of early disease (Fig. 3.3D). The disease possibly remained latent in control cases as ratios of CLST2 and CLST3 are very low. But this still left an open-ended question of whether the occurrence of these variants in control cases was a result of a response to something that is experienced by humans in their lifetime. To answer that question, we extended our study to prenatal terminal ilea of 21- and 22-weeks fetal demise cases. Fetal development occurs in a sterile environment

Fig. 3.13. Pro-fibrotic activities of Crohn's libraries
A. Immunofluorescence detection of antibodies to α-SMA (green) and ECAD (red) on histological sections of xenografts of control and Crohn's stem cell libraries. Scale bar, 100 μm. **B.** Histogram of α-SMA expression fold change between control and Crohn's library xenografts. **C.** Quantitative assessment of submucosal myofibroblasts in xenografts of 11 control and 38 Crohn's stem cell libraries. **D.** Box plot of an extent of submucosal myofibroblasts in xenografts of 11 controls and 38 Crohn's stem cell libraries. Medians, Q1, Q3, and p-values are indicated. **E.** Plot of correlation between the percentage of VSIG1+ cells in a stem cell library and the percent of lumen displaying strong submucosal myofibroblast accumulation. Correlation coefficient and p values are indicated.

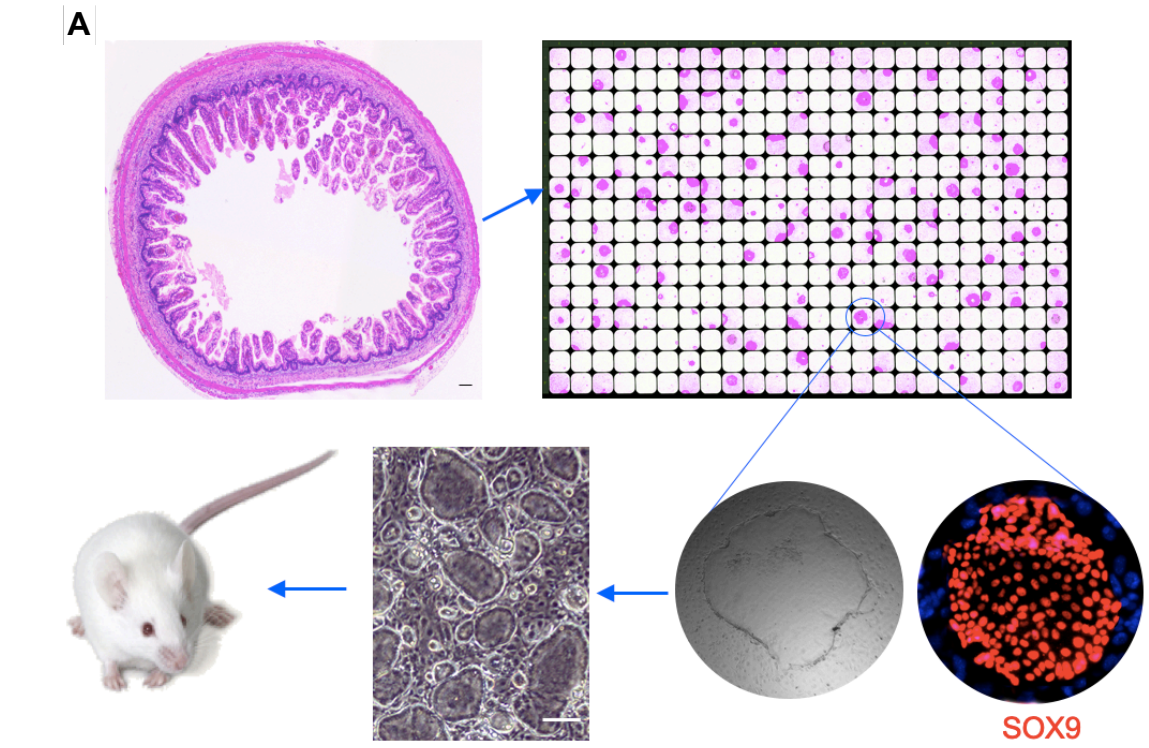

A
Control library
Crohn's library
ECAD
SMA
B
Crohn's
SPN-23
SPN-31
SPN-49
SPN-29
Control
SPN-16
SPN-19
0
80
160
240
CD45 expression fold change
C
Control libraries
Crohn's libraries
Xenograft fibrosis (%)
100
50
0
04 05 06 07 11 12 13 14 16 19 48 15 17 22 28 35 40 42 39 26 46 44 09 10 38 45 41 20 24 27 25 49 23 01 03 29 08 36 37 43 18 02 34 21 47 32 33 31 30
None
Moderate
Strong
D
P < 1.3 e-14
Proportion of fibrosis
1.00
0.75
0.50
0.25
0.00
n=38 Crohn's
n=11 Control
Crohn's
Control
E
R = 0.96, p < 2.2 e-16
VSIG1+ cells (%)
100
75
50
25
0
0
25
50
75
100
Strong sub-lumen fibrosis (%)

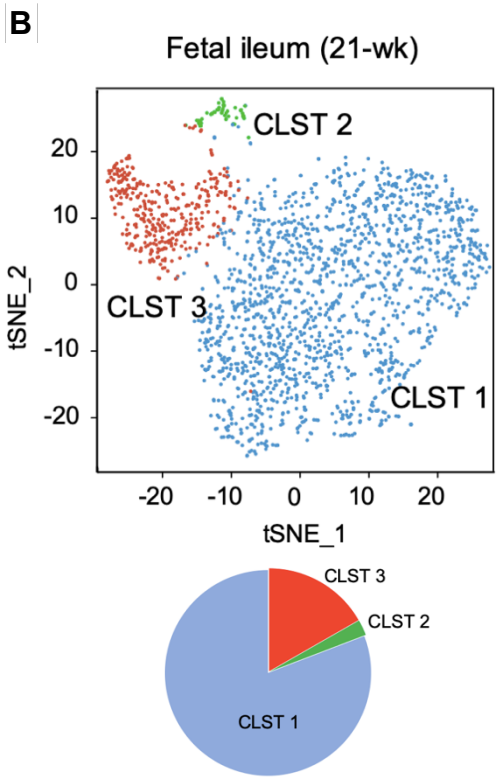
in utero without being exposed to antigens and protected from the environment and its influences (Gluckman et al., 2008). We generated stem cell libraries from terminal ileum of 21- and 22- weeks fetal demise cases (Fig. 3.14A). tSNE profile generated from single-cell RNA-sequencing data of 21-week fetal terminal ileum library showed the majority of stem cells were typical of CLST1 (75%) while remaining 15% belonged to CLST2 and 10% to CLST3 stem cells (Fig. 3.14B). CLST1 had high expression of typical small intestinal genes (GPA33, CLDN3, and FABP2), CLST2 was distinguished by a relatively high expression of CD74, INSM1, and NKX2-2, and CLST3 by TFF3, TFF1, and SPINK1 (Fig. 3.14C).

We made use of 384 well plates to sort single cells and establish clones from the fetal terminal ileum case. Whole-genome microarray expression by PCA showed that the single-cell derived three representative clones (CLST1, 2, and 3) from fetal cases grouped with representative clones from Crohn's patient SPN-29 (Fig. 3.15A). Further, heatmap and qPCR analysis of selected marker genes showed that the gene expression profiles of each variant from both fetal and SPN-29 were similar to each other which explained their grouping in PCA (Fig. 3.15B, C). Most importantly, the xenografts of these fetal clones in NSG mice behaved exactly similar to that of SPN-29 clones (Fig. 3.15D). The cross-section of CLST1 clone xenograft showed no intra- or extra-luminal cellularizations by hematoxylin and eosin (H&E) staining as well as by immunofluorescence to α-SMA and immunohistochemistry to CD45 or LY6G. CLST2 clone xenograft showed a high number of α-SMA positive myofibroblasts with no CD45 or LY6G positive neutrophils infiltration. Both fibrosis (by α-SMA positive myofibroblasts) and inflammation (by CD45 and LY6G positive immune cells) were observed in cross-sections of CLST3 clone xenograft from the fetal case.

The existence of low ratios of CLST2 and CLST3 variant stem cells in fetal and control terminal ileum and higher ratios in Crohn's terminal ileum suggested that inflammation and

Fig. 3.14. CLST 2 and CLST3 variant stem cells in fetal terminal ileum
A. Schematic for generating a library of clonogenic epithelial cells from the 21-week terminal ileum and isolation of clones for subsequent analyses. Scale bar, 100 μm. **B.** tSNE profile of scRNAseq data of stem cell library of the 21-week terminal ileum and a corresponding pie chart of the distribution of clone types. **C.** Superimposition of specific gene expression profiles across tSNE profiles of 21-week fetal ileum library. CLST1 markers include GPA33, CLDN3, and FABP2; CLST2 markers NKX2-2, INSM1, and CD74; CLST3 markers include TFF3, TFF1, and SPINK1.


A
SOX9


B
Fetal ileum (21-wk)
CLST 2
CLST 3
CLST 1
tSNE_1
tSNE_2
CLST 3
CLST 2
CLST 1

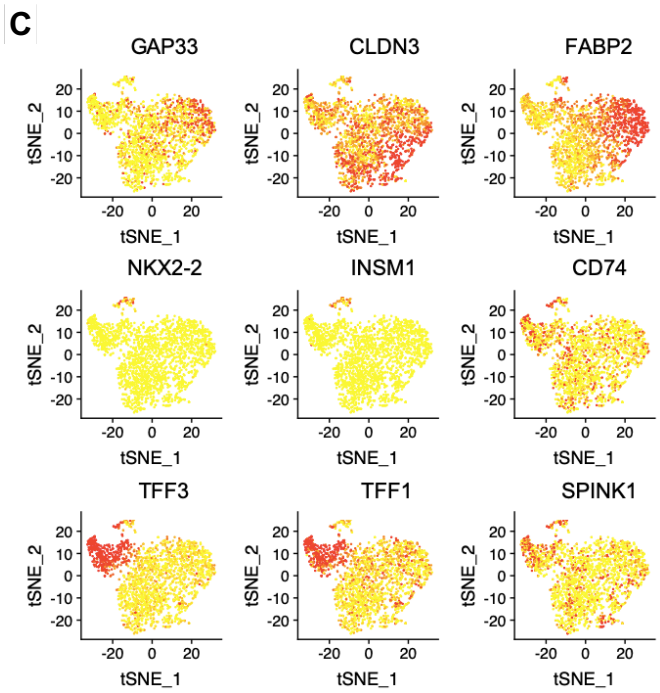
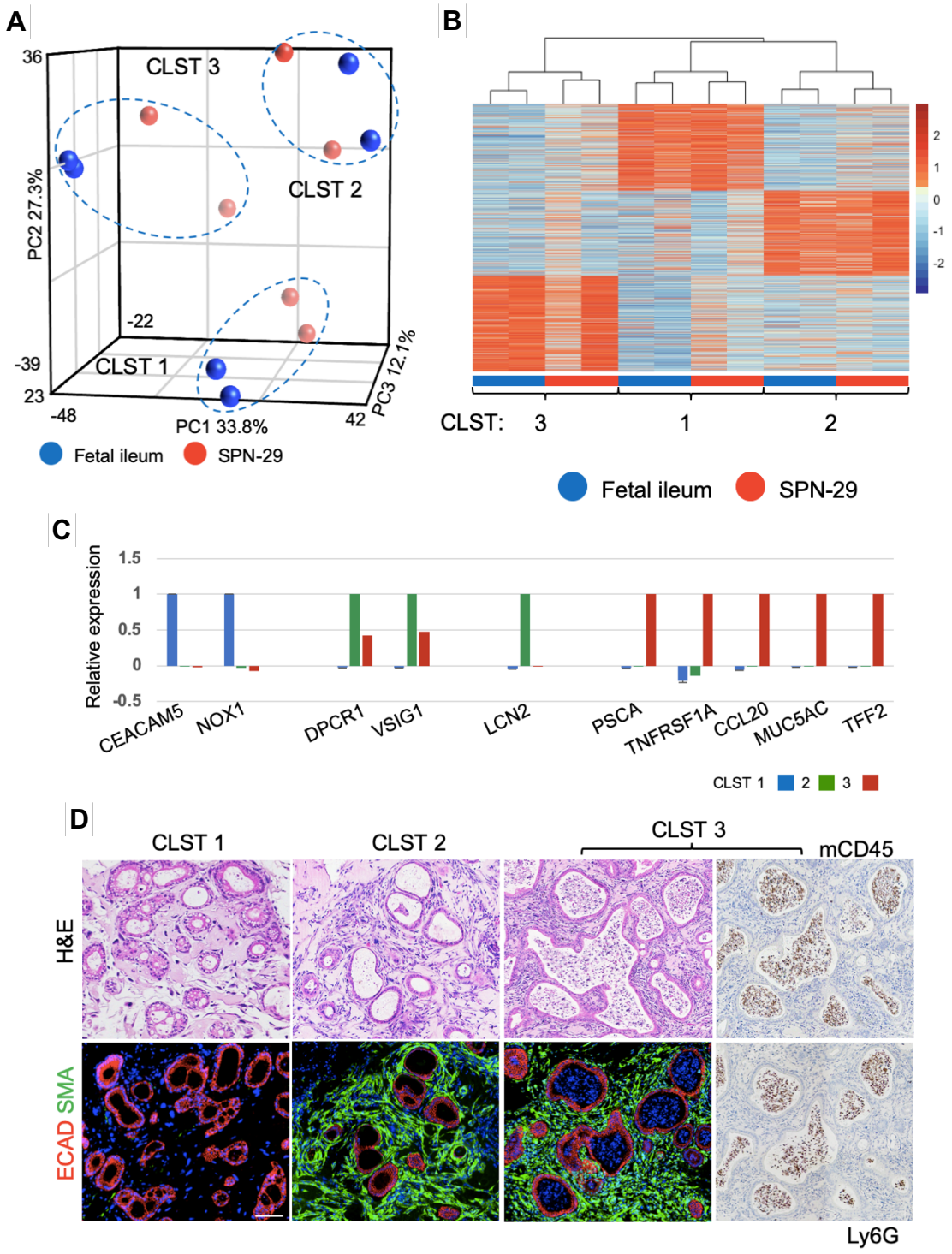

C
GAP33
CLDN3
FABP2
NKX2-2
INSM1
CD74
TFF3
TFF1
SPINK1
tSNE_1
tSNE_2

Fig. 3.15. Variant clones in fetal terminal ileum are similar to that in Crohn's
A. Principal component analysis of whole-genome expression data of nominal clone types corresponding to CLST1, CLST2, and CLST3 from both 21-week fetal ileum and from a pediatric Crohn's case (SPN-29). **B.** Expression heatmaps comparing whole-genome expression data from 21-week fetal ileum clones and those of CLST1, CLST2, and CLST3 from Crohn's case SPN-29. **C.** Quantitative RT-PCR analysis of indicated transcripts across fetal terminal ileum CLST1, CLST2, and CLST3 clones. **D.** Xenograft nodules formed by cloned fetal variants assessed by H&E staining and immunofluorescence using antibodies to ECAD (red) and α-SMA (green). Sections of CLST3 nodules are further stained by immunohistochemistry with antibodies to murine CD45 (mCD45) and LY6G as indicated. Scale bar, 100 μm.

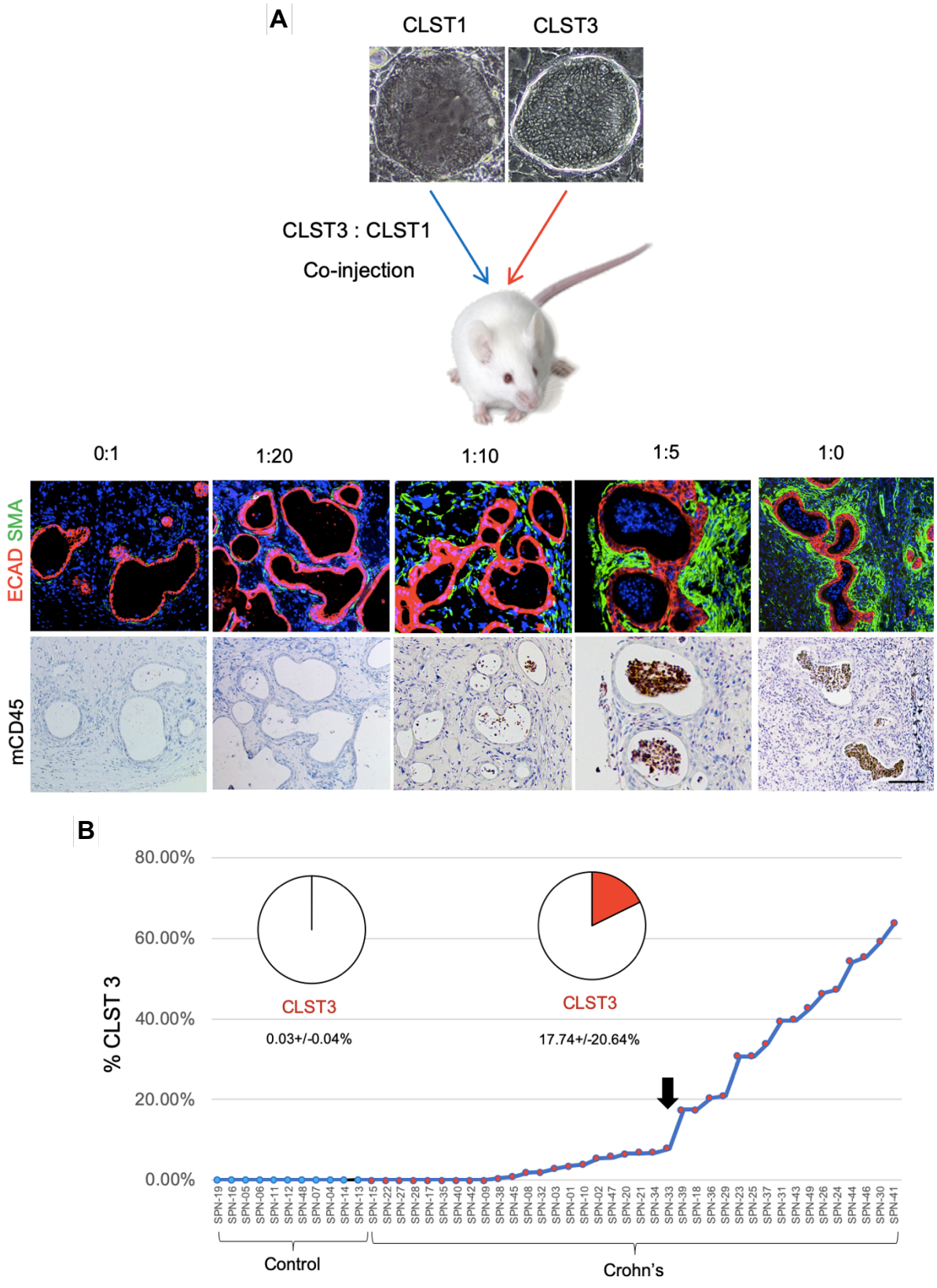

A
CLST 3
CLST 2
CLST 1
36
-22
-39
23
-48
PC1 33.8%
42
PC2 27.3%
PC3 12.1%
Fetal ileum
SPN-29
B
CLST: 3 1 2
Fetal ileum
SPN-29
C
Relative expression
1.5
1
0.5
0
-0.5
CEACAM5
NOX1
DPCR1
VSIG1
LCN2
PSCA
TNFRSF1A
CCL20
MUC5AC
TFF2
CLST 1 2 3
D
CLST 1
CLST 2
CLST 3
mCD45
H&E
ECAD SMA
Ly6G

fibrosis might simply be a manifestation of a higher number of variants. To test this hypothesis, we co-injected CLST1 and CLST3 clones from Crohn's patient SPN-29 in xenograft in various ratios (Fig. 5.3A). We observed xenograft of CLST3 clones co-injected with CLST1 clones in lower ratios such as 5% showed no obvious inflammation or fibrosis-related phenotypes. Increasing the ratio of CLST3 to 10% showed low but detectable inflammation and fibrosis while having 20% or more of CLST3 clones showed extensive inflammation and fibrosis (Fig. 3.16A). The threshold for detection of inflammation and fibrosis from this experiment was around 10-20% of CLST3 clones within a library. This range included the median percentage (between 8-18%) of CLST3 stem cells in libraries of patients with Crohn's (Fig. 3.16B, *arrow*).

3.4 Discussion

The present study suggests a deterministic role of clonogenic variant mucosal stem cells in Crohn's disease that give rise to metaplastic lesions with compromised epithelial barrier function and that drive pathological features as a host response in immunodeficient mice. The intestinal epithelial stem cell cloning technology used in this study generated both patient-derived stem cell libraries, that was representative of a 1 mm endoscopic biopsy, and individual single-cell derived clones, that were grown separately as pure "pedigrees" away from their libraries. Although single-cell RNA sequencing and FACS revealed the heterogeneity within a Crohn's library, the generation of clones was critical to corroborate their molecular and functional properties relevant to the pathology of the disease. Further, isolation and characterization of single cell-derived clones representing each stem cell population that was indicated by single-cell RNA sequencing validated the heterogeneity that was based on the genome-wide expression profile. The technology also allowed to assess their fate commitment upon differentiation and to functionally test their pro-inflammatory and pro-fibrotic gene signatures in an immunodeficient mouse in parallel which was essential to securely assign the

Fig. 3.16. Variant stem cells provoke pathology in lower ratios
A. Co-xenografts with compensating ratios of CLST3 and CLST1 clones assessed by immunofluorescence with antibodies to ECAD (red) and α-SMA (green) and by immunohistochemistry with antibodies to mCD45. Scale bars, 100 μm. **B.** Graphical representation of the percentage of CLST3 clones in stem cell libraries generated from terminal ileum biopsies of control and Crohn's cases. Arrow indicates the median percentage of CLST3 clones in Crohn's cases.


A
CLST1
CLST3
CLST3 : CLST1
Co-injection
0:1
1:20
1:10
1:5
1:0
ECAD SMA
mCD45
B
80.00%
60.00%
40.00%
20.00%
0.00%
% CLST 3
CLST3
0.03+/-0.04%
CLST3
17.74+/-20.64%
Control
Crohn's

disease-related pathology to two epithelial stem cell variants. The co-existence of normal mucosal stem cells with two pathological variants in an epithelium also explains the skip-lesion phenomenon seen in Crohn's where affected regions are separated by completely uninvolved skip areas. Moreover, the rapid and unlimited proliferative capacity of these single cell-derived clones revealed the stability of their fate commitment, molecular characteristics related to the pathology of the disease, and functional properties in an immunodeficient mouse.

The adult stem cell cloning technology used in the present study offers several advantages in contrast to pluripotent stem cells, embryonic stem cells (ESCs), and induced pluripotent stem cells (iPSCs). Unlike intestinal epithelial stem cells which are already committed to a specific lineage of columnar epithelia of the intestinal barrier, iPSCs, and ECSs demand taxing and inefficient processes for commitment to desired lineages (Medvedev et al., 2010; Sun et al., 2014). Most importantly, adult stem cell cloning allowed the discovery of regional epithelial variants. Since intestinal stem cells did not require any further processes for commitment to a specific lineage, the study of molecular and functional characteristics of those variants to link them to Crohn's disease-related pathology was possible. While adult stem cell-derived organoids for Crohn's disease may be useful in many aspects, clonal analysis to dissect and study the epithelial variants at a single cell level would be complex, labor-intensive, and challenging (Kim et al., 2020; Lancaster et al., 2019). Also, the maintenance of high clonogenicity (>70%) and long-term culture of stem cells with the present technology has a tremendous proliferative advantage over organoids (Duleba et al., 2020).

These epithelial variants were found to be dominating both pediatric and adult Crohn's cases while their age-matched control cases and fetal ileum had a minor proportion of them. This strengthens our hypotheses and suggests a common mechanism to Crohn's disease in

both pediatric and adult groups. More than 80% of the Crohn's libraries examined provoked neutrophilic inflammation and myofibroblast accumulation as a host response in immunodeficient mice and these features were highly correlated with the proportion of epithelial variants within that library. But 18% of the Crohn's libraries (7 out of 38) behaved like that of control libraries which did not show signs of inflammation or fibrosis and were present with a minor proportion of variant stem cells. Surprisingly, all the outlier libraries were cases biopsied at the time of stable clinical remission where the intestine appears normal to white light endoscopy. Given the skip-lesion phenomenon in Crohn's, it raises the possibility of having pathological features displayed in these outlier libraries if the biopsy were taken from a different region. The majority of cases examined (29 out of 38) in this study were at a state of stable clinical remission at the time of biopsy collection (Table 3.1). But 75% of those remission cases (22 out of 29) still displayed a similar proportion of variant cells correlated to inflammation and fibrosis in host animals to that of active cases (9 out of 9). This suggests the counteraction of pathogenic impacts of variant stem cells during remission induced by the use of immunosuppressants and anti-inflammatory therapeutics. This also raises important questions concerning therapeutic regimens, the mechanisms by which they induce remission, roles of latent variant stem cells in relapse, and recurrence of Crohn's following ileocolonic resection (Rutgeerts et al., 1984).

The variant stem cells displayed two different characteristics both of which solidify our argument to classify them as pathogenic for the disease. Both variants differentiated into polarized columnar epithelia that is devoid of secretory cells including goblet cells, endocrine cells, and Paneth cells. The epithelia also showed either discontinuous expression or a complete absence of tight junction-related proteins. All of these are known to play critical roles in protecting the intestine from the invading gut microbes (Wehkamp et al., 2005; Rioux et al.,

2007; Adolph et al., 2013). Loss of these components could alone leave the barrier compromised in the containment of gut microbes and expose them to underlying immune cells leading to chronic inflammation and fibrosis of Crohn's.

Another characteristic of variant stem cells was constitutive pro-inflammatory and pro-fibrotic gene expression which suggested their primary role in driving the pathology of Crohn's disease. The pathogenic molecular characteristics of two variants were further confirmed functionally by subcutaneous transplantation in immunodeficient mice. While both CLST2 and CLST3 variants induced strong recruitment of host-derived myofibroblasts, only CLST3 induced a strong trans-epithelial infiltration by host neutrophils. Despite several inflammatory pathways, including IL-17 signaling, shared between CLST2 and CLST3 stem cells, we found no inflammatory activities driven by CLST2 as a host response. The inflammatory gene signatures of CLST3 were biased towards innate immunity response including cytokine and chemokine signaling implicated in neutrophil chemotaxis such as IL-6 and CXCL8 (Fielding et al., 2008; de Oliveira et al., 2013) as well as monocyte signaling. In contrast, CLST2 showed inflammatory gene signatures biased towards adaptive immunity response including T and B cell activation and differentiation. The immunodeficient host strain, NSG, used in this study has deficiencies in hematopoietic lineages including those of B and T cell, natural killer cells, macrophages, and dendritic cells (Shultz et al., 2005; Shultz et al., 2012). Thus, IL-17 signaling pathways seen in both variants, T and B cell activation signaling seen in CLST2, and monocyte signaling seen in CLST3 should not be underestimated. Other strains of humanized mice with more hematopoietic lineages could show profound inflammatory responses to CLST2 and CLST3.

Among inflammatory genes differentially expressed between normal and the two variants were also the genes that have been previously linked to IBD. The signaling pathways

involving IL-1, IL-6, IL-12, IL-17, IL-18, and oncostatin M have been tied to play important roles in IBD before (Kaser et al., 2010; Elinav et al., 2011; Khor et al., 2011). Similarly, there was a stunning overlap of 75 up-regulated and down-regulated genes between those implicated by GWAS in the risk for Crohn's and those differentially expressed by the CLST2 and CLST3 cells. This overlap of constitutive gene expression with GWAS and with genes implicated in IBD support that these variants may be actively contributing to Crohn's pathology.

We also reported the existence of these variant stem cells in control libraries as well as 21- and 22-week fetal terminal ileum libraries maintained at lower proportions without inducing an inflammatory or fibrotic response in host NSG mice. Upon isolation from their libraries, they displayed similar pro-inflammatory and pro-fibrotic gene expression profiles to that of Crohn's variants and also induced pathological activities in xenografts. Moreover, in a co-injection experiment with various ratios of CLST3 to CLST1, we showed 20% of CLST3 variant is enough to provoke extensive inflammation and fibrosis in host mice. This suggests the expansion of CLST2 and CLST3 variants in an individual due to which pathology of the disease is promoted.

The tissue enrichment profile of CLST2 and CLST3 differentiated epithelia showed similarity to gastric epithelia rather than ileal or colonic epithelia. Further, the reaction of epithelia to antibodies of gastric marker genes, such as VSIG1, CLDN18, LCN2, MUC5AC, and many more, validated the gene expression profile and assigned these variants as the source of gastric metaplasia often seen in ulcer associated regions in Crohn's (Liber, 1951; Lee, 1964; Kariv et al., 2010; Agarwal et al., 2013; Thorsvik et al., 2019). But such gastric metaplasia has been argued to function as repair cells in an ongoing state of inflammation (Wright et al., 1990). Nevertheless, in the present study, we showed gastric metaplasia rise from clonogenic epithelial variants and tied them to two pathological features relevant to the disease.

We observed that the pro-inflammatory and pro-fibrotic gene signatures underlying CLST2 and CLST3 were also similar to that of gastric epithelia. In accompanying research by our lab, we made use of previously cloned stem cells from 21- week fetal proximal and distal gastrointestinal tract. These stem cells from different regions of the gastrointestinal tract are epigenetically committed to differentiate to their respective mucosa from which they were derived (Wang et al., 2015). A mapping between inflammation-related genes differentially expressed among two variants from Crohn's and genes differentially expressed among gastric epithelia from fetus revealed a striking overlap (data not shown). This finding suggests that both Crohn's variants and gastric epithelia might play a role as sentinels to prevent the colonization of pathogens by eliciting an immune response against them and that Crohn's may simply be an overrepresentation of variant stem cells. In a recent publication, Rao et al., showed epithelial variants present in minor ratios in controls are drivers of key pathological features of chronic obstructive pulmonary disease (COPD), a condition in the lung (Rao et al., 2020). The research strengthens and supports our finding in Crohn's and argues for a similar etiology for chronic diseases.

The present study comes with some limitations. While Crohn's has regional presentations in the colon, rectum, or more proximal parts of the gastrointestinal tract (Farmer et al., 1975; Silverberg et al., 2005; Levine et al., 2011; Cleynen et al., 2016; Kotze et al., 2017), the research was focused solely on terminal ileum. To assess the nature of Crohn's disease in different segments of the gastrointestinal tract and to investigate the role of variant stem cells in those regions, more studies that are similar to the current one will be necessary. Currently, the study also does not provide any mechanism for the expansion of these variants. Genetics and adverse gastrointestinal events may selectively favor such expansions. But a clinically relevant animal model is necessary to link the variant epithelial cells to causes and symptoms of IBD.

If these clones truly contribute to Crohn's disease, then several strategies can be employed to mitigate the disease. We have related two major pathologies of Crohn's, inflammation, and fibrosis, to an array of cytokines, chemokines, and other relevant genes constitutively expressed by two epithelial variants. Those pathogenically relevant factors could be neutralized to lessen their control on disease progression. Another approach could be targeting the variants specifically with cell surface directed antibodies or through small molecules that selectively target these variant stem cells allowing normal mucosal stem cells to repopulate and maintain homeostasis.

CHAPTER 4: DRUG DISCOVERY USING PATIENT-DERIVED STEM CELLS

4.1. Significance of the proposed study

The current approach to treating IBD including Crohn's is by the use of anti-inflammatory, corticosteroids, immune-suppressors, immune-modulators, and anti-TNFs directed towards alleviating immediate symptoms of the inflammation. The treatment is based on the assumption that dysregulated immune system, microbiome, or environmental factors are the trigger of this disease (McGovern et al., 2015). A large number of patient investigations (Buisine et al., 1999) and genetic knockout animal models (Park et al., 2009; Williams et al., 2015) along with GWAS (Franke et al., 2010) highlight the possibility of the dysfunctional epithelial barrier being a key player of the disease. Despite a good correlation between the disorder and barrier defects, it is unclear whether these defects play a deterministic role in driving the inflammation or are simply secondary consequences of the inflammatory state of this disease. Our cloning and characterization of patient-derived epithelial stem cells support the first hypothesis.

A total of 31 out of 38 Crohn's patient stem cell libraries were dominated by two variant mucosal stem cell populations co-existing with normal mucosal stem cell populations. Our studies showed that these variant stem cells are the source of both gastric metaplasia and dysfunctional epithelial barrier characterized by loss of mucins, anti-microbial peptides, and tight junction related proteins often seen in Crohn's. Together, they present a compromised barrier that exposes the underlying tissue and immune cells to the microbiome in the gut. Most importantly, we showed that these variants are also marked by constitutive inflammatory gene signatures including those associated with inflammatory bowel disease and with the response to multiple human pathogens. Consistently, xenografts of these variant cells in highly

immunodeficient mice drive host inflammatory and fibrotic responses which are two major pathological features seen in Crohn's. Moreover, the existence of these variants in control biopsies and in the fetal terminal ileum in minor proportion without provoking pathological features showed that Crohn's disease is the result of an overrepresentation of intrinsic cells that normally function to protect the gastrointestinal tract from human pathogens.

Therapeutics directed at the immune system might suppress the inflammatory state for the time being but certainly does not eradicate those pathogenic cells. In this context, the other emerging treatment options such as immune cell therapy, hematopoietic stem cell therapy, and mesenchymal stem cell therapy do not hold a suitable rationale. On the other hand, selectively targeting these pathogenic variants might create an opportunity for existing normal mucosal cells to repopulate the intestine and maintain homeostasis.

4.2. Research strategy

4.2.1 Phenotypic approach to drug discovery

Ever since advancements made on genomics, target-based methods have been the focus of drug discovery (Lindsay, 2003). The method includes the identification of a target which is typically a single gene, protein, or molecular mechanism that has an important role in disease pathology via genetic biological approach such as RNA interference methods. It is followed by target validation, screening for hits, and lead compound optimization. The development of advanced genomic (RNA interference) and structure (X-ray crystallography) based tools have tremendously helped identification and optimization for targets (Imming et al., 2006). Along with these, computational modeling and virtual screening methods have also made the discovery process economical and fast (Overington et al., 2006). The target-based approach allows

applying chemical and molecular knowledge to investigate specific compounds, biologically based approaches, and high throughput small molecule screening strategies. Despite these advantages, this approach has limitations due to challenges remaining in research and development productivity and due to the requirement of knowledge of molecular mechanism (Williams, 2005). Also, a solution to the molecular hypothesis might not be relevant to polygenic disease and its pathogenesis (Flordellis et al., 2006).

On the other hand, the phenotypic approach does not require prior knowledge of disease mechanisms as well as recognition of disease-relevant targets. Compared to the target-based approach, the readout from the assays can be directly translated into therapeutic impact in a disease. Phenotypic assays were primarily the drivers of drug discovery before the genomic era and target-based approaches (Williams, 2005). This approach also comes with challenges such as optimization of candidate drugs without the knowledge of the molecular mechanism of action and incorporating new screening technologies. Nonetheless, the success rate of discovery and development of innovative drugs is still higher with the phenotypic approach compared to the target-based approach (Swinney and Anthony, 2011).

The phenotypic approach is especially effective in cancer drug discovery where the main aim is to either eradicate the malignant cells or to reprogram them minimizing cytotoxicity to normal cells. Cancer-derived cell lines that show phenotype related to uncontrolled growth have been the basis for phenotypic screening for anticancer drugs as well as validation in vivo using humanized mouse models. High throughput screening on these cancer cell lines with a phenotypic readout, such as survivability or reporter signal, has proven itself as an effective method to drug discovery (Moffat et al., 2014).

4.2.2 High throughput screening in the hit discovery process

A ‘hit’ is defined as a molecule which gives anticipated activity during screening and whose activity is reiterated upon testing. The desired compound screening assays, such as reporter assays or drug-induced cell death or apoptosis, are developed during the hit identification phase. Several screening strategies are available to discover hit compounds such as high throughput screening, focused screening, fragment screening, and many more (Hughes et al., 2011). Focused and fragment screening requires knowledge about target protein that is likely to be involved in the mechanism of disease as well as knowledge about the smaller subset of molecules to be screened (Boppana et al., 2009). On the other hand, high throughput screening (HTS) involves directly screening the target in an entire compound library. Cell-based assays, such as using cancer cell lines, mostly use this approach where it might require secondary assays to confirm the action of compounds (apoptosis or change in expression of some genes) (Fox et al., 2006). While this screening approach requires automation of complex laboratory equipment, it does not require any prior knowledge of the nature of the compound that is likely to have the desired activity.

High throughput screening (HTS) involves screening a large number of compounds against the target in an assay designed to run in multi-well plates of generally 384 wells or above. Because of its ability to screen hundreds of thousands of molecules in a large number of chemical libraries (related to genomics, protein, peptide, small, and large molecule libraries), it accelerates the hit identification process playing an important role in early drug development. The chemical libraries are a diverse collection of unique small molecules that are focused on a huge spectrum of targets over many signaling pathways (Hughes et al., 2011). HTS also aids in the early exclusion of unsuitable compounds by providing qualitative and quantitative readouts (Martis et al., 2011). Moreover, advancement in technology over the past

decade has supported the growth of HTS as well. The creation of novel platforms, miniaturization of assay plates, automation of fluid handling, and robotics have minimized costs and eliminated labor-intensive steps. High content imaging, the development of advanced database management software, and advances in analytical tools to analyze massive quantities of data have dramatically reduced assay times (Szymański et al., 2012). Thus, high throughput screening is in high demand for drug discovery today.

4.2.3 Cell-based assay development and high content screening

Mostly, cell-based assays have been applied in a phenotypic approach to drug discovery to generate the desired readout as a result of the compound activity (Michelini et al., 2010). The current trend in cell-based assays for drug discovery is the establishment of stable mammalian cell lines to overexpress the target protein of interest or establishment of cancer cell lines which serve as target themselves (Hughes et al., 2011). The use of primary cell lines for compound screening has also increased in recent years (Dunne et al., 2009). Cell-based screening approaches give more complexity and more information compared to simple biochemical assays. The assay gives information on the penetration of compounds through the cell membrane, build an acute cytotoxicity profile, and identifies drugs against a target of an unknown mechanism. The ease of availability of cell lines from different organs and different species has contributed to assess toxicity that may be more reliable in extending tests to animals (Houck and Kavlock, 2008). Further, evaluation of cell signaling and stress-response pathways in addition to classical membrane leakage, ATP utilization, and cell number count techniques have added to the understanding of mechanisms of cytotoxicity (Szymański et al., 2012).

A number of cell-based assays such as second messenger assays, reporter gene

assays, cell proliferation assays, and high content screening (HCS) assays exist. High content screening (HCS) offers measurement of multiple parameters using fluorescence-based reagents to reveal information about cytoskeletal rearrangements, nuclear size and shape changes, mitochondrial potential, and most importantly drug-induced programmed cell death. In combination with high content imaging (HCI), it has become an indispensable tool to quantitatively analyze cellular phenotype assays (Giuliano et al., 1997).

4.2.4 Primary and stem cell lines for high throughput screening

Most cell-based screening assays for drug discovery have used cancer-cell lines, other immortalized cell lines such as CHO or HEK 293, or non-human immortalized cell lines such as 3T3 or Vero to overexpress the target protein of interest at high levels to screen for compounds against the target. These have been successfully working for years in high throughput screening to identify hit compounds. However, lead optimization requires testing in a physiologically similar environment to that of *in vivo* for pharmacological assessment of candidate molecules to reduce failure of compounds in clinical trials especially with assays involving non-human cell lines (Hughes et al., 2011). Hence, the field is recognizing the value of primary cell lines which retain numerous functions seen *in vivo* as well as endogenously express the target of interest (Eglen et al., 2008). Primary cells and stem cells, either embryonic or adult, are beginning to secure a place in drug discovery now.

Primary cell lines are derived either from embryonic tissues, such as neuronal and cardiomyocyte cultures, or from adult tissues, such as hepatocytes and pituitary cells. The cells are usually derived from animals although human cells are now being routinely employed for preclinical drug testing (Eglen et al., 2008). Advances in cell culture have established cell lines from various organs, including brain, pancreas, liver, heart, intestine, pituitary, ovary, and more

organs and employed for screening. Their ability to be cryopreserved and cultured for a larger number of passages *in vitro* has made it easy to adapt them for screening. Human hepatocytes and myocytes are perhaps the most widely used cell lines to assess cytotoxicity and the effect of drugs in metabolism (Nolan, 2007). Pituitary cells have been studied for understanding basic mechanisms regulating hormone secretion as well as in identifying hypothalamic features governing pituitary function. Similarly, neuronal cells have been used to study response towards drugs as assessed by firing frequency and ionic conductance (Daub et al., 2009).

Stem cells provide an unlimited source of cells from various tissues that reflects the cell properties *in vivo* more accurately giving phenotypes that are more clinically relevant. They also allow more precise modeling of human cells as well as a disease when obtained from patients with different diseases. An extensive amount of work has been done to employ both embryonic stem cells (ES) and induced pluripotent stem cells (iPSCs) in drug screening (Primary). For example, cardiomyocytes derived from both mouse and human embryonic stem cells (ES) were used for compound screening to identify cardioselective drugs that simulated electrical pacing activity and contractility (McNeish, 2007; Cezar, 2007). Somatic multipotent stem cells such as those from mesenchyme, blood, brain, and adipose tissue have been used as tools for compound screening in wound healing and therapeutics (Nirmalanandhan and Sittampalam, 2009). For example, neural stem cells (NCS) that differentiate into neurons play an important role in patients suffering from neurodegenerative diseases. Chemical screening in NCS to discover a compound that could promote neurogenesis would be beneficial for those patients (Rishton et al., 2008). Thus, stem cells can act as authentic human cells that are directly derived from patients acting as a crucial tool in high throughput screening for drug discovery.

4.2.5 Animal model for *in vivo* validation of drugs

4.2.5.1 Limitations of current animal models

While chemically induced models are easy and rapid to develop, they provide limited information regards to how chronic disease perpetuates in humans. These models are limited to the study of tissue regeneration, neutrophil infiltration, understanding alleviation of inflammation, and addressing physiology of other aspects of acute flares. Immunological models portray the role of adaptive immunity, regulatory and pathogenic T cells, but fail to investigate innate immunity as hosts possess profound immune disorders. Genetic models reveal the roles of a single gene or a group of genes in the regulation of mucosal immunity. However, enforced genetic manipulations are unlikely to represent underlying defects, explain causative factors, and etiology of the disease. Spontaneous models do not require any extra manipulations to induce inflammation like in human disease, but it tends to resolve by itself with age without recurrence which is not the case in human disease.

Each mouse model described under chemically induced, immunologically mediated, genetically manipulated, and spontaneous has provided valuable insights in understanding one or another major aspect of chronic intestinal inflammation and the underlying mechanism. These models have led to the development of generally accepted principles of chronic intestinal inflammation which has some similarities to human IBD. Yet, no single model has captured all the aspects of human IBD. Furthermore, the mouse models described above do not involve the small intestine like Crohn's disease, so they are more likely to model colitis. Hence, it has fueled current researchers to develop a model that closely resembles disease in terms of location, pathophysiological features, and development of therapy.

4.2.5.2 Humanized mouse model for *in vivo* validation of drugs

Humanized mouse models are created by engraftment of in vitro cultured human cells, human cell suspension, or tissues in an immunocompromised mouse that does not reject the xenografts. The discovery and development of highly immunocompromised mice such as nude, NOG, BRG, or NSG mice allowed humanized mouse models to advance (Bosma et al., 1983; Shultz et al., 1995). These mice generally have impairments in B and T cell function and lack natural killer (NK) cells. The different strains of immunodeficient mice exist with a mutation in the *Il2rg* gene along with other genetic manipulation and are specific in addressing a certain question. Thus, choosing an appropriate model and designing proper experiments plays a critical role in providing valid data.

Humanized mouse models have offered direct translational research on human diseases which was previously not possible on immunocompetent mice. One of the most popular humanized mouse model examples is the engraftment of human hematopoietic stem cells (HSC) and generation of well-differentiated lineages of human myeloid and lymphoid system in an irradiated NOD-SCID IL2Rγnull mouse (Shultz et al., 2005). Such hematolymphoid humanized mice have also allowed the study of infectious diseases such as HIV-1, HTLV-1, and EBV (Zhang et al., 2010). The mice provide a system to study viruses that specifically target lymphocytes and study disease mechanisms along with the development of drugs. Immunodeficient mice are also used to establish various cancer models that investigate the development of tumor, its interactions with various types of host cells as well as an immune response to tumor (Ninomiya et al., 2004; Machida et al., 2009; Shiokawa et al., 2010). In addition, it was demonstrated that graft versus host disease (GVHD) induction was possible in humanized mouse models allowing the study of this severe complication. The transplanted human T cells along with hPBMCs were activated and attacked the host mouse tissues giving

graft versus host disease-like symptoms (Mosier et al., 1988).

Recently, the patient-derived xenograft (PDX) model has also gained traction in basic and translational cancer research over the last decade. This model is created by direct subcutaneous or orthotopic engraftment of patient-derived tumor tissues, bypassing any *in vitro* work, in an immunocompromised mouse. Importantly, the features of the patient tumor such as gene expression profile and drug response are retained by the model. Moreover, the model maintains heterogeneity and microenvironment that comes with a tumor as well as allows interaction with various cell types such as fibroblast, blood vessels, and immune cells in the host. Hence, the PDX model has been considered as an advanced preclinical cancer model and as a valuable tool to assess personalized medicine (Siolas et al., 2013).

4.3 Results

4.3.1 Adapting stem cells for high throughput screening

Three variant clones representative of CLST 1, 2, and 3 from Crohn's patient SPN-29, CLST1 and CLST2 clones from patient SPN-49, CLST2 clone from patient SPN-20, CLST3 clone from patient SPN-25, and CLST1 clone from control patient SPN-19 were subjected to high throughput screening (HTS) on 1 micromolar (μM) chemical libraries. Each well of a clonal line, labeled with green fluorescent protein (GFP), after treatment with a chemical library was represented in terms of survival rate as a percentage of untreated wells (Fig. 4.1A). In general, a threshold was set for the hit from a small molecule to be at 80% or lower survival rate. The selection of small molecule libraries included Custom Clinical Library, Prestwick Chemical Library, Selleck Bioactive Chemical Library, and UT Austin Kinase Inhibitors Library with a total of more than 4000 compounds and around 2000 unique compounds (Fig. 4.1B).

A plot between the number of compounds and their effect represented as survival rate from entire clonal lines after treatment with four chemical libraries showed that 70% of chemicals were ineffective against the clones as the histogram is skewed towards higher survival rate (Fig. 4.1C). We also observed some compounds having a promotion effect on some clonal lines, the threshold for which was set at 120% or higher survival rate (Fig. 4.1C). An assessment of small molecule targets from all the hit compounds revealed the target pathways such as *Angiogenesis, Cell Cycle, JAK/STAT, NF-kB, TGF-B/SMAD, GPCR, PI3K/AKT/mTOR, Apoptosis, Cytoskeleton Signaling, Metabolism,* and many more (data not shown). Many of these pathways such as *JAK/STAT, NF-kB, TGF-B/SMAD, PI3K/mTOR,* and *Metabolism* have been implicated in Crohn's and Ulcerative Colitis (Coskun et al., 2013; Atreya et al., 2008; Sedda et al., 2015; Tokuhira et al., 2015).

4.3.2 HTS allows the discovery of drugs against variant stem cells

A total of 72 common hits from all CLST2 and CLST3 clones across patients and all four libraries were found, whose target enriched pathways/proteins were among *PI3K/AKT/mTOR, HSP, Protein Tyrosine Kinase,* and *Cytoskeletal Signaling* (Fig. 4.2A). We used the median survival rate of CLST1 clones against CLST2 and CLST3 clones from all the patients in screening and constructed a multi-patient normalized scatter plot for each library (Fig. 4.2B). From the plot, we observed a number of candidate small molecules which seemed to inhibit CLST2 and CLST3 clones more compared to CLST1 clones. Tanespimycin, a HSP90 inhibitor, was one of such compounds (Fig. 4.2C). Further, the selection of compounds for validation was based on two criteria - at least 50% inhibition of CLST2 and CLST3 clones and at least 2-fold lesser effect on CLST1 clones (Fig. 4.2B, C).

Fig. 4.1. Stem cells adapted for high throughput screening
A. *Top*, A 384 well plate showing a single cell-derived clone labeled by green fluorescent protein (GFP) from control or Crohn's patients stem cell libraries after treatment with a chemical library at 1uM in high throughput screening. The wells treated with Positive control, negative control, and various library compounds are indicated in the plate. *Bottom*, Heatmap representing survival rate of clones in each well of a 384 well plate after treatment with a chemical library at 1 μM. Scale bars, 200 μm. **B.** Four chemical libraries – custom clinical, Prestwick chemical library, Selleck bioactive chemical library, and UT Austin kinase inhibitors library selected for high throughput screening showing a number of unique and overlapping compounds among them. **C.** Color histogram showing a plot between the number of compounds and their effect represented as survival rate from entire clonal lines after treatment with four chemical libraries.

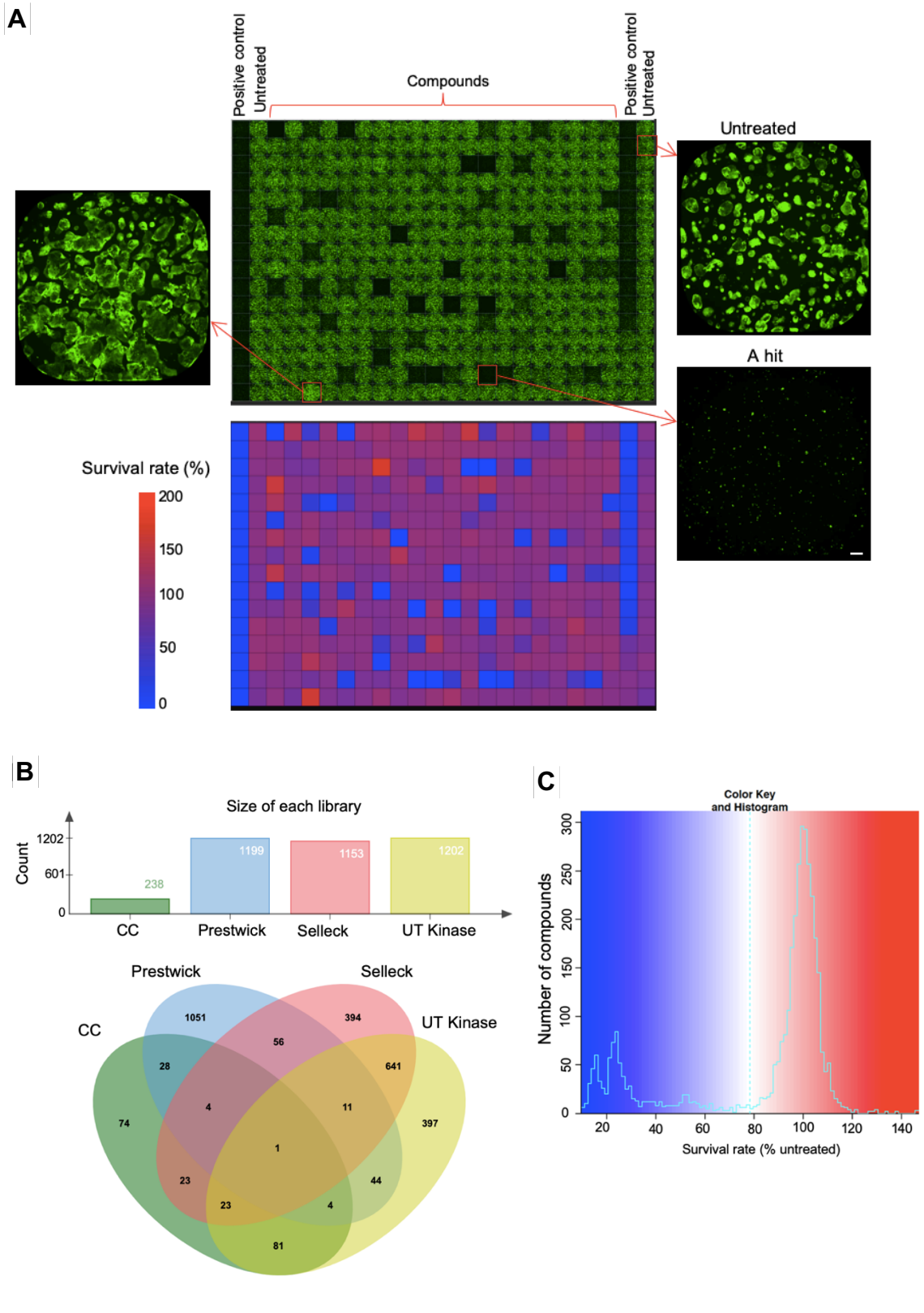

A
Positive control
Untreated
Compounds
Positive control
Untreated
Untreated
A hit
Survival rate (%)
200
150
100
50
0
B
Size of each library
Count
1202
601
0
238
1199
1153
1202
CC
Prestwick
Selleck
UT Kinase
Prestwick
Selleck
CC
UT Kinase
1051
394
56
28
641
4
11
74
397
1
23
44
23
4
81
C
Color Key
and Histogram
Number of compounds
300
250
200
150
100
50
0
20
40
60
80
100
120
140
Survival rate (% untreated)

Fig. 4.2. Small molecules specifically target variant clones
A. *Top*, A Venn diagram showing unique and common hits from four chemical libraries across CLST2 and CLST3 variant clones from four Crohn's patients. *Bottom*, Enriched targets or pathways from the hits that are common across all the variant clones screened. **B.** A scatter plot showing the median survival rate from all CLST1 clones compared to the median survival rate from all CLST2 and CLST3 clones screened in the Selleck library at 1 μM. **C.** A small molecule circled in scatter plot from ***B*** showing differential effect in variant clones compared to CLST1 clone. Scale bars, 200 μm.

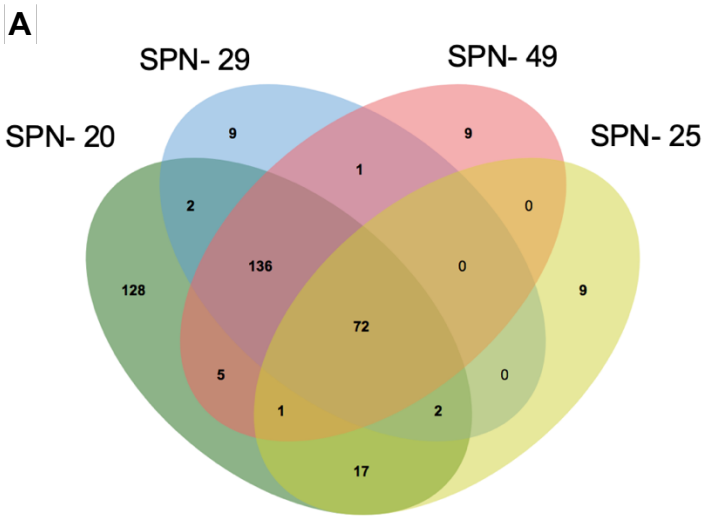
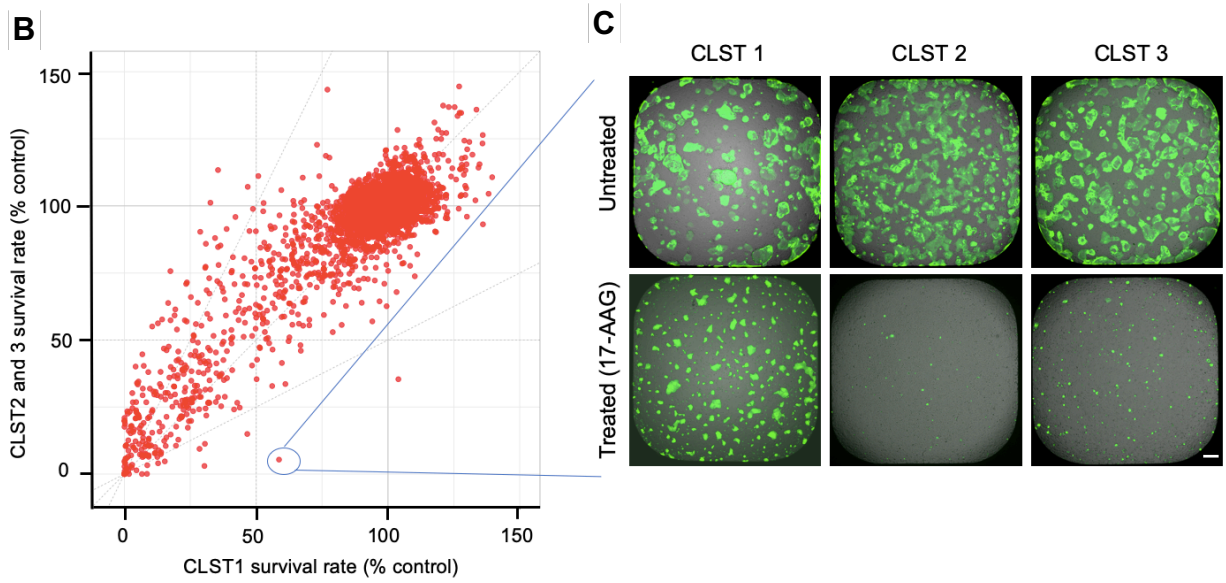


Table 4.1. Target or pathway enrichment from common hits

Pathway/Protein	p-value	q-value	Number of compounds
Cytoskeletal Signaling	4.65E-07	5.35E-06	21
Protein Tyrosine Kinase	0.00072378	0.0033294	20
PI3K/Akt/mTOR	0.00609694	0.02337161	24
HSP	0.018496	0.086375	7



From a total of 4000 compounds in HTS, only nine compounds were selected for dose-response and validation *in vitro* (Fig 4.3A, B). Among them were four compounds (Alvespimycin, Tanespimycin, Luminespib, and KW-2478) that target the HSP protein, three compounds (Gedatolisib, INK 128, and Omipalisib) that target the *PI3K/Akt/mTOR* pathway, one compound (TTNPB) that targets the retinoic acid receptor protein, and one compound (ONX 0914) that targets immunoproteasome (Fig. 4.3B).

The dosage response curve for the selected four compounds (one representing each target protein or pathway) in three representative clones from multiple Crohn's patients showed at least 5-fold differential inhibition in CLST2 and CLST3 clones compared to CLST1 clones (Fig. 4.3A, *left*). Next, we tested these compounds in a co-culture of representative CLST1 to CLST2 clones and CLST1 to CLST3 clones choosing the concentration at which maximum differential was seen (Fig. 4.3A, *right*). The drugs added to the culture two days post-seeding of clones inhibited CLST2 and CLST3 significantly allowing CLST1 to take over the culture by five days. Red fluorescent protein (RFP) tagged CLST1 and green fluorescent protein (GFP) tagged CLST2 and CLST3 were used to assess the co-culture treatment (Fig. 4.3A, *right*). Since we used a single representative clone from each variant in high throughput screening (HTS) and dosage response, the question remained whether the compound showed comparable results in the cell populations representing three clusters. To resolve this, we took the same effective concentration of compounds that worked in the co-culture experiment and treated two Crohn's patient libraries on the second day post-seeding. Fluorescent Activated Cell Sorting (FACS) to CEACAM5 (CLST1 marker) and VSIG1 (CLST2 and CLST3 marker) 5 days post-treatment with two selected drugs showed that only CLST1 stem cells (CEACAM5 positive and VSIG1 negative) survived the treatment (Fig. 4.3C).

Fig. 4.3. Validation of selected small molecules from HTS
A. *Left*, Dosage response curve from treatment of representative CLST1, 2 and 3 clones with four compounds (Tanespimycin, Gedatolisib, ONX-0914, and TTNPB) out of nine selected for validation *in vitro* in B. Each compound represents different target protein/pathway listed in B. *Right*, Co-culture of CLST1 with CLST2 and of CLST1 with CLST3 treated with the same compounds. CLST1 was genetically labeled using retrovirus carrying red fluorescent protein (RFP). CLST2 and CLST3 were genetically labeled using retrovirus carrying green fluorescent protein (GFP). Arrows indicate the concentration of each compound used in the co-culture. Scale bars, 100 μm. **B.** List of nine compounds selected for validation and dose response curve *in vitro* from high throughput screening. **C.** Fluorescence-activated cell sorting (FACS) profiles of Crohn's patient SPN-49 stem cell library stained by antibodies to the CLST1-specific marker CEACAM5 and the CLST2+CLST3 marker VSIG1 before and after the treatment with Tanespimycin and Gedatolisib with a concentration that was used in A.

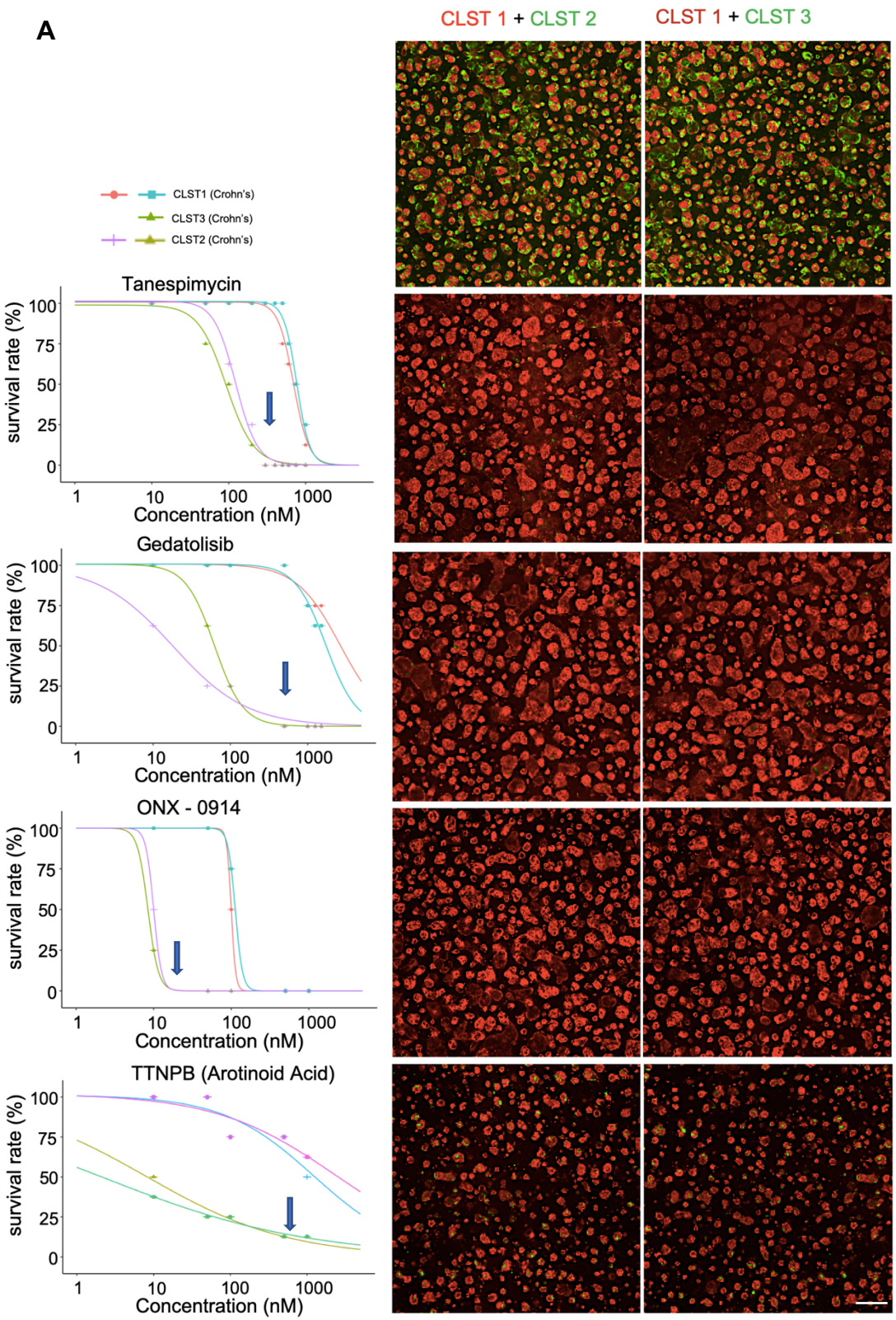

A
CLST 1 + CLST 2
CLST 1 + CLST 3
CLST1 (Crohn's)
CLST3 (Crohn's)
CLST2 (Crohn's)
Tanespimycin
Gedatolisib
ONX - 0914
TTNPB (Arotinoid Acid)
survival rate (%)
Concentration (nM)
100
75
50
25
0
1
10
100
1000

Table 4.2 Compounds selected for validation *in vitro*

B

Compounds for validation	Target	Clinical Trial Stage
17-DMAG (Alvespimycin)	HSP90	Phase I
17-AAG (Tanespimycin)	HSP90	Phase II
Luminespib (AUY-922)	HSP90	Phase II
KW-2478	HSP	Phase I
INK 128 (MLN0128)	mTOR	Phase I
Omipalisib (GSK2126458)	mTOR	Phase I
Gedatolisib (PF-05212384)	mTOR/PI3K	Phase I
TTNPB (Arotinoid Acid)	RAR	
ONX-0914 (PR-957)	Immunoproteasome	

C

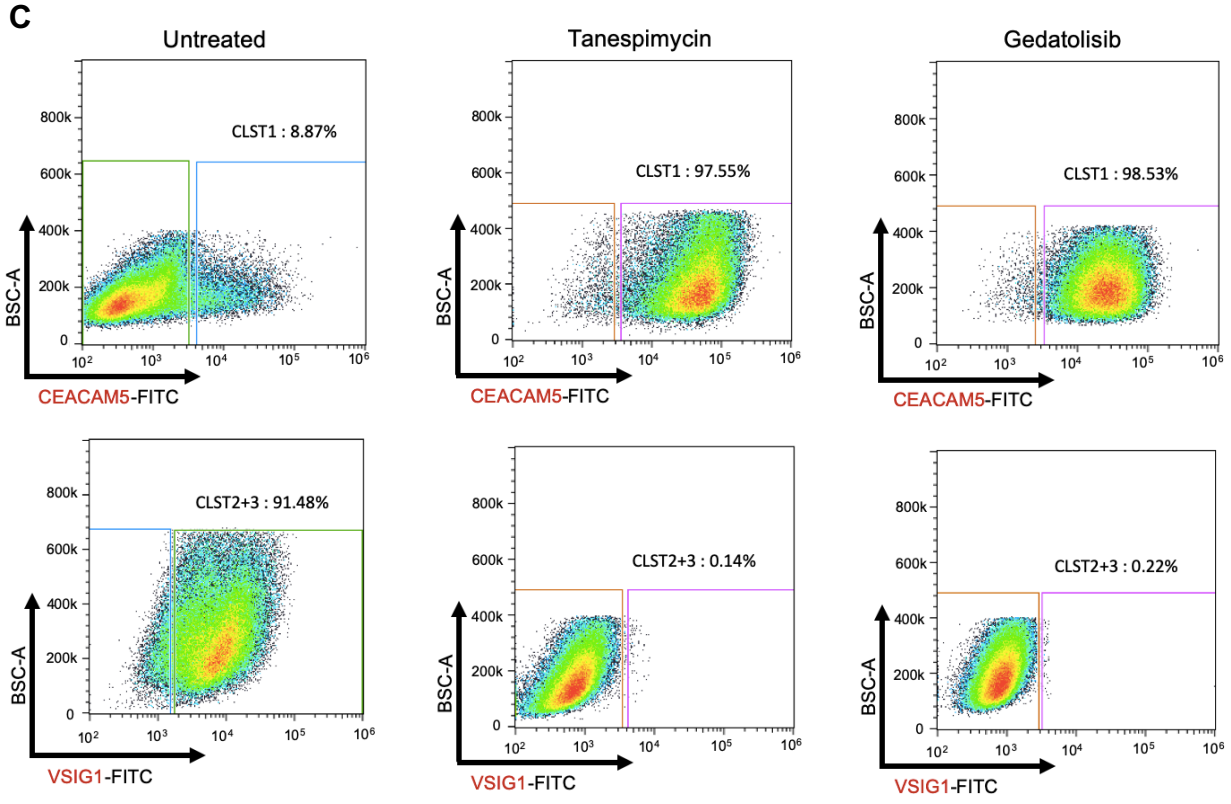


Fig. 4.3. (cont.)

4.3.3 HTS allows the discovery of promoters

Tanespimycin, an HSP90 inhibitor, and Gedatolisib, an mTOR/PI3K inhibitor, were selected as the best candidates out of nine drugs screened for validation. The selection was based on the maximum differential seen among control and two variants and on the number of variant clones from multiple Crohn's patients the drugs showed inhibition. An important aspect of drug discovery is finding compounds that can work synergistically with each other allowing the use of lower concentrations in combination to eradicate the target. This is crucial in minimizing cytotoxicity and side effects to the normal population of cells and organs nearby (Ashburn and Thor, 2004; Jia et al., 2009). We attempted to discover compounds that could work synergistically with two selected CLST2 and CLST3 inhibitors. Representative clones from each variant treated with a sub-lethal dose of Tanespimycin were subjected to Selleck chemical library treatment at 1 micromolar (μM) (Fig.4.4A, *right*). Surprisingly, the scatter plot showed several compounds that promoted CLST1 clones while CLST2 and CLST3 clones were still inhibited. Such compounds favored all three clones in the absence of Tanespimycin before (Fig. 4.4A, *left*).

To validate this observation made in high throughput screening (HTS), we chose eight compounds to combine with two inhibitors of CLST2 and CLST3. The targets of such compounds were *Focal Adhesion Kinase, PI3K/AKT, VEGFR, JAK-STAT1/2, Phosphatase, beta-secretase*, and *IGF-1R* (Fig. 4.4B). The dose-response curve generated for Tanespimycin and Gedatolisib in presence of those eight compounds at 1 micromolar (μM) in three variant clones validated our finding from the screening (data not shown). These eight compounds in presence of CLST2 and CLST3 inhibitors significantly increased the survivability of CLST1 clones yet keeping the dose-response of variant clones unchanged in a manner similar to that of before adding the compounds (Fig. 4.4C). Although these compounds greatly reduced the

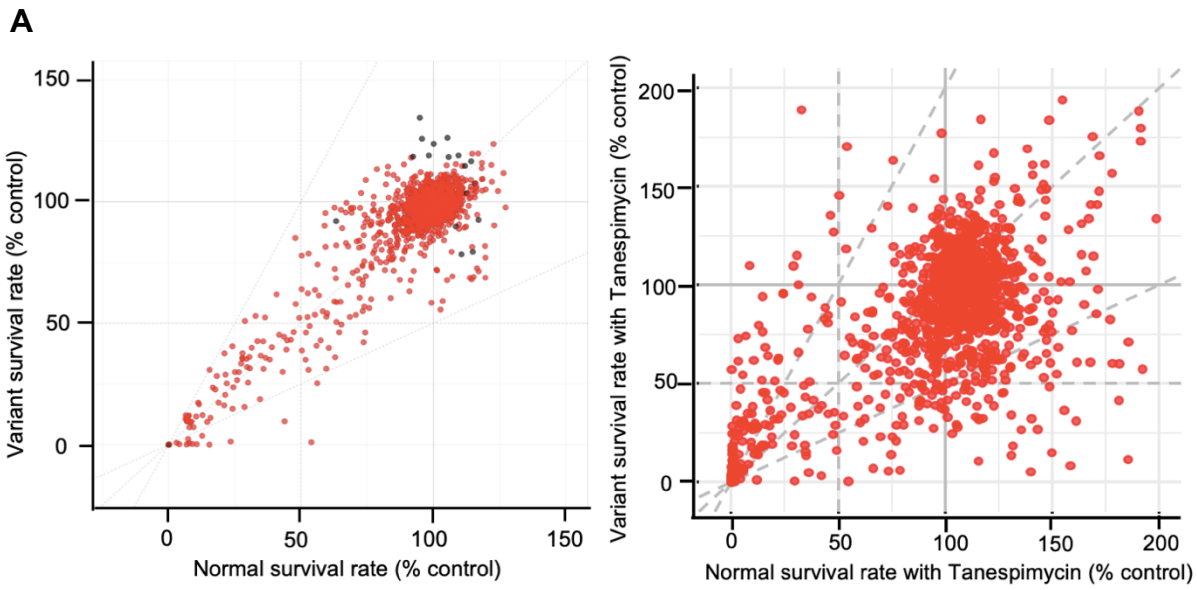
toxicity of two inhibitors in CLST1 clones compared to CLST2 and CLST3 clones, we noticed that most of them were also increasing the survivability of CLST2 and CLST3 clones to some extent (data not shown). Thus, the complete elimination of two variants with these synergies of drugs was not possible. Only one out of those eight compounds promoted CLST1 clones without having any protective effects in CLST2 and CLST3 clones (Fig. 4.4C, *left*). LY2811376, a beta-secretase inhibitor, in combination with Tanespimycin increased the differential window by more than 25-fold compared to 5-fold previously. Similarly, LY2811376 with Gedatolisib increased the differential window by more than 100-fold compared to 50-fold previously (Fig. 4.4C, *right*).

4.3.4 Drugs recover the phenotype in the xenograft model

Despite an enormous amount of data from thousands of Crohn's patients and advances in genome-wide association and immune studies, we still lack a good model system of the disease (Pizarro et al., 2000). We showed that inflammation and fibrosis, two major complications of the disease, were major distinguishable features of Crohn's patient stem cell libraries in xenograft (Fig.3.10C, Fig. 3.13D). CLST3 marked by PSCA and both variants marked by VSIG1 showed a high correlation with inflammation and fibrosis seen in the library's xenograft respectively (Fig. 3.11A, Fig. 3.13E). Xenograft mouse model allows us to test the compounds *in vivo* not only to assess effective targeting of variant clones but also the pathological phenotype associated with them. Thus, we made use of this humanized xenograft mouse model to test the drugs we discovered through high throughput screening (HTS).

CLST1 clones were co-injected at a ratio of 1:1 with either CLST2 or CLST3 clones in xenograft. Xenografts harvested from the mice treated with Tanespimycin or Gedatolisib two days post co-injection of the clones showed necrosis by hematoxylin and eosin (H&E) staining

Fig. 4.4. Synergistic HTS identifies drugs that promote CLST1
A. *Left*, A scatter plot showing survival rate from one CLST1 clone compared to survival rate from one CLST3 clone screened in Selleck library at 1 μM. *Right,* A scatter plot showing survival rate from same clones (in *Left*) screened in Selleck library at 1 μM in presence of Tanespimycin (sub-lethal dose to CLST3 clones). **B.** *Left***,** Example of two small molecules circled in A which in combination with Tanespimycin showed synergistic promotion effects in CLST1. *Right,* List of top 8 compounds selected for validation *in vitro* to combine with Tanespimycin. Scale bars, 200 μm. **C.** Dosage response curve of three variants with CLST2 and CLST3 specific inhibitors, Tanespimycin (*Left*) and Gedatolisib (*Right*) in presence of the best candidate promoter, LY2811376 at 1 μM.



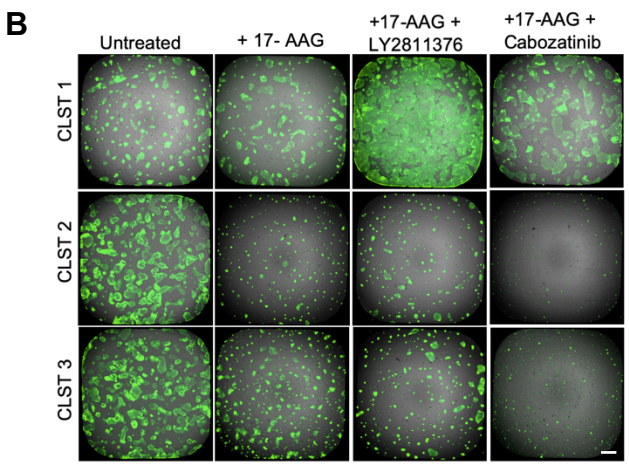
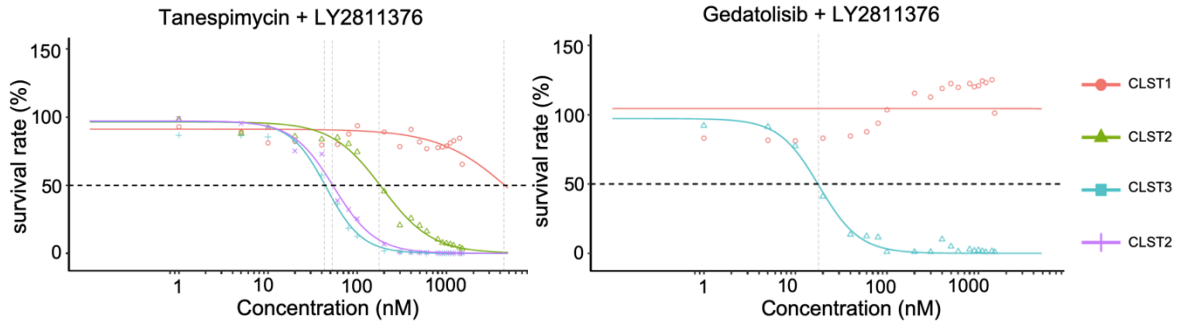


Table 4.3. Promoters selected for validation *in vitro*

Promoters for validation	**Target**
PP1	Phosphatase
AT7867	Akt
Tivozanib	VEGFR
PF-573228	FAK
Cabozatinib	VEGFR
LY2811376	Beta secretase
NVP-ADW742	IGF-1R
LY2784544	JAK-STAT1/2



compared to their untreated control (data not shown). The immunofluorescence staining of CLST1, 2, and 3 specific markers in cross-sections of xenografts from treated mice detected only CLST1 markers in intact epithelial cells suggesting normal epithelial cells were left to regenerate. In contrast, all three cluster-specific markers were still detected in untreated controls (data not shown). The columnar epithelial cells in xenografts that survived the treatment reacted positively to CEACAM5, a CLST1 specific marker, and negatively to VSIG1, a CLST2, and CLST3 specific marker. VSIG1 stained positively in necrotic epithelial cells suggesting CLST2 and CLST3 clones did not survive the drug treatment. Most importantly, neutrophils (by CD45/LY6G staining) and myofibroblasts (by α-SMA/FN1 staining) in CLST1 and CLST3 co-injection xenografts were mostly absent after the treatment while in untreated they were still prominent. Similarly, myofibroblasts in CLST1 and CLST2 co-injection xenograft mostly disappeared after the treatment. The disappearance of neutrophils and myofibroblasts from xenografts correlated with necrosis of CLST2 and CLST3 clones after the treatment with Tanespimycin and Gedatolisib in mice (data not shown).

Since we used a single representative clone from each variant in the xenograft mouse model, the question remained whether the treatment would show comparable results using the cell populations representing three clusters. To resolve this, we injected four Crohn's patients and one control patient library and took the same effective dosage of compounds that worked in the co-injection xenograft experiment. Unlike their corresponding libraries xenograft from untreated mice, Crohn's libraries xenografts from mice treated with both Tanespimycin and Gedatolisib showed extensive necrosis visible by hematoxylin and eosin staining (H&E) (Fig. 4.5A). The xenograft of the control library showed no signs of necrosis from either of the mice treated with two compounds. The results suggested only CLST2 and CLST3 variants epithelia were targeted in Crohn's patient libraries xenograft leaving out CLST1 epithelia to regenerate.

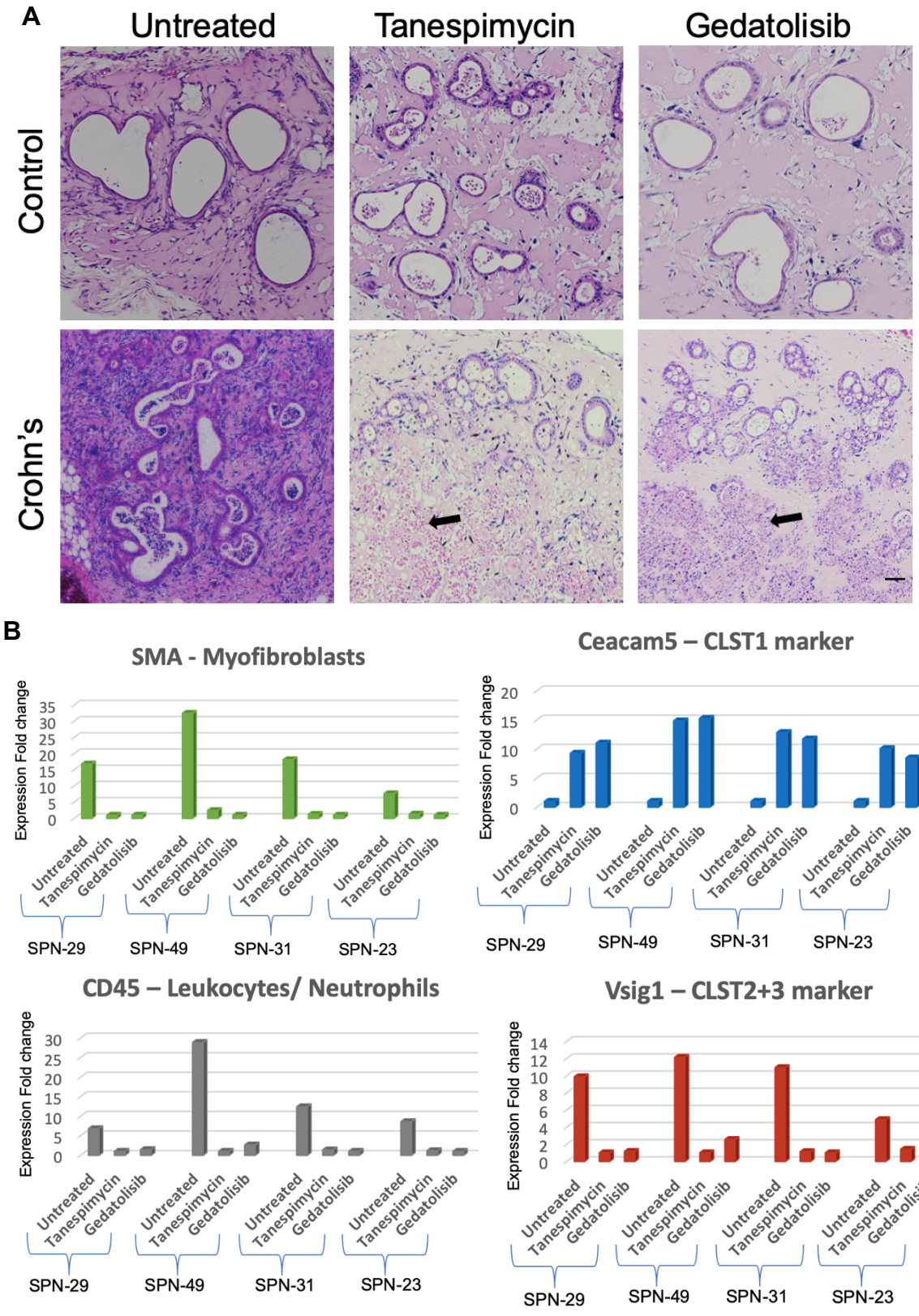
This notion was supported by significantly lesser reads of VSIG1, a CLST2, and CLST3 specific marker, by qPCR after treatment with two drugs. Conversely, CEACAM5 (CLST1 specific marker) reads increased after the treatment (Fig. 4.5B). Most importantly, a significantly lesser number of α-SMA and CD45 reads were seen after the treatment.

The qPCR assessment was further corroborated by immunofluorescence staining of CLST1, 2, and 3 specific markers in cross-section of xenografts from both untreated and treated mice (Fig.4.5C). The epithelial cells in xenografts of Crohn's patients' libraries that survived the treatment reacted positively to CEACAM5, a CLST1 specific marker, and negatively to VSIG1, a CLST2 and CLST3 specific marker. VSIG1 stained positively only in necrotic epithelial cells suggesting CLST2 and CLST3 clones did not survive the drug treatment. Further, neutrophils (by CD45/LY6G staining) and myofibroblasts (by α-SMA/FN1 staining) in Crohn's patients' libraries xenografts mostly disappeared after the treatment while in untreated libraries xenografts they were still significant (Fig.4.5C). In addition, the treated Crohn's libraries xenografts also showed MUC2, a marker of goblet cells specific of normal CLST1, staining by immunofluorescence. Incorporating the data from *in vivo* drug validation in the xenograft mouse model, it was evident that not only CLST2 and CLST3 variant epithelia can be specifically targeted, but the inflammation and fibrosis phenotype related to those variants can also be prevented. This leaves the normal CLST1 epithelia to regenerate and maintain intestinal homeostasis.

4.4 Discussion

Despite undisputable positive impact of current treatment strategies, the patients with stable clinical remission are as low as 10% (Baumgart and Sandborn, 2012). This suggests the possibility of biological therapies being directed towards only maintaining the condition and

Fig. 4.5. Validation of two best candidate drugs in a xenograft mouse model
A. Histology of xenografts of control SPN-19 and Crohn's SPN-23 libraries stained by H&E before and after the treatment with Tanespimycin and Gedatolisib *in vivo* intraperitoneally. Arrows indicate the region of necrosis after the treatment with two candidate small molecules. Scale bars, 100 μm. **B.** Histogram showing relative expression of α–SMA (a myofibroblast marker), CD45 (a leukocyte marker), CEACAM5 (a CLST1 specific marker), and VSIG1 (a CLST2 and CLST3 specific marker) from xenografts of untreated and treated libraries of four independent Crohn's patients by qPCR. **C.** Histology of xenografts of control SPN-19 and Crohn's SPN-23 libraries stained by H&E and by immunofluorescence with antibodies to α–SMA (green), ECAD (red), CEACAM5 (red), FN1 (green), VSIG1 (green), MUC2 (red), MUC5AC (red) before and after the treatment with Tanespimycin and Gedatolisib *in vivo* intraperitoneally. Scale bars, 100 μm.


A
Untreated
Tanespimycin
Gedatolisib
Control
Crohn's
B
SMA - Myofibroblasts
Expression Fold change
Untreated
Tanespimycin
Gedatolisib
SPN-29
SPN-49
SPN-31
SPN-23
Ceacam5 – CLST1 marker
CD45 – Leukocytes/ Neutrophils
Vsig1 – CLST2+3 marker

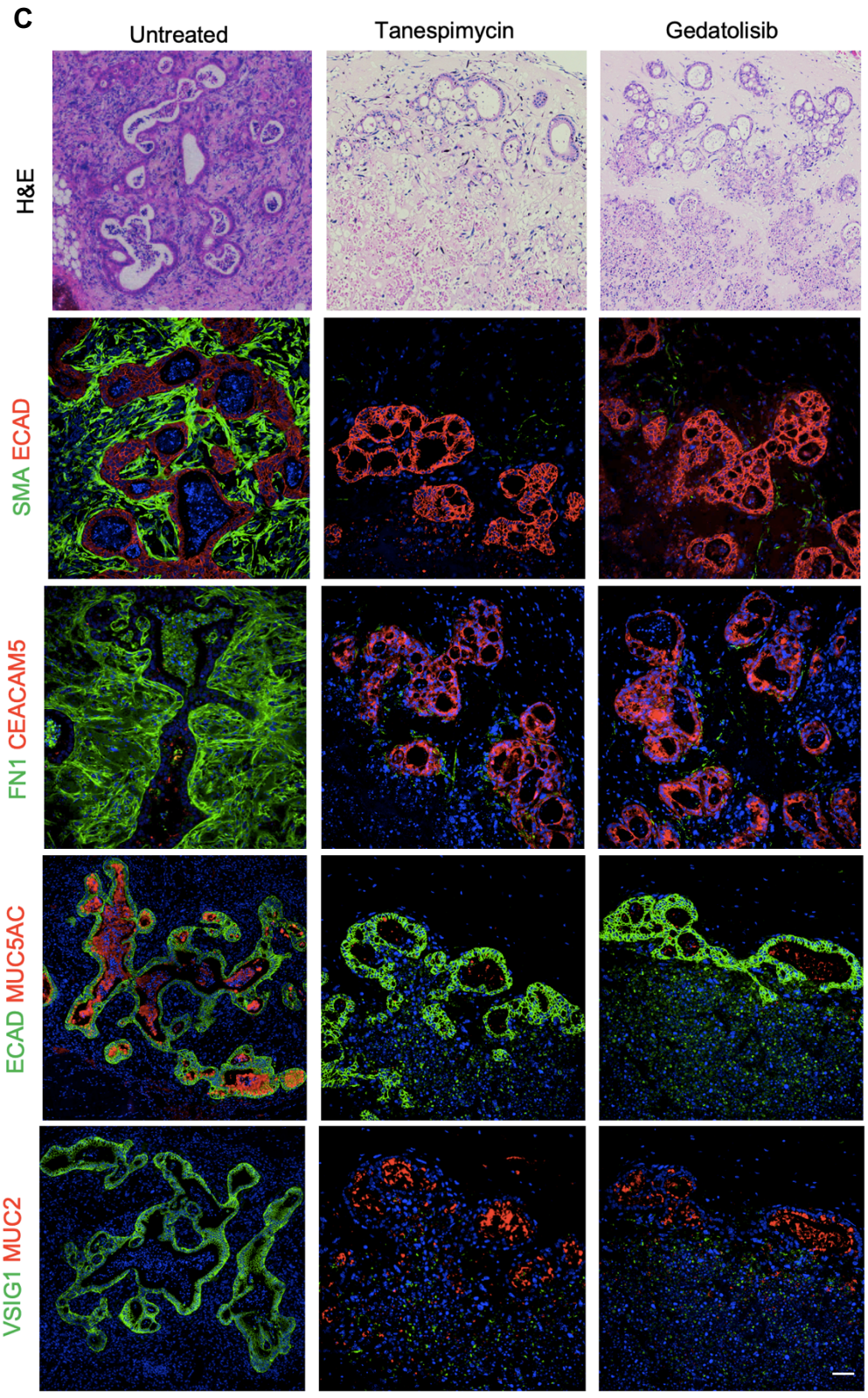

C
Untreated
Tanespimycin
Gedatolisib
H&E
SMA ECAD
FN1 CEACAM5
ECAD MUC5AC
VSIG1 MUC2

Fig. 4.5. (cont.)

illuminates the continuation of a therapeutic gap in the treatment of Crohn's disease. The present work illustrates the discovery of small molecules that could be more central to the disease. The clonogenic mucosal stem cell variants tied to pathological features of the disease were selectively targeted in the xenograft model diminishing those features in parallel and leaving normal mucosal stem cells to compensate for the loss of those variants.

The intestinal epithelial stem cell cloning technology used here was essential to isolate and characterize single cell-derived clones representing three groups of stem cells within a Crohn's patient stem cell library. Without the expansion of these clones as separate lines away from their libraries, it would be challenging to resolve the heterogeneity and to render two variants pathogenic. Most importantly, it would be impossible to assess those target variant stem cells against control normal stem cells for differential sensitivity with drugs or drug combinations. Moreover, the rapid and unlimited proliferative capacity of these single cell-derived clones allowed us to employ them in high throughput to screen for more than 4000 small molecules treatment over four chemical libraries. The four chemical libraries with a diverse collection of around 2000 unique small molecules focused on a large spectrum of targets revealed candidate drugs that selectively targeted particular pathways critical to the survival of pathogenic variants. Lastly, single cell-derived clones revealed markers consistently expressed specific to each cluster which was essential to validate selective targeting of variant stem cells by the candidate drugs in Crohn's stem cell libraries and coculture.

The genetic labeling of individual clones representative of CLST1, CLST2, and CLST3 by retrovirus carrying fluorescent proteins allowed constitutive expression of the reporter in the stem cells. The detection of the fluorescent signal made it extremely easy to quantify the phenotypic effect of the small molecules, inhibition, or promotion, without the use of other

complex secondary assays (Blucher and McWeeney, 2014). Genetic labeling offered several advantages in this context. For one, it made the entire process more efficient bypassing secondary assays and made quantification more accurate avoiding noise that was likely to present as a result of those assays. Both of these are central to high throughput screening in expediting the hit discovery process while minimizing the false discovery rate (Fox et al., 2006). Most notably, we could trace activities of variant and normal stem cells during drug treatment *in vitro*. This feature was extremely valuable in showing that CLST1 clones took the opportunity of dominating the culture upon selective targeting of CLST2 and CLST3 clones in a co-culture experiment *in vitro*. While *in vivo*, the only endpoint of the experiment was visualized which also revealed the domination of xenograft nodules by CLST1 epithelia.

The screening of clones in four chemical libraries revealed 72 hit compounds common to all the CLST2 and CLST3 clones from multiple patients. The target enriched pathway or enriched target protein of those small molecules showed only a handful of pathways or proteins. This suggests a common survival mechanism in pathogenic stem cell variants and the possibility of generalizing the therapeutics. Also, none of the common hits belonged to the class of anti-inflammatory or immune modulators suggesting the inability of current therapeutics to target variant clones which might be central to the disease. Detailed dosage response with selected best nine compounds validated the result from HTS performed at a single concentration of 1 micromolar (μM). The maximum differential window between CLST1 clones and the two variant clones was at best, 100-fold with TTNPB, 50-fold with Gedatolisib, and only 5 to 10-fold with other small molecules. Even though they successfully eliminated both variant stem cells *in vitro* in coculture and a Crohn's library, the narrow therapeutic window is likely to cause potential cytotoxicity problems in normal cells and organs nearby. Further, subjecting the clones to ultra-high-throughput screens for a larger array of compounds will possibly identify

candidates with a larger therapeutic window in the future (Sundberg, 2000).

Meanwhile, we aimed to resolve the issue by establishing "synthetic lethal" screens (McCormick, 2015; Thompson et al., 2017; Aguirre and Hahn, 2018) against variant stem cells using Selleck chemical library in a background of sub-lethal concentration of Tanespimycin that selectively targeted those stem cells. Interestingly, the synthetic lethal screen identified several small molecules that, in the background of Tanespimycin, augmented only CLST1 stem cells while still inhibiting variant stem cells effectively. Comparing their behavior in previous screenings without the background of a sub-lethal dose of Tanespimycin showed that these compounds promoted all three variants. This suggested that CLST2 and CLST3 clones sensitized by their specific inhibitor, which might have triggered apoptosis in them, were unable to recover or to respond to promoters. In contrast, the CLST1 clone which shows a 5-fold lesser effect of Tanespimycin compared to the other two variants still responded to promoters exactly like before. This phenomenon allowed us to make use of such promoters to widen the therapeutic window which was the primary objective of finding synergistic compounds. Among them, LY2811376 in a synergistic manner with both Tanespimycin and Gedatolisib was the most effective in promoting the normal stem cells while having no protective effects in variant stem cells. Most importantly, dosage-response curves for these variant stem cells with those combinations, in the background of a fixed concentration of the LY2811376, showed a remarkable differential of nearly 100-fold.

The variant stem cells linked to neutrophilic inflammation and fibrosis were selectively targeted *in vivo* rescuing those features in parallel and leaving normal mucosal stem cells to regenerate. Four independent Crohn’s libraries xenografts from mice treated with Tanespimycin and Gedatolisib as single agents showed extensive cell death while that of control libraries

showed no such signs. This notion was supported by both gene expression quantification and immunofluorescence staining of antibodies to markers of specific clusters in treated xenografts. In addition, treated Crohn's libraries epithelia were marked by the presence of goblet cells, shown by MUC2 staining in immunofluorescence, suggesting remaining CLST1 epithelia were resuming functions typical of epithelial barrier (Cornick et al., 2015).

While this model offers several advantages such as testing efficacy against multiple patients, quantification of variant clusters and disease-related pathology before and after treatment, and ease of operation, it also comes with some limitations. One of the limitations posed by the current *in vivo* model is the inability to organize long-term studies because of the relatively short life (two weeks) of intestinal epithelial cells in subcutaneous regions of the mouse. For example, we showed myofibroblast accumulation, which is an early stage of fibrosis (D'Haens et al., 2019), as an assessment of fibrosis and showed the disappearance of myofibroblasts from Crohn's libraries after the treatment with candidate drugs. Hence, the study of the fully developed diseased state, including late fibrosis stages with excess collagen deposition, as well as the assessment of drugs on later stages of the disease remain unexplored. Perhaps, immunodeficient mice compensated with human-specific factors required for epithelial cell activity would allow long-term maintenance of this lineage in subcutaneous regions (Zhu et al., 2019). Another limitation of this mouse model is the challenge to accurately quantify regeneration by remaining CLST1 epithelia after the treatment. We showed a resumption of normal barrier functions by CLST1 epithelia that survived the treatment but could not find a way to show if the loss of CLST2 and CLST3 epithelia were fully compensated by remaining CLST1 epithelia. Such valuation is critical in terms of clinical perspective in not letting the intestine completely deprived of epithelial cells which can bring severe consequences (Odenwald and Turner, 2017). Possibly, a more subtle approach would be using a lower dosage

of inhibitors in combination therapy with promoters such as LY2811376 so that the synergy gives a competitive advantage to normal mucosal stem cells. To answer those questions, more studies with various dosage with promoter are required in this xenograft model or in any other relevant models which link variant epithelia to the pathology of Crohn's disease.

Dextran Sodium Sulfate (DSS) and TNBS- induced rodent colitis are two well-established chemical injury models in the study of inflammatory bowel disease (IBD), including both Crohn's and ulcerative colitis (UC). The two major phenotypes of IBD, chronic inflammation and fibrosis, have been recapitulated in these models following injury with these chemicals that damage the distal gastrointestinal tract (Okayashu et al., 1990; Morris et al., 1989). In accompanying research by our lab, the rats treated with two or more cycles of DSS or with TNSB displayed extensive epithelial damage and displayed multiple mucosal regions that reacted positively to variant markers, LCN2 and VSIG1, in both colon and terminal ileum. These regions were also devoid of goblet cells (by MUC2) and showed a generalized expansion of α–SMA-expressing myofibroblasts and inflammation around them (data not shown). Normal-looking mucosal regions surrounded those VSIG1+/LCN2+ mucosal regions, an observation which is in accordance with the co-existence of variants in Crohn's patients. Immunofluorescence staining to KI67, a proliferation marker, showed that such abnormal regions were in an active state of cell division suggesting they are expanding in the mucosa. Toxicity of *in vitro* cultured epithelial cells has been shown with both chemicals before (Ni et al., 1996; Yamada et al., 1992). Perhaps, while causing damage to epithelia they allow variant stem cells existing in lower proportion to expand. These might be more resistant to these chemicals or located deep in the mucosa protected from the treatment. Nevertheless, we see their emergence, and these models might provide us a mechanism of expansion of these variants. But more comprehensive research will be required to confirm the mechanism and to link these

models to the variant cells that were found in Crohn's patients' biopsies. Cloning these epithelial variants from rodents and confirming their pro-inflammatory and pro-fibrotic nature is necessary to conclude that these variants drive inflammation and fibrosis in the long run. Such models present an excellent opportunity to connect current human-related observations to an animal model as well as give a chance to test candidate drugs in a regenerating intestine.

BIBLIOGRAPHY

Abraham, C., & Cho, J. H. (2009). Inflammatory bowel disease. The New England Journal of Medicine, 361(21), 2066–2078.

Adegbola, S. O., Sahnan, K., Warusavitarne, J., Hart, A., & Tozer, P. (2018). Anti-TNF therapy in Crohn's disease. International Journal of Molecular Sciences, 19(8), 2244.

Adolph, T. E., Tomczak, M. F., Niederreiter, L., Ko, H. J., Böck, J., Martinez-Naves, E., Glickman, J. N., Tschurtschenthaler, M., Hartwig, J., Hosomi, S., Flak, M. B., Cusick, J. L., Kohno, K., Iwawaki, T., Billmann-Born, S., Raine, T., Bharti, R., Lucius, R., Kweon, M. N., Marciniak, S. J., Blumberg, R. S. (2013). Paneth cells as a site of origin for intestinal inflammation. Nature, 503(7475), 272–276.

Agarwal, S., Stucchi, A.F., Dendrinos, K., Cerda, S., O'Brien, M.J., Becker, J.M., Heeren, T., and Farraye, F.A. (2013). Is pyloric gland metaplasia in ileal pouch biopsies a marker for Crohn's disease? Digestive Disease & Sciences, 58(10), 2918-2925.

Aguirre, A. J., & Hahn, W. C. (2018). Synthetic lethal vulnerabilities in KRAS-mutant cancers. Cold Spring Harbor Perspectives in Medicine, 8(8), a031518.

Amabile, G., Welner, R. S., Nombela-Arrieta, C., D'Alise, A. M., Di Ruscio, A., Ebralidze, A. K., Kraytsberg, Y., Ye, M., Kocher, O., Neuberg, D. S., Khrapko, K., Silberstein, L. E., & Tenen, D. G. (2013). In vivo generation of transplantable human hematopoietic cells from induced pluripotent stem cells. Blood, 121(8), 1255–1264.

Antoniou, E., Margonis, G. A., Angelou, A., Pikouli, A., Argiri, P., Karavokyros, I., Papalois, A., & Pikoulis, E. (2016). The TNBS-induced colitis animal model: An overview. Annals of Medicine and Surgery (2012), 11, 9–15.

Araya, J., and Nishimura, S.L. (2010). Fibrogenic reactions in lung disease. Annual Review of Pathology *5*, 77-98.

Ashburn, T. T., & Thor, K. B. (2004). Drug repositioning: identifying and developing new uses for existing drugs. Nature Reviews. Drug Discovery, 3(8), 673–683.

Atreya, I., Atreya, R., & Neurath, M. F. (2008). NF-kappaB in inflammatory bowel disease. Journal of Internal Medicine, 263(6), 591–596.

Barrett, J. C., Hansoul, S., Nicolae, D. L., Cho, J. H., Duerr, R. H., Rioux, J. D., Brant, S. R., Silverberg, M. S., Taylor, K. D., Barmada, M. M., Bitton, A., Dassopoulos, T., Datta, L. W., Green, T., Griffiths, A. M., Kistner, E. O., Murtha, M. T., Regueiro, M. D., Rotter, J. I., Schumm, L. P., … Daly, M. J. (2008). Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. Nature Genetics, 40(8), 955–962.

Baumgart, D.C., and Sandborn, W.J. (2012). Crohn's disease. Lancet, 380(9853), 1590-1605.

Baumgart, D. C., Thomas, S., Przesdzing, I., Metzke, D., Bielecki, C., Lehmann, S. M., Lehnardt, S., Dörffel, Y., Sturm, A., Scheffold, A., Schmitz, J., & Radbruch, A. (2009). Exaggerated inflammatory response of primary human myeloid dendritic cells to lipopolysaccharide in patients with inflammatory bowel disease. Clinical and Experimental Immunology, 157(3), 423–436.

Bindea, G., Galon, J., and Mlecnik, B. (2013). CluePedia Cytoscape plugin: Pathway insights using integrated experimental and in silico data. Bioinformatics, 29(5), 661-663.

Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., Fridman, W.-H., Pag.s, F., Trajanoski, Z., and Galon, J. (2009). ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics, 25(8), 1091-1093.

Bloom, S. M., Bijanki, V. N., Nava, G. M., Sun, L., Malvin, N. P., Donermeyer, D. L., Dunne, W. M., Jr, Allen, P. M., & Stappenbeck, T. S. (2011). Commensal Bacteroides species induce colitis in host-genotype-specific fashion in a mouse model of inflammatory bowel disease. Cell Host & Microbe, 9(5), 390–403.

Blucher, A. S., & McWeeney, S. K. (2014). Challenges in secondary analysis of high throughput screening data. Pacific Symposium on Biocomputing, 114–124.

Boppana, K., Dubey, P. K., Jagarlapudi, S. A., Vadivelan, S., & Rambabu, G. (2009). Knowledge based identification of MAO-B selective inhibitors using pharmacophore and structure based virtual screening models. European Journal of Medicinal Chemistry, 44(9), 3584–3590.

Bosma, G. C., Custer, R. P., & Bosma, M. J. (1983). A severe combined immunodeficiency mutation in the mouse. Nature, 301(5900), 527–530.

Both, H., Torp-Pedersen, K., Kreiner, S., Hendriksen, C., & Binder, V. (1983). Clinical appearance at diagnosis of ulcerative colitis and Crohn's disease in a regional patient group. Scandinavian Journal of Gastroenterology, 18(7), 987–991.

Brazil, J. C., & Parkos, C. A. (2016). Pathobiology of neutrophil-epithelial interactions. Immunological Reviews, 273(1), 94–111.

Buisine, M. P., Desreumaux, P., Debailleul, V., Gambiez, L., Geboes, K., Ectors, N., Delescaut, M. P., Degand, P., Aubert, J. P., Colombel, J. F., & Porchet, N. (1999). Abnormalities in mucin gene expression in Crohn's disease. Inflammatory Bowel Diseases, 5(1), 24–32.

Burke, J. P., Mulsow, J. J., O'Keane, C., Docherty, N. G., Watson, R. W., & O'Connell, P. R. (2007). Fibrogenesis in Crohn's disease. The American Journal of Gastroenterology, 102(2), 439–448.

Cadwell, K., Liu, J. Y., Brown, S. L., Miyoshi, H., Loh, J., Lennerz, J. K., Kishi, C., Kc, W., Carrero, J. A., Hunt, S., Stone, C. D., Brunt, E. M., Xavier, R. J., Sleckman, B. P., Li, E., Mizushima, N., Stappenbeck, T. S., & Virgin, H. W., 4th (2008). A key role for autophagy and

the autophagy gene Atg16l1 in mouse and human intestinal Paneth cells. Nature, 456(7219), 259–263.

Cezar G. G. (2007). Can human embryonic stem cells contribute to the discovery of safer and more effective drugs?. Current Opinion in Chemical Biology, 11(4), 405–409.

Chande, N., Townsend, C. M., Parker, C. E., & MacDonald, J. K. (2016). Azathioprine or 6-mercaptopurine for induction of remission in Crohn's disease. The Cochrane Database of Systematic Reviews, 10(10), CD000545.

Cheifetz A. S. (2013). Management of active Crohn disease. The Journal of the American Medical Association, 309(20), 2150–2158.

Chen, E. Y., Tan, C. M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G. V., Clark, N. R., & Ma'ayan, A. (2013). Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics, 14, 128.

Cleynen, I., Boucher, G., Jostins, L., Schumm, L. P., Zeissig, S., Ahmad, T., Andersen, V., Andrews, J. M., Annese, V., Brand, S., Brant, S. R., Cho, J. H., Daly, M. J., Dubinsky, M., Duerr, R. H., Ferguson, L. R., Franke, A., Gearry, R. B., Goyette, P., Hakonarson, H., … Lees, C. W. (2016). Inherited determinants of Crohn's disease and ulcerative colitis phenotypes: a genetic association study. Lancet, 387(10014), 156–167.

Cooney, R., Baker, J., Brain, O., Danis, B., Pichulik, T., Allan, P., Ferguson, D. J., Campbell, B. J., Jewell, D., & Simmons, A. (2010). NOD2 stimulation induces autophagy in dendritic cells influencing bacterial handling and antigen presentation. Nature Medicine, 16(1), 90–97.

Cornick, S., Tawiah, A., & Chadee, K. (2015). Roles and regulation of the mucus barrier in the gut. Tissue Barriers, 3(1-2), e982426.

Coskun, M., Salem, M., Pedersen, J., & Nielsen, O. H. (2013). Involvement of JAK/STAT signaling in the pathogenesis of inflammatory bowel disease. Pharmacological Research, 76, 1–8.

Dambha, F., Tanner, J., & Carroll, N. (2014). Diagnostic imaging in Crohn's disease: what is the new gold standard?. Best practice & research. Clinical gastroenterology, 28(3), 421–436.
Daub, A., Sharma, P., & Finkbeiner, S. (2009). High-content screening of primary neurons: ready for prime time. Current Opinion in Neurobiology, 19(5), 537–543.

de Lange, K. M., Moutsianas, L., Lee, J. C., Lamb, C. A., Luo, Y., Kennedy, N. A., Jostins, L., Rice, D. L., Gutierrez-Achury, J., Ji, S. G., Heap, G., Nimmo, E. R., Edwards, C., Henderson, P., Mowat, C., Sanderson, J., Satsangi, J., Simmons, A., Wilson, D. C., Tremelling, M., … Barrett, J. C. (2017). Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. Nature Genetics, 49(2), 256–261.

de Oliveira, S., Reyes-Aldasoro, C. C., Candel, S., Renshaw, S. A., Mulero, V., & Calado, A. (2013). Cxcl8 (IL-8) mediates neutrophil recruitment and behavior in the zebrafish inflammatory response. Journal of Immunology, 190(8), 4349–4359.

Dethlefsen, L., & Relman, D. A. (2011). Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. Proceedings of the National Academy of Sciences, 108 Suppl 1(Suppl 1), 4554–4561.

D'Haens, G., Rieder, F., Feagan, B. G., Higgins, P., Panes, J., Maaser, C., Rogler, G., Löwenberg, M., van der Voort, R., Pinzani, M., Peyrin-Biroulet, L., Danese, S., & IOIBD Fibrosis Working Group (2019). Challenges in the pathophysiology, diagnosis and management of intestinal fibrosis in inflammatory bowel disease. Gastroenterology, S0016-5085(19)41035-4.

Dieleman, L. A., Ridwan, B. U., Tennyson, G. S., Beagley, K. W., Bucy, R. P., & Elson, C. O. (1994). Dextran sulfate sodium-induced colitis occurs in severe combined immunodeficient mice. Gastroenterology, 107(6), 1643–1652.

D'Incà, R., Di Leo, V., Corrao, G., Martines, D., D'Odorico, A., Mestriner, C., Venturi, C., Longo, G., & Sturniolo, G. C. (1999). Intestinal permeability test as a predictor of clinical course in Crohn's disease. The American Journal of Gastroenterology, 94(10), 2956–2960.

Dulai, P. S., Siegel, C. A., & Peyrin-Biroulet, L. (2014). Anti-tumor necrosis factor-α monotherapy versus combination therapy with an immunomodulator in IBD. Gastroenterology Clinics of North America, 43(3), 441–456.

Duleba, M., Qi, Y., Mahalingam, R., Liew, A. A., Neupane, R., Flynn, K., Rinaldi, F., Vincent, M., Crum, C. P., Ho, K. Y., Hou, J. K., Hyams, J. S., Sylvester, F. A., McKeon, F., & Xian, W. (2019). An efficient method for cloning gastrointestinal stem cells from patients via endoscopic biopsies. Gastroenterology, 156(1), 20–23.

Duleba, M., Yamamoto, Y., Neupane, R., Rao, W., Xie, J., Qi, Y., Liew, A. A., Niroula, S., Zhang, Y., Mahalingam, R., Wang, S., Goller, K., Ajani, J. A., Vincent, M., Ho, K. Y., Hou, J. K., Hyams, J. S., Sylvester, F. A., Crum, C. P., McKeon, F., Xian, W. (2020). Cloning of ground-state intestinal stem cells from endoscopic biopsy samples. Nature Protocols, 15(5), 1612–1627.

Dunne, A., Jowett, M., & Rees, S. (2009). Use of primary human cells in high-throughput screens. Methods in Molecular Biology, 565, 239–257.

Eglen, R. M., Gilchrist, A., & Reisine, T. (2008). An overview of drug screening using primary and embryonic stem cells. Combinatorial Chemistry & High Throughput screening, 11(7), 566–572.

Eglen, R. M., Gilchrist, A., & Reisine, T. (2008). The use of immortalized cell lines in GPCR screening: the good, bad and ugly. Combinatorial Chemistry & High Throughput Screening, 11(7), 560–565.

Elinav, E., Strowig, T., Kau, A. L., Henao-Mejia, J., Thaiss, C. A., Booth, C. J., Peaper, D. R., Bertin, J., Eisenbarth, S. C., Gordon, J. I., & Flavell, R. A. (2011). NLRP6 inflammasome regulates colonic microbial ecology and risk for colitis. Cell, 145(5), 745–757.

Farmer, R. G., Hawk, W. A., & Turnbull, R. B., Jr (1975). Clinical patterns in Crohn's disease: a statistical study of 615 cases. Gastroenterology, 68(4 Pt 1), 627–635.

Ferguson, A., & Chen, K. (2020). Analysis of Transcriptional Profiling of Immune Cells at the Single-Cell Level. Methods in Molecular Biology, 2111, 47–57.

Fielding, C. A., McLoughlin, R. M., McLeod, L., Colmont, C. S., Najdovska, M., Grail, D., Ernst, M., Jones, S. A., Topley, N., & Jenkins, B. J. (2008). IL-6 regulates neutrophil trafficking during acute inflammation via STAT3. Journal of Immunology, 181(3), 2189–2195.

Fiocchi C. (1998). Inflammatory bowel disease: etiology and pathogenesis. Gastroenterology, 115(1), 182–205.

Flordellis, C. S., Manolis, A. S., Paris, H., & Karabinis, A. (2006). Rethinking target discovery in polygenic diseases. Current Topics in Medicinal Chemistry, 6(16), 1791–1798.

Florén, C. H., Benoni, C., & Willén, R. (1987). Histologic and colonoscopic assessment of disease extension in ulcerative colitis. Scandinavian Journal of Gastroenterology, 22(4), 459–462.

Fox, S., Farr-Jones, S., Sopchak, L., Boggs, A., Nicely, H. W., Khoury, R., & Biros, M. (2006). High-throughput screening: update on practices and success. Journal of Biomolecular Screening, 11(7), 864–869.

Frank, D. N., St Amand, A. L., Feldman, R. A., Boedeker, E. C., Harpaz, N., & Pace, N. R. (2007). Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. Proceedings of the National Academy of Sciences, 104(34), 13780–13785.

Franke, A., McGovern, D. P., Barrett, J. C., Wang, K., Radford-Smith, G. L., Ahmad, T., Lees, C. W., Balschun, T., Lee, J., Roberts, R., Anderson, C. A., Bis, J. C., Bumpstead, S., Ellinghaus, D., Festen, E. M., Georges, M., Green, T., Haritunians, T., Jostins, L., Latiano, A., … Parkes, M. (2010). Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. Nature Genetics, 42(12), 1118–1125.

Frolkis, A. D., Dykeman, J., Negrón, M. E., Debruyn, J., Jette, N., Fiest, K. M., Frolkis, T., Barkema, H. W., Rioux, K. P., Panaccione, R., Ghosh, S., Wiebe, S., & Kaplan, G. G. (2013). Risk of surgery for inflammatory bowel diseases has decreased over time: a systematic review and meta-analysis of population-based studies. Gastroenterology, 145(5), 996–1006.

García Rodríguez, L. A., Ruigómez, A., & Panés, J. (2006). Acute gastroenteritis is followed by an increased risk of inflammatory bowel disease. Gastroenterology, 130(6), 1588–1594.

Geboes K. (2001). Pathology of inflammatory bowel diseases (IBD): variability with time and treatment. Colorectal Disease, 3(1), 2–12.

Gisbert JP, Gomollón F, Maté J, Pajares JM. (2002). Role of 5-aminosalicylic acid (5-ASA) in treatment of inflammatory bowel disease: a systemic review. Digestive Diseases and Sciences, 47(3), 471-488.

Giuliano K.A., DeBiasio R.L., Dunlay R.T., Gough A., Volosky J.M., Zock J., Pavlakis G.N., Taylor D.L. High-content screening: A new approach to easing key bottlenecks in the drug discovery process. Journal of Biomolecular Screening, 2:249–259.

Glocker, E. O., Kotlarz, D., Boztug, K., Gertz, E. M., Schäffer, A. A., Noyan, F., Perro, M., Diestelhorst, J., Allroth, A., Murugan, D., Hätscher, N., Pfeifer, D., Sykora, K. W., Sauer, M., Kreipe, H., Lacher, M., Nustede, R., Woellner, C., Baumann, U., Salzer, U., … Klein, C. (2009). Inflammatory bowel disease and mutations affecting the interleukin-10 receptor. The New England Journal of Medicine, 361(21), 2033–2045.

Gluckman, P. D., Hanson, M. A., Cooper, C., & Thornburg, K. L. (2008). Effect of in utero and early-life conditions on adult health and disease. The New England Journal of Medicine, 359(1), 61–73.

Gomollón, F., Dignass, A., Annese, V., Tilg, H., Van Assche, G., Lindsay, J. O., Peyrin-Biroulet, L., Cullen, G. J., Daperno, M., Kucharzik, T., Rieder, F., Almer, S., Armuzzi, A., Harbord, M., Langhorst, J., Sans, M., Chowers, Y., Fiorino, G., Juillerat, P., Mantzaris, G. J., ECCO (2017). 3rd European evidence-based consensus on the diagnosis and management of Crohn's disease 2016: Part 1: Diagnosis and medical management. Journal of Crohn's & Colitis, 11(1), 3–25.

Green H. (2008). The birth of therapy with cultured cells. BioEssays, 30(9), 897–903.

Halfvarson, J. (2011). Genetics in twins with Crohn's disease: less pronounced than previously believed? Inflammatory Bowel Disease, 17, 6-12.

Hampe, J., Franke, A., Rosenstiel, P., Till, A., Teuber, M., Huse, K., Albrecht, M., Mayr, G., De La Vega, F. M., Briggs, J., Günther, S., Prescott, N. J., Onnie, C. M., Häsler, R., Sipos, B., Fölsch, U. R., Lengauer, T., Platzer, M., Mathew, C. G., Krawczak, M., … Schreiber, S. (2007). A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in ATG16L1. Nature Genetics, 39(2), 207–211.

Hart, A. L., Al-Hassi, H. O., Rigby, R. J., Bell, S. J., Emmanuel, A. V., Knight, S. C., Kamm, M. A., & Stagg, A. J. (2005). Characteristics of intestinal dendritic cells in inflammatory bowel diseases. Gastroenterology, 129(1), 50–65.

Heresbach, D., Alexandre, J. L., Branger, B., Bretagne, J. F., Cruchant, E., Dabadie, A., Dartois-Hoguin, M., Girardot, P. M., Jouanolle, H., Kerneis, J., Le Verger, J. C., Louvain, V., Politis, J., Richecoeur, M., Robaszkiewicz, M., Seyrig, J. A., & ABERMAD (Association Bretonne d'Etude et de Recherche sur les Maladies de l'Appareil Digestif) (2005). Frequency and significance of granulomas in a cohort of incident cases of Crohn's disease. Gut, 54(2), 215–222.

Hirsch, T., Rothoeft, T., Teig, N., Bauer, J. W., Pellegrini, G., De Rosa, L., Scaglione, D., Reichelt, J., Klausegger, A., Kneisz, D., Romano, O., Secone Seconetti, A., Contin, R., Enzo, E., Jurman, I., Carulli, S., Jacobsen, F., Luecke, T., Lehnhardt, M., Fischer, M., … De Luca, M. (2017). Regeneration of the entire human epidermis using transgenic stem cells. Nature, 551(7680), 327–332.

Hollander D. (1988). Crohn's disease--a permeability disorder of the tight junction?. Gut, 29(12), 1621–1624.

Holländer, G. A., Simpson, S. J., Mizoguchi, E., Nichogiannopoulou, A., She, J., Gutierrez-Ramos, J. C., Bhan, A. K., Burakoff, S. J., Wang, B., & Terhorst, C. (1995). Severe colitis in mice with aberrant thymic selection. Immunity, 3(1), 27–38.

Houck, K. A., & Kavlock, R. J. (2008). Understanding mechanisms of toxicity: insights from drug discovery research. Toxicology and Applied Pharmacology, 227(2), 163–178.

Howell, K. J., Kraiczy, J., Nayak, K. M., Gasparetto, M., Ross, A., Lee, C., Mak, T. N., Koo, B. K., Kumar, N., Lawley, T., Sinha, A., Rosenstiel, P., Heuschkel, R., Stegle, O., & Zilbauer, M. (2018). DNA methylation and transcription patterns in intestinal epithelial cells from pediatric patients with inflammatory bowel diseases differentiate disease subtypes and associate with outcome. Gastroenterology, 154(3), 585–598.

Hughes, J. P., Rees, S., Kalindjian, S. B., & Philpott, K. L. (2011). Principles of early drug discovery. British Journal of Pharmacology, 162(6), 1239–1249.

Hugot, J. P., Chamaillard, M., Zouali, H., Lesage, S., Cézard, J. P., Belaiche, J., Almer, S., Tysk, C., O'Morain, C. A., Gassull, M., Binder, V., Finkel, Y., Cortot, A., Modigliani, R., Laurent-Puig, P., Gower-Rousseau, C., Macry, J., Colombel, J. F., Sahbatou, M., & Thomas, G. (2001). Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. Nature, 411(6837), 599–603.

Hyams, J.S. (2014). Standardized recording of parameters related to the natural history of inflammatory bowel disease: from Montreal to Paris. Digestive Disease, 32, 337-344.

Imam, T., Park, S., Kaplan, M. H., & Olson, M. R. (2018). Effector T helper cell subsets in inflammatory bowel diseases. Frontiers in Immunology, 9, 1212.

Imming, P., Sinning, C., & Meyer, A. (2006). Drugs, their targets and the nature and number of drug targets. Nature Reviews. Drug discovery, 5(10), 821–834.

Ito, R., Katano, I., Ida-Tanaka, M., Kamisako, T., Kawai, K., Suemizu, H., Aiso, S., and Ito, M. (2012). Efficient xenoengraftment in severe immunodeficient NOD/Shi-scid IL2rgammanull mice is attributed to a lack of CD11c+B220+CD122+ cells. The Journal of Immunology, 189, 4313-4320.

Ito, R., Shin-Ya, M., Kishida, T., Urano, A., Takada, R., Sakagami, J., Imanishi, J., Kita, M., Ueda, Y., Iwakura, Y., Kataoka, K., Okanoue, T., & Mazda, O. (2006). Interferon-gamma is causatively involved in experimental inflammatory bowel disease in mice. Clinical and Experimental Immunology, 146(2), 330–338.

Iversen, P. W., Eastwood, B. J., Sittampalam, G. S., & Cox, K. L. (2006). A comparison of assay performance measures in screening assays: signal window, Z' factor, and assay variability ratio. Journal of Biomolecular Screening, 11(3), 247–252.

Jia, J., Zhu, F., Ma, X., Cao, Z., Cao, Z. W., Li, Y., Li, Y. X., & Chen, Y. Z. (2009). Mechanisms of drug combinations: interaction and network perspectives. Nature Reviews. Drug Discovery, 8(2), 111–128.

Joossens, M., Simoens, M., Vermeire, S., Bossuyt, X., Geboes, K., & Rutgeerts, P. (2007). Contribution of genetic and environmental factors in the pathogenesis of Crohn's disease in a large family with multiple cases. Inflammatory Bowel Diseases, 13(5), 580–584.

Joossens, M., Huys, G., Cnockaert, M., De Preter, V., Verbeke, K., Rutgeerts, P., Vandamme, P., & Vermeire, S. (2011). Dysbiosis of the faecal microbiota in patients with Crohn's disease and their unaffected relatives. Gut, 60(5), 631–637.

Jostins, L., Ripke, S., Weersma, R. K., Duerr, R. H., McGovern, D. P., Hui, K. Y., Lee, J. C., Schumm, L. P., Sharma, Y., Anderson, C. A., Essers, J., Mitrovic, M., Ning, K., Cleynen, I., Theatre, E., Spain, S. L., Raychaudhuri, S., Goyette, P., Wei, Z., Abraham, C., … Cho, J. H. (2012). Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature, 491(7422), 119–124.

Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Research, 28, 27-30.

Kaplan, G. G., & Ng, S. C. (2016). Globalisation of inflammatory bowel disease: perspectives from the evolution of inflammatory bowel disease in the UK and China. Lancet. Gastroenterology & Hepatology, 1(4), 307–316.

Kaplan, G. G., & Ng, S. C. (2017). Understanding and preventing the global increase of inflammatory bowel disease. Gastroenterology, 152(2), 313–321.e2.

Karagiannis, P., Takahashi, K., Saito, M., Yoshida, Y., Okita, K., Watanabe, A., Inoue, H., Yamashita, J. K., Todani, M., Nakagawa, M., Osawa, M., Yashiro, Y., Yamanaka, S., & Osafune, K. (2019). Induced pluripotent stem cells and their use in human models of disease and development. Physiological Reviews, 99(1), 79–114.

Kariv, R., Plesec, T.P., Gaffney, K., Lian, L., Fazio, V.W., Remzi, F.H., Lopez, R., Goldblum, J.R., and Shen, B. (2010). Pyloric gland metaplasia and pouchitis in patients with ileal pouch-anal anastomoses. Alimentary Pharmacology & Therapeutics, 31, 862-873.

Kaser, A., & Blumberg, R. S. (2010). Endoplasmic reticulum stress and intestinal inflammation. Mucosal Immunology, 3(1), 11–16.

Kaser, A., Lee, A. H., Franke, A., Glickman, J. N., Zeissig, S., Tilg, H., Nieuwenhuis, E. E., Higgins, D. E., Schreiber, S., Glimcher, L. H., & Blumberg, R. S. (2008). XBP1 links ER stress to intestinal inflammation and confers genetic risk for human inflammatory bowel disease. Cell, 134(5), 743–756.

Kashani, A., & Schwartz, D. A. (2019). The expanding role of anti-il-12 and/or anti-il-23 antibodies in the treatment of inflammatory bowel disease. Gastroenterology & Hepatology, 15(5), 255–265.

Kelly, S., Bliss, T. M., Shah, A. K., Sun, G. H., Ma, M., Foo, W. C., Masel, J., Yenari, M. A., Weissman, I. L., Uchida, N., Palmer, T., & Steinberg, G. K. (2004). Transplanted human fetal neural stem cells survive, migrate, and differentiate in ischemic rat cerebral cortex. Proceedings of the National Academy of Sciences, 101(32), 11839–11844.

Khor, B., Gardet, A., and Xavier, R.J. (2011). Genetics and pathogenesis of inflammatory bowel disease. Nature, 474, 307-317.

Kiesler, P., Fuss, I. J., & Strober, W. (2015). Experimental models of inflammatory bowel diseases. Cellular and Molecular Gastroenterology and Hepatology, 1(2), 154–170.

Kim, J., Koo, B. K., & Knoblich, J. A. (2020). Human organoids: model systems for human biology and medicine. Nature reviews. Molecular Cell Biology, 21(10), 571–584.

Kostic, A. D., Xavier, R. J., & Gevers, D. (2014). The microbiome in inflammatory bowel disease: current status and the future ahead. Gastroenterology, 146(6), 1489–1499.

Kotze, P. G., Ma, C., Almutairdi, A., Al-Darmaki, A., Devlin, S. M., Kaplan, G. G., Seow, C. H., Novak, K. L., Lu, C., Ferraz, J., Stewart, M. J., Buresi, M., Jijon, H., Mathivanan, M., Heatherington, J., Martin, M. L., & Panaccione, R. (2018). Real-world clinical, endoscopic and radiographic efficacy of vedolizumab for the treatment of inflammatory bowel disease. Alimentary Pharmacology & Therapeutics, 48(6), 626–637.

Kühn, R., Löhler, J., Rennick, D., Rajewsky, K., & Müller, W. (1993). Interleukin-10-deficient mice develop chronic enterocolitis. Cell, 75(2), 263–274.

Kumar, P. A., Hu, Y., Yamamoto, Y., Hoe, N. B., Wei, T. S., Mu, D., Sun, Y., Joo, L. S., Dagher, R., Zielonka, E. M., Wang, d., Lim, B., Chow, V. T., Crum, C. P., Xian, W., & McKeon, F. (2011). Distal airway stem cells yield alveoli in vitro and during lung regeneration following H1N1 influenza infection. Cell, 147(3), 525–538.

Lachmann, A., Torre, D., Keenan, A.B., Jagodnik, K.M., Lee, H.J., Wang, L., Silverstein, M.C., and Ma'ayan, A. (2018). Massive mining of publicly available RNA-seq data from human and mouse. Nature Communications, 9, 1366.

Lancaster, M. A., & Huch, M. (2019). Disease modelling in human organoids. Disease Models & Mechanisms, 12(7), dmm039347.

Lee, F.D. (1964). Pyloric metaplasia in the small intestine. The Journal of Pathology Bacteriology, 87, 267-277.

Levine, J. S., & Burakoff, R. (2011). Extraintestinal manifestations of inflammatory bowel disease. Gastroenterology & Hepatology, 7(4), 235–241.

Levine, A., Griffiths, A., Markowitz, J., Wilson, D. C., Turner, D., Russell, R. K., Fell, J., Ruemmele, F. M., Walters, T., Sherlock, M., Dubinsky, M., & Hyams, J. S. (2011). Pediatric modification of the Montreal classification for inflammatory bowel disease: the Paris classification. Inflammatory Bowel Diseases, 17(6), 1314–1321.

Li, C., & Kuemmerle, J. F. (2020). The fate of myofibroblasts during the development of fibrosis in Crohn's disease. Journal of Digestive Diseases, 21(6), 326–331.

Li, G. M., Zhang, C. L., Rui, R. P., Sun, B., & Guo, W. (2018). Bioinformatics analysis of common differential genes of coronary artery disease and ischemic cardiomyopathy. European Review for Medical and Pharmacological Sciences, 22(11), 3553–3569.

Liber, A.F. (1951). Aberrant pyloric glands in regional ileitis. American Medical Association Archives of Pathology, 51, 205-212.

Lindsay M. A. (2003). Target discovery. Nature Reviews. Drug Discovery, 2(10), 831–838.

Liu, J. Z., van Sommeren, S., Huang, H., Ng, S. C., Alberts, R., Takahashi, A., Ripke, S., Lee, J. C., Jostins, L., Shah, T., Abedian, S., Cheon, J. H., Cho, J., Dayani, N. E., Franke, L., Fuyuno, Y., Hart, A., Juyal, R. C., Juyal, G., Kim, W. H., … Weersma, R. K. (2015). Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. Nature Genetics, 47(9), 979–986.

Loftus E. V., Jr (2004). Clinical epidemiology of inflammatory bowel disease: Incidence, prevalence, and environmental influences. Gastroenterology, 126(6), 1504–1517.

Machida, K., Suemizu, H., Kawai, K., Ishikawa, T., Sawada, R., Ohnishi, Y., & Tsuchiya, T. (2009). Higher susceptibility of NOG mice to xenotransplanted tumors. The Journal of Toxicological Sciences, 34(1), 123–127.

Maerten, P., Shen, C., Colpaert, S., Liu, Z., Bullens, D. A., van Assche, G., Penninckx, F., Geboes, K., Vanham, G., Rutgeerts, P., & Ceuppens, J. L. (2004). Involvement of interleukin 18 in Crohn's disease: evidence from in vitro analysis of human gut inflammatory cells and from experimental colitis models. Clinical and Experimental Immunology, 135(2), 310–317.

Mahid, S. S., Minor, K. S., Soto, R. E., Hornung, C. A., & Galandiuk, S. (2006). Smoking and inflammatory bowel disease: a meta-analysis. Mayo Clinic Proceedings, 81(11), 1462–1471.

Malo, N., Hanley, J. A., Cerquozzi, S., Pelletier, J., & Nadon, R. (2006). Statistical practice in high-throughput screening data analysis. Nature Biotechnology, 24(2), 167–175.

Martis, E.A.; Radhakrishnan, R.; Badve, R.R. (2011). High-throughput screening: The hits and leads of drug discovery—An overview. Journal of Applied Pharmaceutical Science, 1, 2–10.

McCole D. F. (2014). IBD candidate genes and intestinal barrier regulation. Inflammatory Bowel Diseases, 20(10), 1829–1849.

McCormick F. (2015). KRAS as a therapeutic target. Clinical Cancer Research, 21(8), 1797–1801.

McGovern, D. P., Kugathasan, S., & Cho, J. H. (2015). Genetics of inflammatory bowel diseases. Gastroenterology, 149(5), 1163–1176.e2.

McNeish J. D. (2007). Stem cells as screening tools in drug discovery. Current Opinion in Pharmacology, 7(5), 515–520.

Medvedev, S. P., Shevchenko, A. I., & Zakian, S. M. (2010). Induced pluripotent stem cells: problems and advantages when applying them in regenerative medicine. Acta Naturae, 2(2), 18–28.

Michelini, E., Cevenini, L., Mezzanotte, L., Coppa, A., & Roda, A. (2010). Cell-based assays: fuelling drug discovery. Analytical and Bioanalytical Chemistry, 398(1), 227–238.

Moffat, J. G., Rudolph, J., & Bailey, D. (2014). Phenotypic screening in cancer drug discovery - past, present and future. Nature Reviews. Drug Discovery, 13(8), 588–602.

Molodecky, N. A., Soon, I. S., Rabi, D. M., Ghali, W. A., Ferris, M., Chernoff, G., Benchimol, E. I., Panaccione, R., Ghosh, S., Barkema, H. W., & Kaplan, G. G. (2012). Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. Gastroenterology, 142(1), 46–e30.

Mombaerts, P., Mizoguchi, E., Grusby, M. J., Glimcher, L. H., Bhan, A. K., & Tonegawa, S. (1993). Spontaneous development of inflammatory bowel disease in T cell receptor mutant mice. Cell, 75(2), 274–282.

Morrissey, P. J., Charrier, K., Braddy, S., Liggitt, D., & Watson, J. D. (1993). CD4+ T cells that express high levels of CD45RB induce wasting disease when transferred into congenic severe combined immunodeficient mice. Disease development is prevented by cotransfer of purified CD4+ T cells. The Journal of Experimental Medicine, 178(1), 237–244.

Mosier, D. E., Gulizia, R. J., Baird, S. M., & Wilson, D. B. (1988). Transfer of a functional human immune system to mice with severe combined immunodeficiency. Nature, 335(6187), 256–259.

Muegge, B. D., Kuczynski, J., Knights, D., Clemente, J. C., González, A., Fontana, L., Henrissat, B., Knight, R., & Gordon, J. I. (2011). Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. Science, 332(6032), 970–974.

Müller, A. M., & Dzierzak, E. A. (1993). ES cells have only a limited lymphopoietic potential after adoptive transfer into mouse recipients. Development, 118(4), 1343–1351.

Ng, S. C., Bernstein, C. N., Vatn, M. H., Lakatos, P. L., Loftus, E. V., Jr, Tysk, C., O'Morain, C., Moum, B., Colombel, J. F., & Epidemiology and Natural History Task Force of the International Organization of Inflammatory Bowel Disease (IOIBD) (2013). Geographical variability and environmental risk factors in inflammatory bowel disease. Gut, 62(4), 630–649.

Nestorowa, S., Hamey, F. K., Pijuan Sala, B., Diamanti, E., Shepherd, M., Laurenti, E., Wilson, N. K., Kent, D. G., & Göttgens, B. (2016). A single-cell resolution map of mouse hematopoietic stem and progenitor cell differentiation. Blood, 128(8), e20–e31.

Neurath M. F. (2014). Cytokines in inflammatory bowel disease. Nature Reviews. Immunology, 14(5), 329–342.

Neurath M. F. (2017). Current and emerging therapeutic targets for IBD. Nature Reviews. Gastroenterology & Hepatology, 14(5), 269–278.

Ni, J., Chen, S. F., & Hollander, D. (1996). Effects of dextran sulphate sodium on intestinal epithelial cells and intestinal lymphocytes. Gut, 39(2), 234–241.

Nighot, P., Al-Sadi, R., Rawat, M., Guo, S., Watterson, D. M., & Ma, T. (2015). Matrix metalloproteinase 9-induced increase in intestinal epithelial tight junction permeability contributes to the severity of experimental DSS colitis. American Journal of Physiology. Gastrointestinal and Liver Physiology, 309(12), G988–G997.

Ninomiya, M., Kiyoi, H., Ito, M., Hirose, Y., Ito, M., & Naoe, T. (2004). Retinoic acid syndrome in NOD/scid mice induced by injecting an acute promyelocytic leukemia cell line. Leukemia, 18(3), 442–448.

Nirmalanandhan, V. S., & Sittampalam, G. S. (2009). Stem cells in drug discovery, tissue engineering, and regenerative medicine: emerging opportunities and challenges. Journal of Biomolecular Screening, 14(7), 755–768.

Nolan G. P. (2007). What's wrong with drug screening today. Nature Chemical Biology, 3(4), 187–191.

Odenwald, M. A., & Turner, J. R. (2017). The intestinal epithelial barrier: a therapeutic target? Nature Reviews. Gastroenterology & Hepatology, 14(1), 9–21.

Ogura, Y., Bonen, D. K., Inohara, N., Nicolae, D. L., Chen, F. F., Ramos, R., Britton, H., Moran, T., Karaliuskas, R., Duerr, R. H., Achkar, J. P., Brant, S. R., Bayless, T. M., Kirschner, B. S., Hanauer, S. B., Nuñez, G., & Cho, J. H. (2001). A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. Nature, 411(6837), 603–606.

Overington, J. P., Al-Lazikani, B., & Hopkins, A. L. (2006). How many drug targets are there? Nature Reviews. Drug Discovery, 5(12), 993–996.

Palmer, C., Bik, E. M., DiGiulio, D. B., Relman, D. A., & Brown, P. O. (2007). Development of the human infant intestinal microbiota. PLoS Biology, 5(7), e177.

Panés, J., Bouzas, R., Chaparro, M., García-Sánchez, V., Gisbert, J. P., Martínez de Guereñu, B., Mendoza, J. L., Paredes, J. M., Quiroga, S., Ripollés, T., & Rimola, J. (2011). Systematic review: the use of ultrasonography, computed tomography and magnetic resonance imaging for the diagnosis, assessment of activity and abdominal complications of Crohn's disease. Alimentary Pharmacology & Therapeutics, 34(2), 125–145.

Parikh, K., Antanaviciute, A., Fawkner-Corbett, D., Jagielowicz, M., Aulicino, A., Lagerholm, C., Davis, S., Kinchen, J., Chen, H. H., Alham, N. K., Ashley, N., Johnson, E., Hublitz, P., Bao, L., Lukomska, J., Andev, R. S., Björklund, E., Kessler, B. M., Fischer, R., Goldin, R., … Simmons, A. (2019). Colonic epithelial cell diversity in health and inflammatory bowel disease. Nature, 567(7746), 49–55.

Park, S. W., Zhen, G., Verhaeghe, C., Nakagami, Y., Nguyenvu, L. T., Barczak, A. J., Killeen, N., & Erle, D. J. (2009). The protein disulfide isomerase AGR2 is essential for production of intestinal mucus. Proceedings of the National Academy of Sciences, 106(17), 6950–6955.

Park, J. H., Wacholder, S., Gail, M. H., Peters, U., Jacobs, K. B., Chanock, S. J., & Chatterjee, N. (2010). Estimation of effect size distribution from genome-wide association studies and implications for future discoveries. Nature Genetics, 42(7), 570–575.

Pearson, A. D., Eastham, E. J., Laker, M. F., Craft, A. W., & Nelson, R. (1982). Intestinal permeability in children with Crohn's disease and coeliac disease. British Medical Journal, 285(6334), 20–21.

Peeters, M., Joossens, S., Vermeire, S., Vlietinck, R., Bossuyt, X., & Rutgeerts, P. (2001). Diagnostic value of anti-Saccharomyces cerevisiae and antineutrophil cytoplasmic autoantibodies in inflammatory bowel disease. The American Journal of Gastroenterology, 96(3), 730–734.

Peloquin, J. M., Goel, G., Villablanca, E. J., & Xavier, R. J. (2016). Mechanisms of pediatric inflammatory bowel disease. Annual Review of Immunology, 34, 31–64.

Peyrin-Biroulet, L., Loftus, E. V., Jr, Colombel, J. F., & Sandborn, W. J. (2010). The natural history of adult Crohn's disease in population-based cohorts. The American Journal of Gastroenterology, 105(2), 289–297.

Philpott, D. J., & Viala, J. (2004). Towards an understanding of the role of NOD2/CARD15 in the pathogenesis of Crohn's disease. Best practice & research. Clinical Gastroenterology, 18(3), 555–568.

Pizarro, T. T., Arseneau, K. O., Bamias, G., & Cominelli, F. (2003). Mouse models for the study of Crohn's disease. Trends in Molecular Medicine, 9(5), 218–222.

Poppe, D., Tiede, I., Fritz, G., Becker, C., Bartsch, B., Wirtz, S., Strand, D., Tanaka, S., Galle, P. R., Bustelo, X. R., & Neurath, M. F. (2006). Azathioprine suppresses ezrin-radixin-moesin-dependent T cell-APC conjugation through inhibition of Vav guanosine exchange activity on Rac proteins. Journal of Immunology, 176(1), 640–651.

Prasad, S., Mingrino, R., Kaukinen, K., Hayes, K. L., Powell, R. M., MacDonald, T. T., & Collins, J. E. (2005). Inflammatory processes have differential effects on claudins 2, 3 and 4 in colonic epithelial cells. Laboratory Investigation, 85(9), 1139–1162.

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K. S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., Yamada, T., Mende, D. R., Li, J., Xu, J., Li, S., Li, D., Cao, J., Wang, B., Liang, H., Zheng, H., Xie, Y., … Wang, J. (2010). A human gut microbial gene catalogue established by metagenomic sequencing. Nature, 464(7285), 59–65.

Rakoff-Nahoum, S., Paglino, J., Eslami-Varzaneh, F., Edberg, S., & Medzhitov, R. (2004). Recognition of commensal microflora by toll-like receptors is required for intestinal homeostasis. Cell, 118(2), 229–241.

Rama, P., Matuska, S., Paganoni, G., Spinelli, A., De Luca, M., & Pellegrini, G. (2010). Limbal stem-cell therapy and long-term corneal regeneration. The New England Journal of Medicine, 363(2), 147–155.

Rao, W., Wang, S., Duleba, M., Niroula, S., Goller, K., Xie, J., Mahalingam, R., Neupane, R., Liew, A. A., Vincent, M., Okuda, K., O'Neal, W. K., Boucher, R. C., Dickey, B. F., Wechsler, M. E., Ibrahim, O., Engelhardt, J. F., Mertens, T., Wang, W., Jyothula, S., … Xian, W. (2020). Regenerative metaplastic clones in copd lung drive inflammation and fibrosis. Cell, 181(4), 848–864.e18.

Raychaudhuri, S., Plenge, R. M., Rossin, E. J., Ng, A. C., International Schizophrenia Consortium, Purcell, S. M., Sklar, P., Scolnick, E. M., Xavier, R. J., Altshuler, D., & Daly, M. J. (2009). Identifying relationships among genomic disease regions: predicting genes at pathogenic SNP associations and rare deletions. PLoS Genetics, 5(6), e1000534.

Regueiro, M., Feagan, B. G., Zou, B., Johanns, J., Blank, M. A., Chevrier, M., Plevy, S., Popp, J., Cornillie, F. J., Lukas, M., Danese, S., Gionchetti, P., Hanauer, S. B., Reinisch, W., Sandborn, W. J., Sorrentino, D., Rutgeerts, P., & PREVENT Study Group (2016). Infliximab reduces endoscopic, but not clinical, recurrence of crohn's disease after ileocolonic resection. Gastroenterology, 150(7), 1568–1578.

Rheinwald, J. G., & Green, H. (1975). Serial cultivation of strains of human epidermal keratinocytes: the formation of keratinizing colonies from single cells. Cell, 6(3), 331–343.

Rieder, F., Zimmermann, E. M., Remzi, F. H., & Sandborn, W. J. (2013). Crohn's disease complicated by strictures: a systematic review. Gut, 62(7), 1072–1084.

Rieder, F., Latella, G., Magro, F., Yuksel, E. S., Higgins, P. D., Di Sabatino, A., de Bruyn, J. R., Rimola, J., Brito, J., Bettenworth, D., van Assche, G., Bemelman, W., d'Hoore, A., Pellino, G., & Dignass, A. U. (2016). European Crohn's and colitis organisation topical review on prediction, diagnosis and management of fibrostenosing Crohn's disease. Journal of Crohn's & Colitis, 10(8), 873–885.

Rioux, J. D., Xavier, R. J., Taylor, K. D., Silverberg, M. S., Goyette, P., Huett, A., Green, T., Kuballa, P., Barmada, M. M., Datta, L. W., Shugart, Y. Y., Griffiths, A. M., Targan, S. R., Ippoliti, A. F., Bernard, E. J., Mei, L., Nicolae, D. L., Regueiro, M., Schumm, L. P., Steinhart, A. H., … Brant, S. R. (2007). Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis. Nature Genetics, 39(5), 596–604.

Rishton G. M. (2008). Small molecules that promote neurogenesis in vitro. Recent Patents on CNS Drug Discovery, 3(3), 200–208.

Ritz, C., Baty, F., Streibig, J. C., & Gerhard, D. (2015). Dose-response analysis using R. PloS One, 10(12), e0146021.

Roche, K.C., Gracz, A.D., Liu, X.F., Newton, V., Akiyama, H., and Magness, S.T. (2015). SOX9 maintains reserve stem cells and preserves radioresistance in mouse small intestine. Gastroenterology, 149, 1553-1563 e1510.

Rose, S., Misharin, A., & Perlman, H. (2012). A novel Ly6C/Ly6G-based strategy to analyze the mouse splenic myeloid compartment. Cytometry. Part A : The Journal of the International Society for Analytical Cytology, 81(4), 343–350.

Rudolph, U., Finegold, M. J., Rich, S. S., Harriman, G. R., Srinivasan, Y., Brabet, P., Boulay, G., Bradley, A., & Birnbaumer, L. (1995). Ulcerative colitis and adenocarcinoma of the colon in G alpha i2-deficient mice. Nature Genetics, 10(2), 143–150.

Rutgeerts, P., Geboes, K., Vantrappen, G., Kerremans, R., Coenegrachts, J.L., and Coremans, G. (1984). Natural history of recurrent Crohn's disease at the ileocolonic anastomosis after curative surgery. Gut, 25(6), 665-672.

Sadlack, B., Merz, H., Schorle, H., Schimpl, A., Feller, A. C., & Horak, I. (1993). Ulcerative colitis-like disease in mice with a disrupted interleukin-2 gene. Cell, 75(2), 253–261.

Saidel-Odes, L., Borer, A., & Odes, S. (2011). Clostridium difficile infection in patients with inflammatory bowel disease. Annals of Gastroenterology, 24(4), 263–270.

Salzman, N. H., Hung, K., Haribhai, D., Chu, H., Karlsson-Sjöberg, J., Amir, E., Teggatz, P., Barman, M., Hayward, M., Eastwood, D., Stoel, M., Zhou, Y., Sodergren, E., Weinstock, G. M., Bevins, C. L., Williams, C. B., & Bos, N. A. (2010). Enteric defensins are essential regulators of intestinal microbial ecology. Nature Immunology, 11(1), 76–83.

Sandborn, W. J., Feagan, B. G., Rutgeerts, P., Hanauer, S., Colombel, J. F., Sands, B. E., Lukas, M., Fedorak, R. N., Lee, S., Bressler, B., Fox, I., Rosario, M., Sankoh, S., Xu, J., Stephens, K., Milch, C., Parikh, A., & GEMINI 2 Study Group (2013). Vedolizumab as induction and maintenance therapy for Crohn's disease. The New England Journal of Medicine, 369(8), 711–721.

Sandborn, W. J., Gasink, C., Gao, L. L., Blank, M. A., Johanns, J., Guzzo, C., Sands, B. E., Hanauer, S. B., Targan, S., Rutgeerts, P., Ghosh, S., de Villiers, W. J., Panaccione, R., Greenberg, G., Schreiber, S., Lichtiger, S., Feagan, B. G., & CERTIFI Study Group (2012). Ustekinumab induction and maintenance therapy in refractory Crohn's disease. The New England Journal of Medicine, 367(16), 1519–1528.

Satija, R., Farrell, J. A., Gennert, D., Schier, A. F., & Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. Nature Biotechnology, 33(5), 495–502.

Sauer, C.G., and Kugathasan, S. (2009). Pediatric inflammatory bowel disease: highlighting pediatric differences in IBD. Gastroenterology Clinics of North America, 38, 611-628.

Scharl, M., & Rogler, G. (2014). Pathophysiology of fistula formation in Crohn's disease. World Journal of Gastrointestinal Pathophysiology, 5(3), 205–212.

Sedda, S., Marafini, I., Dinallo, V., Di Fusco, D., & Monteleone, G. (2015). The TGF-β/Smad system in IBD pathogenesis. Inflammatory Bowel Diseases, 21(12), 2921–2925.

Senoo, M., Pinto, F., Crum, C. P., & McKeon, F. (2007). p63 Is essential for the proliferative potential of stem cells in stratified epithelia. Cell, 129(3), 523–536.

Seow, C. H., Benchimol, E. I., Griffiths, A. M., Otley, A. R., & Steinhart, A. H. (2008). Budesonide for induction of remission in Crohn's disease. The Cochrane Database of Systematic Reviews, (3), CD000296.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Research, 13, 2498-2504.

Sheehan, A. L., Warren, B. F., Gear, M. W., & Shepherd, N. A. (1992). Fat-wrapping in Crohn's disease: pathological basis and relevance to surgical practice. The British Journal of Surgery, 79(9), 955–958.

Shiokawa, M., Takahashi, T., Murakami, A., Kita, S., Ito, M., Sugamura, K., & Ishii, N. (2010). In vivo assay of human NK-dependent ADCC using NOD/SCID/gammac(null) (NOG) mice. Biochemical and Biophysical Research Communications, 399(4), 733–737.

Shultz, L.D., Brehm, M.A., Garcia-Martinez, J.V., and Greiner, D.L. (2012). Humanized mice for immune system investigation: progress, promise and challenges. Nature Reviews. Immunology, 12, 786-798.

Shultz, L.D., Lyons, B.L., Burzenski, L.M., Gott, B., Chen, X., Chaleff, S., Kotb, M., Gillies, S.D., King, M., Mangada, J.*, et al.* (2005). Human lymphoid and myeloid cell development in NOD/LtSz-scid IL2R gamma null mice engrafted with mobilized human hemopoietic stem cells. The Journal of Immunology, 174, 6477-6489.

Shultz, L.D., Schweitzer, P.A., Christianson, S.W., Gott, B., Schweitzer, I.B., Tennent, B., McKenna, S., Mobraaten, L., Rajan, T.V., Greiner, D.L.*, et al.* (1995). Multiple defects in innate and adaptive immunologic function in NOD/LtSz-scid mice. The Journal of Immunology, 154, 180-191.

Silverberg, M. S., Satsangi, J., Ahmad, T., Arnott, I. D., Bernstein, C. N., Brant, S. R., Caprilli, R., Colombel, J. F., Gasche, C., Geboes, K., Jewell, D. P., Karban, A., Loftus, E. V., Jr, Peña, A. S., Riddell, R. H., Sachar, D. B., Schreiber, S., Steinhart, A. H., Targan, S. R., Vermeire, S., … Warren, B. F. (2005). Toward an integrated clinical, molecular and serological classification of inflammatory bowel disease: report of a Working Party of the 2005 Montreal World Congress of Gastroenterology. Canadian Journal of Gastroenterology, 19 Suppl A, 5A–36A.

Siolas, D., & Hannon, G. J. (2013). Patient-derived tumor xenografts: transforming clinical samples into mouse models. Cancer Research, 73(17), 5315–5319.

Slenter, D. N., Kutmon, M., Hanspers, K., Riutta, A., Windsor, J., Nunes, N., Mélius, J., Cirillo, E., Coort, S. L., Digles, D., Ehrhart, F., Giesbertz, P., Kalafati, M., Martens, M., Miller, R., Nishida, K., Rieswijk, L., Waagmeester, A., Eijssen, L., Evelo, C. T., … Willighagen, E. L. (2018). WikiPathways: a multifaceted pathway database bridging metabolomics to other omics research. Nucleic Acids Research, 46(D1), D661–D667.

Sokol, H., Pigneur, B., Watterlot, L., Lakhdari, O., Bermúdez-Humarán, L. G., Gratadoux, J. J., Blugeon, S., Bridonneau, C., Furet, J. P., Corthier, G., Grangette, C., Vasquez, N., Pochart, P., Trugnan, G., Thomas, G., Blottière, H. M., Doré, J., Marteau, P., Seksik, P., & Langella, P. (2008). Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. Proceedings of the National Academy of Sciences, 105(43), 16731–16736.

Spehlmann, M. E., Begun, A. Z., Burghardt, J., Lepage, P., Raedler, A., & Schreiber, S. (2008). Epidemiology of inflammatory bowel disease in a German twin cohort: results of a nationwide study. Inflammatory Bowel Diseases, 14(7), 968–976.

Stappenbeck, T. S., & McGovern, D. (2017). Paneth cell alterations in the development and phenotype of Crohn's disease. Gastroenterology, 152(2), 322–326.

Steinhart, A. H., Ewe, K., Griffiths, A. M., Modigliani, R., & Thomsen, O. O. (2000). Corticosteroids for maintaining remission of Crohn's disease. The Cochrane Database of Systematic Reviews, (2), CD000301.

Steinhoff, U., Brinkmann, V., Klemm, U., Aichele, P., Seiler, P., Brandt, U., Bland, P. W., Prinz, I., Zügel, U., & Kaufmann, S. H. (1999). Autoimmune intestinal pathology induced by hsp60-specific CD8 T cells. Immunity, 11(3), 349–358.

Sun, Q., Zhang, Z., & Sun, Z. (2014). The potential and challenges of using stem cells for cardiovascular repair and regeneration. Genes & Diseases, 1(1), 113–119.

Sundberg S. A. (2000). High-throughput and ultra-high-throughput screening: solution- and cell-based approaches. Current Opinion in Biotechnology, 11(1), 47–53.

Sundberg, J. P., Cordy, W. R., & King, L. E., Jr (1994). Alopecia areata in aging C3H/HeJ mice. The Journal of Investigative Dermatology, 102(6), 847–856.

Suzuki, N., Yamazaki, S., Yamaguchi, T., Okabe, M., Masaki, H., Takaki, S., Otsu, M., & Nakauchi, H. (2013). Generation of engraftable hematopoietic stem cells from induced pluripotent stem cells by way of teratoma formation. Molecular Therapy : The Journal of the American Society of Gene Therapy, 21(7), 1424–1431.

Suzuki, K., Yokoyama, J., Kawauchi, Y., Honda, Y., Sato, H., Aoyagi, Y., Terai, S., Okazaki, K., Suzuki, Y., Sameshima, Y., Fukushima, T., Sugahara, K., Atreya, R., Neurath, M. F., Watanabe, K., Yoneyama, H., & Asakura, H. (2017). Phase 1 clinical study of siRNA targeting carbohydrate sulphotransferase 15 in Crohn's disease patients with active mucosal lesions. Journal of Crohn's & Colitis, 11(2), 221–228.

Swaminath, A., & Kornbluth, A. (2007). Optimizing drug therapy in inflammatory bowel disease. Current Gastroenterology Reports, 9(6), 513–520.

Swinney, D. C., & Anthony, J. (2011). How were new medicines discovered? Nature Reviews. Drug Discovery, 10(7), 507–519.

Szymański, P., Markowicz, M., & Mikiciuk-Olasik, E. (2012). Adaptation of high-throughput screening in drug discovery-toxicological screening tests. International Journal of Molecular Sciences, 13(1), 427–452.

Thompson, N., Adams, D. J., & Ranzani, M. (2017). Synthetic lethality: emerging targets and opportunities in melanoma. Pigment Cell & Melanoma Research, 30(2), 183–193.

Thorsvik, S., van Beelen Granlund, A., Svendsen, T. D., Bakke, I., Røyset, E. S., Flo, T. H., Damås, J. K., Østvik, A. E., Bruland, T., & Sandvik, A. K. (2019). Ulcer-associated cell lineage expresses genes involved in regeneration and is hallmarked by high neutrophil gelatinase-associated lipocalin (NGAL) levels. The Journal of Pathology, 248(3), 316–325.

Tokuhira, N., Kitagishi, Y., Suzuki, M., Minami, A., Nakanishi, A., Ono, Y., Kobayashi, K., Matsuda, S., & Ogura, Y. (2015). PI3K/AKT/PTEN pathway as a target for Crohn's disease therapy (Review). International Journal of Molecular Medicine, 35(1), 10–16.

Tuvlin, J. A., Raza, S. S., Bracamonte, S., Julian, C., Hanauer, S. B., Nicolae, D. L., King, A. C., & Cho, J. H. (2007). Smoking and inflammatory bowel disease: trends in familial and sporadic cohorts. Inflammatory Bowel Diseases, 13(5), 573–579.

Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., Olsson, I., Edlund, K., Lundberg, E., Navani, S., Szigyarto, C. A., Odeberg, J., Djureinovic, D., Takanen, J. O., Hober, S., Alm, T., … Pontén, F. (2015). Proteomics. Tissue-based map of the human proteome. Science, 347(6220), 1260419.

Ukabam, S. O., Clamp, J. R., & Cooper, B. T. (1983). Abnormal small intestinal permeability to sugars in patients with Crohn's disease of the terminal ileum and colon. Digestion, 27(2), 70–74.

van der Maaten L and Hinton, G. Visualizing data using t-SNE (2008). Journal of Machine Learning Research, 9, 2579-2605.

Van Limbergen, J., Russell, R. K., Nimmo, E. R., Drummond, H. E., Smith, L., Anderson, N. H., Davies, G., Gillett, P. M., McGrogan, P., Weaver, L. T., Bisset, W. M., Mahdi, G., Arnott, I. D., Wilson, D. C., & Satsangi, J. (2008). Autophagy gene ATG16L1 influences susceptibility and disease location but not childhood-onset in Crohn's disease in northern Europe. Inflammatory Bowel Diseases, 14(3), 338–346.

Vijay-Kumar, M., Sanders, C. J., Taylor, R. T., Kumar, A., Aitken, J. D., Sitaraman, S. V., Neish, A. S., Uematsu, S., Akira, S., Williams, I. R., & Gewirtz, A. T. (2007). Deletion of TLR5 results in spontaneous colitis in mice. The Journal of Clinical Investigation, 117(12), 3909–3921.

Wang, X., Yamamoto, Y., Wilson, L. H., Zhang, T., Howitt, B. E., Farrow, M. A., Kern, F., Ning, G., Hong, Y., Khor, C. C., Chevalier, B., Bertrand, D., Wu, L., Nagarajan, N., Sylvester, F. A., Hyams, J. S., Devers, T., Bronson, R., Lacy, D. B., Ho, K. Y., … Xian, W. (2015). Cloning and variation of ground state intestinal stem cells. Nature, 522(7555), 173–178.

Wehkamp, J., Salzman, N. H., Porter, E., Nuding, S., Weichenthal, M., Petras, R. E., Shen, B., Schaeffeler, E., Schwab, M., Linzmeier, R., Feathers, R. W., Chu, H., Lima, H., Jr, Fellermann,

K., Ganz, T., Stange, E. F., & Bevins, C. L. (2005). Reduced Paneth cell alpha-defensins in ileal Crohn's disease. Proceedings of the National Academy of Sciences, 102(50), 18129–18134.

Williams M. (2005). Systems and integrative biology as alternative guises for pharmacology: prime time for an iPharm concept? Biochemical Pharmacology, 70(12), 1707–1716.

Wright, N. A., Pike, C., & Elia, G. (1990). Induction of a novel epidermal growth factor-secreting cell lineage by mucosal ulceration in human gastrointestinal stem cells. Nature, 343(6253), 82–85.

Williams, B. B., Tebbutt, N. C., Buchert, M., Putoczki, T. L., Doggett, K., Bao, S., Johnstone, C. N., Masson, F., Hollande, F., Burgess, A. W., Scott, A. M., Ernst, M., & Heath, J. K. (2015). Glycoprotein A33 deficiency: a new mouse model of impaired intestinal epithelial barrier function and inflammatory disease. Disease Models & Mechanisms, 8(8), 805–815.

Wirtz, S., Neufert, C., Weigmann, B., & Neurath, M. F. (2007). Chemically induced mouse models of intestinal inflammation. Nature Protocols, 2(3), 541–546.

Wodehouse, T., Demopoulos, M., Petty, R., Miraki-Moud, F., Belhaj, A., Husband, M., Fulton, L., Randive, N., Oksche, A., Mehta, V., Gribben, J., & Langford, R. (2019). A randomized pilot study to investigate the effect of opioids on immunomarkers using gene expression profiling during surgery. Pain, 160(12), 2691–2698.

Wu, G., and Haw, R. (2017). Functional interaction network construction and analysis for disease discovery. Methods in Molecular Biology, 1558, 235-253.

Wyatt, J., Vogelsang, H., Hübl, W., Waldhöer, T., & Lochs, H. (1993). Intestinal permeability and the prediction of relapse in Crohn's disease. Lancet, 341(8858), 1437–1439.

Yamada, Y., Marshall, S., Specian, R. D., & Grisham, M. B. (1992). A comparative analysis of two models of colitis in rats. Gastroenterology, 102(5), 1524–1534.

Yoo, J. H., Holubar, S., & Rieder, F. (2020). Fibrostenotic strictures in Crohn's disease. Intestinal Research, 18(4), 379–401.

Zeileis, A. & Hothorn, T. (2002). Diagnostic checking in regression relationships. The R Journal, 2, 7-10.

Zeileis, A. (2004). Econometric computing with HC and HAC covariance matrix estimators. Journal of Statistical Software, 1, 1-17.

Zeissig, S., Bürgel, N., Günzel, D., Richter, J., Mankertz, J., Wahnschaffe, U., Kroesen, A. J., Zeitz, M., Fromm, M., & Schulzke, J. D. (2007). Changes in expression and distribution of claudin 2, 5 and 8 lead to discontinuous tight junctions and barrier dysfunction in active Crohn's disease. Gut, 56(1), 61–72.

Zhao, F., Edwards, R., Dizon, D., Afrasiabi, K., Mastroianni, J. R., Geyfman, M., Ouellette, A. J., Andersen, B., & Lipkin, S. M. (2010). Disruption of Paneth and goblet cell homeostasis and

increased endoplasmic reticulum stress in Agr2-/- mice. Developmental Biology, 338(2), 270–279.

Zhang, L., Meissner, E., Chen, J., & Su, L. (2010). Current humanized mouse models for studying human immunology and HIV-1 immuno-pathogenesis. Science China. Life Sciences, 53(2), 195–203.

Zhang, J. H., Chung, T. D., & Oldenburg, K. R. (1999). A simple statistical parameter for use in evaluation and validation of high throughput screening assays. Journal of Biomolecular Screening, 4(2), 67–73.

Zhou, L., Lopes, J. E., Chong, M. M., Ivanov, I. I., Min, R., Victora, G. D., Shen, Y., Du, J., Rubtsov, Y. P., Rudensky, A. Y., Ziegler, S. F., & Littman, D. R. (2008). TGF-beta-induced Foxp3 inhibits T(H)17 cell differentiation by antagonizing RORgammat function. Nature, 453(7192), 236–240.

Zhu, F., Nair, R. R., Fisher, E., & Cunningham, T. J. (2019). Humanizing the mouse genome piece by piece. Nature Communications, 10(1), 1845.