

UCLA

UCLA Electronic Theses and Dissertations

Title

Molecular Mechanisms in Irritable Bowel Syndrome

Permalink

<https://escholarship.org/uc/item/2wb0z1q6>

Author

Videlock, Elizabeth Jane

Publication Date

2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Molecular Mechanisms in Irritable Bowel Syndrome

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy
in Molecular, Cellular, and Integrative Physiology

by

Elizabeth Jane Videlock

2019

© Copyright by

Elizabeth Jane Videlock

2019

ABSTRACT OF THE DISSERTATION

Molecular Mechanisms of Irritable Bowel Syndrome

by

Elizabeth Jane Videlock

Doctor of Philosophy in Molecular, Cellular and Integrative Physiology

University of California, Los Angeles, 2019

Professor Charalabos Pothoulakis, Chair

Irritable bowel syndrome (IBS) is a common functional gastrointestinal disorder characterized by abdominal pain associated with alteration in stool form and/or frequency. The pathophysiology of IBS is poorly understood. Although structural or histological changes in the gastrointestinal tract are absent in IBS, changes in intestinal function (e.g. increased intestinal permeability, increased neuronal sensitivity) and changes at the molecular level (e.g. gene methylation, gene/protein expression) have been described.

This dissertation describes a translational research project which began with data from gastrointestinal biopsies and concluded with studies in cell culture models. We performed a microarray gene expression profiling analysis of colonic mucosal biopsies in patients with IBS and healthy controls. I analyzed the differential gene and described the pathways involved in the regulation of IBS. These studies have direct clinical impact to patient care. Since data in the literature suggests a role for epigenetic changes (changes affecting gene expression) in IBS, I next focused on the role of differentially expressed long non-coding RNAs in IBS. The long non-

coding RNA afadin-divergent transcript (*AFDN-DT*), which was up-regulated in IBS, was selected for further *in vitro* studies. I generated intestinal epithelial cell lines capable of inducible expression of *AFDN-DT* and tested the hypothesis that overexpression of *AFDN-DT* resulted in altered epithelial permeability.

The dissertation of Elizabeth Jane Videlock is approved.

Lin Chang

Emeran A. Mayer

Joseph R. Pisegna

Charalabos Pothoulakis, Committee Chair

University of California, Los Angeles

2019

DEDICATION

To my parents. My love of science and education comes from you.

TABLE OF CONTENTS

List of tables and figures	viii
Acknowledgments	xi
Biographical sketch	xv
Chapter 1: Introduction	1
Abstract.....	2
Irritable bowel syndrome.....	3
Summary	21
Introduction to dissertation research.....	21
References:	23
Chapter 2: Sigmoid colon mucosal gene expression supports alterations of neuronal signaling in irritable bowel syndrome with constipation	39
Abstract.....	40
Introduction	41
Materials and Methods	42
Results	47
Discussion	61
Conclusions	67
References:	68
Chapter 3: Differential Expression of Long non-Coding RNAs in Irritable Bowel Syndrome.....	80
Abstract.....	81
Introduction	82
Methods	84
Results	86
Discussion	100
References:	106
Chapter 4 The Role of the lncRNA <i>AFDN-DT</i> in epithelial permeability.....	113
Abstract.....	114
Introduction	115
Methods	117
Results	125
Discussion	148
References:	154
Chapter 5: Gene Expression Profiling of the Gut-Brain Axis in Healthy Deceased Donors	158
Abstract.....	159
Introduction	160

Methods	163
Results	165
Discussion	172
References:	178
Chapter 6: Summary and Future Directions	184
References:	187
Appendix: R code	188

LIST OF TABLES AND FIGURES

Chapter 1

Tables

Table 1: Diagnostic criteria for IBS	6
Table 2: IBS subtyping	6

Figures

Figure 1: Bowel habit subtypes	7
Figure 2: Interrelatedness of IBS pathophysiology concepts	8
Figure 3: Enhanced visceral perception	10

Chapter 2

Tables

Table 1: Participant characteristics	48
Table 2: External sample and study characteristics	49
Table 3: Genes differentially expressed in IBS-C vs. HC concordant in Mayo	51
Table 4: Gene ontology terms overrepresented among genes in the cAMP/PKA module	54
Table 5: Description of IBS-associated modules	56
Table 6: Correlation of cAMP-related genes with module eigengenes	57
Table 7: Results of nCounter validation	60
Table 8: Comparison of findings from current study for genes found to be differentially expressed in the literature	62
Table 9: Differentially expressed tight junction genes	63
Table 10: Expression of neuronal and glial/astrocytic markers	66

Figures

Figure 1: Heatmaps of differentially expressed genes	50
Figure 2: WGCNA modules associated with IBS	53
Figure 3: IPA pathway analysis network for the cAMP/PKA module	55
Figure 4: cAMP signaling roles for hub genes in IBS-associated modules	57
Figure 5: nCounter expression of 9 validated genes in IBS-C vs. HC	59

Chapter 3

Tables

Table 1: Primer sequences	85
Table 2: Differentially expressed lncRNAs in IBS vs. HCs	86
Table 3: Fold change (p value) of candidate lncRNAs	87
Table 4: Top ten gene ontology terms associated with the “DNA repair” module	89
Table 5: “DNA repair” genes correlated with <i>GHRLOS</i> and IBS-C	90
Table 6: <i>GHRLOS</i> PCR sample demographics	91
Table 7: Differentially expressed adherens junction genes	94
Table 8: Adherens junction genes PCR sample demographics	95
Table 9: <i>AFDN-DT</i> PCR cohort demographics	97

Figures

Figure 1: Correlation of candidate lncRNAs with IBS-associated mRNA modules	88
Figure 2: <i>GHRLOS</i> is down-regulated in IBS	89
Figure 3: <i>GHRLOS</i> microarray probe and PCR primer alignments	89
Figure 4: <i>GHRLOS</i> and <i>MLN</i> are negatively correlated	91
Figure 5: <i>GHRLOS</i> PCR	92
Figure 6: <i>GHRLOS</i> gene expression by <i>in-situ</i> hybridization	92
Figure 7: Cellular junction formation	94
Figure 8: Expression of adherens junction genes by microarray and RT-PCR	95
Figure 9: Afadin divergent transcript (<i>AFDN-DT</i>)	96
Figure 10: <i>AFDN-DT</i> microarray differential expression and correlation with <i>AFDN</i>	97
Figure 11: PCR validation of <i>AFDN-DT</i>	98
Figure 12: PCR correlation of <i>AFDN-DT</i> and <i>AFDN</i>	98
Figure 13: <i>AFDN-DT</i> expression by <i>in situ</i> hybridization	99
Figure 14: Hypothesized role of <i>GHRLOS</i>	102
Figure 15: Mechanisms of antisense transcripts	103

Chapter 4

Tables

Table 1: Primer sequences	121
Table 2: RNA quality and reads	137
Table 3: Summary of differential expression	138
Table 4: Differential gene expression in T84-AFDN-DT with vs. without doxycycline	139
Table 5: Top genes in modules correlated with doxycycline in T84-TR-AFDN-DT	144

Table 6: GSEA results for T84-TR-AFDN-DT vs. T84-TR-EV	146
Table 7: Comparison of cell line gene set enrichment by doxycycline treatment.....	147

Figures

Figure 1: <i>AFDN-DT</i> vector maps	119
Figure 2: PCR verification of inducible <i>AFDN-DT</i> overexpression	126
Figure 3: Effect of <i>AFDN-DT</i> overexpression on TEER in differentiated cells	127
Figure 4: Pooled results for TEER response to <i>AFDN-DT</i> overexpression.....	127
Figure 5: Effect of <i>AFDN-DT</i> overexpression during differentiation on TEER.....	128
Figure 6: Effect of <i>AFDN-DT</i> overexpression on TEER response to IFN- γ	130
Figure 7: Pooled results for the effect of <i>AFDN-DT</i> overexpression on the TEER response to IFN- γ	131
Figure 8: Effect of <i>AFDN-DT</i> overexpression on paracellular permeability	132
Figure 9: Inducible <i>AFDN-DT</i> overexpression in Caco2-TR-AFDN-DT	133
Figure 10: Effect of <i>AFDN-DT</i> overexpression on TEER of Caco2 monolayers	134
Figure 11: Effect of <i>AFDN-DT</i> overexpression on TEER response to IFN- γ and TNF- α in Caco2	135
Figure 12: Effect of <i>AFDN-DT</i> overexpression on paracellular permeability in Caco2.....	136
Figure 13: RNA-seq principal components plot.....	138
Figure 14: Plots of differentially expressed genes in T84-AFDN-DT with vs. without doxycycline	139
Figure 15: WGCNA, single network	141
Figure 16: WGCNA, consensus networks.....	142

Chapter 5

Tables

Table 1: Analyses and aims	162
Table 2: Immune gut-brain-axis genes differentially expressed in IBS-D (Mayo Clinic cohort) 171	

Figures

Figure 1: The gut-brain axis	160
Figure 2: Outline of network analyses	162
Figure 3: Donor characteristics	165
Figure 4: Tissue-specific networks	166
Figure 5: Consensus networks.....	167
Figure 6: Unified network	168
Figure 7: Inflammation hub genes.....	169

ACKNOWLEDGMENTS

Chapter 1 contains text that is reprinted or derived from Videlock EJ, Chang L. “Irritable Bowel Syndrome”. In: Camilleri M, Podolsky D, editors. *Yamada’s Textbook of Gastroenterology*. 6 ed. Oxford, United Kingdom: Wiley-Blackwell; Copyright 2014, with permission from John Wiley & Sons, Ltd.

Chapter 1, Table 1 is reprinted from Videlock EJ, Chang L. “Irritable Bowel Syndrome”. In: Talley NJ, Kane SV, Wallace MB, editors. *Practical Gastroenterology and Hepatology: Small and Large Intestine and Pancreas*. 1 ed. Oxford, United Kingdom: Wiley-Blackwell; Copyright 2010, with permission from John Wiley & Sons, Ltd.

Chapter 1, Figure 1 is reprinted from *Gastroenterology*:150(6), Lacy BE et al., Bowel Disorders, page 1396; Copyright 2016, with permission from Elsevier and the Rome Foundation.

Chapter 1, Figure 3 is reprinted from Videlock EJ, Chang L. “Irritable Bowel Syndrome”. In: Camilleri M, Podolsky D, editors. *Yamada’s Textbook of Gastroenterology*. 6 ed. Oxford, United Kingdom: Wiley-Blackwell; Copyright 2014, with permission from John Wiley & Sons, Ltd.

The text of Chapter 2 is the submission version and tables and figures are reprinted from Videlock EJ, Mahurkar-Joshi S, Hoffman JM, Iliopoulos D, Pothoulakis C, Mayer EA, Chang L. Sigmoid colon mucosal gene expression supports alterations of neuronal signaling in irritable bowel syndrome with constipation. *American Journal of Physiology – Gastrointestinal & Liver Physiology*. 2018;315(1):G140-G157, Copyright 2018, with permission from the American Physiological Society.

Chapter 3, Figure 7 is modified from Current Opinion in Cell Biology, Vol. 19, Issue 5, Sakisaka et al, The roles of nectins in cell adhesions: cooperation with other cell adhesion molecules and growth factor receptors, pages 593-602, Copyright 2007, with permission from Elsevier.

Chapter 3, Figure 15 is reprinted from Villegas VE and Zaphiropoulos PG, Neighboring gene regulation by antisense long non-coding RNAs. Int J Mol Sci, 2015. 16(2): 3251-66. Licensed under CC BY 4.0.

Chapter 5, Figure 1 is modified from Kennedy PJ, Clarke G, Quigley EM, Groeger JA, Dinan TG, Cryan JF. Gut memories: towards a cognitive neurobiology of irritable bowel syndrome. Neuroscience & Biobehavioral Reviews, Vol 36, Issue 1, pages 310-340, Copyright 2012, with permission from Elsevier.

This work was supported by the National Institutes of Health (NIH) P50 DK64539 (EAM, LC), T32 DK07180 (EJV), the Barbara and Joel Marcus Foundation GI Fellowship Seed Grant (EJV), the Vatche and Tamar Manoukian Division of Digestive Diseases, the Specialty Training and Advanced Research (STAR) Program, and a Research Grant for Training and Innovating Projects for Junior Faculty from the Eli and Edythe Broad Foundation (CP).

The UCLA CURE Digestive Diseases Research Center (NIH P30 DK 41301, PI: Rozengurt) provided funding for research personnel and for the UCLA Integrated Molecular Technologies Core. Additional core services were supported by NIH NCATS UL1TR000124.

I would like to extend my deepest gratitude to my thesis advisor Dr. Charalabos “Harry” Pothoulakis, for his constant encouragement, advocacy and support.

I would also like to extend a special acknowledgment to the members of my doctoral thesis committee: Lin Chang, who paved the road for my success to this point through over a decade’s worth of career, scientific and personal guidance, and who continues to guide me as I forge my

own path; Emeran Mayer, who has been a source of inspiration; and Joseph Pisegna, who has provided invaluable scientific and career advice. I would additionally like to thank Dimitrios Iliopoulos who served on my committee while at UCLA and provided mentorship.

Several individuals warrant special mention for the teaching and assistance they provided in experimental technique and data analysis. Swapna Joshi and Ivy Law taught me everything I know about molecular biology techniques, cell culture and data analysis. I could not have asked for more patient and generous teachers. I would also like to thank David Padua for his friendship and advice.

I would like to thank the following additional current and past members of Center for Inflammatory Bowel Disease: Iordannis Karagiannidis, Marina Koutsiompa, Robert Rankin, Kai Fang, Jill Hoffman, Michael Koon, Kyriaki Bakirtzi, Jonathan Ruiz and Sofia Quezada.

I would also like to thank the following additional institutions and individuals for their role in making this work possible:

Vatche and Tamar Division of Digestive Diseases: Eric Esrailian and Gary Gitnick for their generous support. Division and GI Fellowship Staff including Teresa Olivas, Trisha Navarro, Trisha James, Tonishia Boyle, Aida Santos and Farzana Hossain.

The UCLA CURE Digestive Diseases Research Center.

The Oppenheimer Center for Neurobiology of Stress and Resilience (CNSR): Study participants and staff. In particular Cathy Liu, Melanie Rosales, Melissa Alberto, Tiffany Ju and Nafeesa Islam.

The UCLA Center for Systems Biomedicine: Janet Treger and Emmanuelle Faure-Kumar.

The UCLA STAR Program: Directors Linda Demer and Mitch Wong and staff Laura Juhl and Nadia Barreda.

The UCLA Neurogenomics Core: Joseph DeYoung and staff.

I would like to thank the following additional people:

For their mentorship: Yvette Taché, Michael Camilleri, Jeff Conklin, David Reuben, Bruce Wainer.

For their friendship and support: Adriana LaCorte, Davi Bernstein.

Finally, I would like to thank my family, especially my sister Ali who keeps me smiling.

BIOGRAPHICAL SKETCH

NAME: Elizabeth Jane Videlock

EDUCATION/TRAINING

INSTITUTION AND LOCATION	DEGREE	YEAR(s)	FIELD OF STUDY
Yale University, New Haven, CT	BS	1998-2002	Chemistry
UCLA, Los Angeles, CA	MD	2004-2009	Medicine
Beth Israel Deaconess Medical Center, UCLA, Los Angeles, CA	Residency	2009-2012	Medicine
	Fellowship	2013-2016	Gastroenterology

Positions

- 2012 – 2013 Research Fellow, P.I. Anthony Lembo, M.D, Division of Gastroenterology, Department of Medicine, Beth Israel Deaconess Medical Center, Boston, MA
- 2012 – 2013 Hospitalist, Locum Tenens, Portland and Bar Harbor Maine
- 2014 – 2018 Gastroenterology Fellow, Research Track, Geffen School of Medicine-UCLA, Los Angeles, CA.
- 2015 – 2018 PhD Candidate, Molecular Cellular and Integrative Physiology, University of California, Los Angeles, CA.
- 2016 – 2017 Postdoctoral Scholar, UCLA Division of Digestive Diseases, Los Angeles, CA
- 2017 – 2019 Clinical Instructor of Medicine, UCLA Division of Digestive Diseases, Los Angeles, CA

Other Experience and Professional Memberships

- 2008 - Member, American Gastroenterological Association (AGA)
- 2009 - Member, American College of Physicians
- 2010 - Member, American Neurogastroenterology and Motility Society
- 2011 - Member, American College of Gastroenterology
- 2012 - Ad-hoc Reviewer, American Journal of Gastroenterology
- 2017 - Ad-hoc Reviewer, PLoS One, Clinical Gastroenterology and Hepatology, Scientific Reports, Neurogastroenterology and Motility, Digestive Diseases and Sciences, American Journal of Physiology – Gastrointestinal and Liver Physiology
- 2017 - Abstract Review Committee for Digestive Disease Week: Irritable Bowel Syndrome: Pathophysiology, American Gastroenterological Society
- 2018 - Ad-hoc Reviewer, Journal of Clinical Gastroenterology, Psychosomatic Medicine, Gastroenterology
- 2019 - Ad-hoc Reviewer, Psychoneuroendocrinology

Honors and Awards

- 2008 Award for best clinical poster, Center of Neurobiology of Stress, UCLA
- 2015 Abstract Travel Award, 2015 James W. Freston Single Topic Conference, A Renaissance in the Understanding and Management of IBS. AGA
- 2016 Young Investigators Forum, Federation of Neurogastroenterology and Motility
- 2016 Recipient of GI Fellowship Seed Grant, Barbara and Joel Marcus Foundation
- 2018 Basic Science Travel Award, American Gastroenterological Association
- 2018 Attendee, Little Brain Big Brain Neurogastroenterology Young Investigators Meeting
- 2018 Basic Science Poster Award, UCLA Research Conference on Aging
- 2018 Poster Award, Basic Science Category, UCLA Department of Medicine
- 2018: Scientific Accomplishment for Early Career Investigator, AGA, Washington, D.C.

Selected Publications

1. Vidlock EJ, Adeyemo M, Licudine A, Hirano M, Ohning G, Mayer M, Mayer EA, Chang L. Childhood trauma is associated with hypothalamic-pituitary-adrenal axis responsiveness in irritable bowel syndrome. *Gastroenterology* 137: 1954-1962, 2009.
2. Vidlock EJ, Cremonini F. Meta-analysis: probiotics in antibiotic-associated diarrhoea. *Aliment Pharmacology and Therapeutics*. 2012;35(12):1355-69.
3. Vidlock EJ, Lembo A, Cremonini F. Diagnostic testing for dyssynergic defecation in chronic constipation: meta-analysis. *Neurogastroenterology and Motility*. 2013;25(6):509-20.
4. Vidlock EJ, Cheng V, Cremonini F. Effects of linaclotide in patients with irritable bowel syndrome with constipation or chronic constipation: a meta-analysis. *Clinical Gastroenterology and Hepatology*. 2013;11(9):1084-92.e3; quiz e68.
5. Vidlock EJ, Shih W, Adeyemo M, Mahurkar-Joshi S, Presson AP, Polytarchou C, Alberto M, Iliopoulos D, Mayer EA, Chang L. The effect of sex and irritable bowel syndrome on HPA axis response and peripheral glucocorticoid receptor expression. *Psychoneuroendocrinology*. 2016;69:67-76.
6. Vidlock EJ, Mahurkar-Joshi S, Hoffman JM, Iliopoulos D, Pothoulakis C, Mayer EA, Chang L. Sigmoid colon mucosal gene expression supports alterations of neuronal signaling in irritable bowel syndrome with constipation. *American Journal of Physiology - Gastrointestinal and Liver Physiology*. 2018;315(1):G140-G157.

Selected Abstracts

1. Vidlock EJ, Mahurkar-Joshi S, Karagiannides S, Pothoulakis C, Iliopoulos D, Mayer EA, Chang L. Differential Colonic Mucosal mRNA Expression in IBS with Constipation. *Neurogastroenterology and Motility*. 2016; 28(1):10.
2. Vidlock EJ, Mahurkar-Joshi S, Iliopoulos D, Pothoulakis C, Mayer EA, and Chang L, Dysregulation of the Long-Noncoding RNA, GHRLOS, in Irritable Bowel Syndrome. *Gastroenterology*, 2017. 152(5): S722.
3. Vidlock EJ, Mahurkar-Joshi S, Iliopoulos D, Pothoulakis C, Mayer EA, and Chang L, Expression Profiling of Sigmoid Biopsies in Irritable Bowel Syndrome Vs Healthy Controls. *Gastroenterology*, 2017. 152(5): S722.
4. Vidlock EJ, Mahurkar-Joshi S, Hoffman JM, and Pothoulakis C, Gene Expression Network Analysis of the Gut-Brain Axis Supports an Association Between Alpha-Synuclein and Markers of Enteric Glial Cells. *Gastroenterology*, 2018. 154(6): S117.
**Received Basic Science Travel Award
5. Vidlock EJ, Mahurkar-Joshi S, Hoffman JM, and Pothoulakis C, Gene Expression Network Analysis of the Gut-Brain Axis Supports an Association Between Alpha-Synuclein and Markers of Enteric Glial Cells. Presented at 2018 UCLA Research Conference on Aging. ** Poster Award for Biological Sciences
6. Vidlock EJ, Mahurkar-Joshi S, Hoffman JM, and Pothoulakis C, Inter-organ correlations between gut and brain gene networks exist in neuro-glial and inflammation related pathways: An analysis of human post-mortem samples. *Neurogastroenterology and Motility*, 2018. 30(S1):119.
7. Vidlock EJ, Mahurkar-Joshi S, Iliopoulos D, Pothoulakis C, Mayer EA, and Chang L, The Long Non-Coding RNA AFDN-AS1 is Expressed in Colonic Epithelial Cells. *Am J Gastroenterol*, 2018. 113: S450.

Complete List of Published Work in My Bibliography:

<https://www.ncbi.nlm.nih.gov/myncbi/elizabeth.vidlock.1/bibliography/40588530/public/>

Chapter 1:

Introduction

ABSTRACT

Irritable bowel syndrome (IBS) is a common disorder characterized by abdominal pain associated with alterations in stool form and/or frequency. IBS is prevalent and accounts for significant cost and morbidity. While there have been significant advances in the field, there is still no reliable biomarker for IBS, nor is there a unified pathophysiologic mechanism. Therefore, both diagnosis and treatment are based on symptoms. The goal of the research presented in this dissertation was to improve understanding of IBS at the molecular level. Ultimately, my aim was to advance progress toward development of diagnostic tools and specific therapies for IBS and thereby improve quality of life for IBS patients.

IRRITABLE BOWEL SYNDROME

Definition of irritable bowel syndrome

Irritable bowel syndrome (IBS) is characterized by abdominal pain associated with alterations in stool form and/or frequency. It is frequently diagnosed in both primary care and specialty practice.(1) IBS occurs in the context of a grossly and histologically normal gastrointestinal (GI) tract. For this reason, it has been referred to as a “functional” GI disorder. Over the past decade, the traditional dichotomy between “functional” and “organic” has been called into question by the growing body of evidence for changes at the molecular level associated with IBS and other functional disorders. Furthermore, there is a biochemical basis for all symptoms at the peripheral (gut-based) or central (brain, spinal cord) level. The current preferred classification of IBS is a *Disorder of Gut-Brain Interaction*.(2) It has also been suggested that IBS could be viewed as multiple "organic" diseases.(3) Despite the increasing identification of molecular differences in IBS, there is not a unifying putative mechanism and patients are grouped clinically by symptoms, so that IBS is appropriately designated as a syndrome rather than a disease.

Epidemiology

Prevalence and cost

The worldwide prevalence of IBS is around 11%.(4) Patients with IBS have been consistently shown to have poor health-related quality of life (HRQOL), comparable to other chronic disorders such as gastroesophageal reflux disease (GERD), diabetes, hypertension and depression.(5-7) In comparison to patients with end-stage renal disease dependent on dialysis, patients with IBS score lower on domains assessing bodily pain, energy/fatigue, and social functioning.(6) Detriments in HRQOL are severe enough that, in tertiary care settings, there is an increased risk of suicidal ideation in IBS that is independent of psychological diagnoses.(8,9). IBS accounts for significant health care costs,(10) as patients with IBS are more likely to have blood tests, radiologic tests, endoscopic procedures, and abdominal or

pelvic surgeries than those without IBS.(11) In addition to direct medical expenses, costs to society come in the form of work absenteeism and impaired work productivity.(10)

IBS is more prevalent in women. In a meta-analysis pooling data from population-based studies (total N=188,229), the odds of having IBS were higher among women with an odds ratio (OR) of 1.7 and a 95% confidence interval of 1.5 to 1.8.(12) The increased prevalence was stable across geographic regions. Women with IBS were more likely to have IBS with constipation (IBS-C) with an odds ratio of 2.38 (95% confidence interval: 1.45-3.92), and less likely to have IBS with diarrhea (IBS-D) with an odds ratio of 0.45 (95% confidence interval: 0.32-0.65).(12) IBS is also more commonly diagnosed in patients under the age of 50 (13-15), and is more prevalent among patients with lower income.(13,15)

Risk factors

Family history

Although IBS aggregates in families,(16-19) and twin studies suggest heritability,(20,21) a strong influence of environment and parental response to a child's abdominal pain may also be influential.(22-26)

Prior GI infection

Post-infection IBS (PI-IBS) is the onset of IBS symptoms following resolution of acute infectious gastroenteritis, characterized by two or more of the following: fever, vomiting, diarrhea, or a positive bacterial stool culture, in an individual without a history of IBS.(27) GI infection is the strongest risk factor for IBS and is associated with an approximately 4-fold increase in risk of IBS-symptoms at twelve months in comparison to uninfected individuals.(28)

Stressful life events

There is an association between IBS symptoms and current life stress for many patients with IBS.(29) IBS patients also have a higher prevalence of early adverse life events (EALs), or traumatic experiences during childhood including, but not limited to, maladjusted relationships

with a parent or primary caregiver, severe illness or death of a parent, and physical, sexual or emotional abuse.(30,31) The development of PI-IBS is also associated with stressful life events.(32)

Symptom-based diagnosis

In the absence of a validated biomarker, diagnosis is based on symptoms. The rationale for the use of symptom-based criteria are that patients present to health care providers with symptoms (33,34) and symptom clusters are consistent across clinical samples and in population studies.(35,36) The key symptom of IBS is chronic or recurrent abdominal pain associated with altered bowel habits. The current criteria for diagnosis of IBS are the Rome IV criteria.(37) The patients included in the studies in this dissertation were diagnosed according to the previous, Rome III, criteria, which were published in 2006.(1) Both criteria are listed in Table 1. The main differences in the current (Rome IV) criteria are increased symptom frequency, and removal of the term “discomfort.” In addition, language was revised to be less restrictive regarding the timing of pain and bowel symptoms and whether pain was improved with defecation. These changes are emphasized by bold text in Table 1. Rome IV were slightly less sensitive but more specific (Table 1) in classifying IBS patients with both negative evaluation for alternative diagnoses and diagnosis of IBS by an experienced clinician.(38) Supportive symptoms that are not used for diagnosis include: abnormal stool frequency (> 3 bowel movements/day or < 3 bowel movements/week), abnormal stool form (types 1–2 or 6–7 of the Bristol scale), excessive straining during defecation, defecatory urgency, feelings of incomplete evacuation, and mucus with bowel movements. In addition, an unpredictable bowel pattern (≥ 3 different stool form types/week) reinforces the diagnosis of IBS-D (39), while consecutive days without a bowel movement is associated with IBS-C.(40) Although the majority of patients have bloating, it is not required for diagnosis. IBS is sub-grouped by predominant bowel habit. Because stool form was the best predictor of predominant bowel habit in IBS, stool form rather than frequency

determines bowel habit sub-classification according to both the Rome III and Rome IV criteria.(1,37) These criteria are outlined in Table 2 and Figure 1.

Table 1: Diagnostic criteria for IBS

Rome IV (37)	Rome III (1)
Recurrent abdominal pain, on average, at least 1 day per week in the last 3 months , associated with 2 or more of the following criteria: 1. Related to defecation 2. Associated with a change in frequency of stool 3. Associated with a change in form (appearance) of stool Criteria fulfilled for the last 3 months with symptom onset at least 6 months before diagnosis.	Recurrent abdominal pain or discomfort at least 3 days per month in the last 3 months that is associated with 2 or more of the following: 1. Improvement with defecation 2. Onset associated with a change in frequency of stool 3. Onset associated with a change in form (appearance) of stool Criteria fulfilled for the last 3 months with symptom onset at least 6 months prior to diagnosis.
Sensitivity/Specificity (38)	
62.7%/97.1%	73.1%/93.1%

Differences are highlighted in bold text.

Table 2: IBS subtyping

Subtype*	Percent of stools that meet the description (over the preceding 3 months)	
	Hard or lumpy stools (Bristol type 1-2)	Loose or watery stools (Bristol type 6-7)
IBS with constipation (IBS-C)	≥ 25%	< 25%
IBS with diarrhea (IBS-D)	< 25%	≥ 25%
Mixed IBS (IBS-M)	≥ 25%	≥ 25%
Unsubtyped IBS (IBS-U)**	< 25%	< 25%

*For Rome IV, percent of stools refers to days with at least one abnormal bowel movement.(37)

**Patients who meet diagnostic criteria for IBS but whose bowel habits cannot be accurately categorized into 1 of the 3 groups above should be categorized as having IBS unclassified.

Reprinted from Videlock EJ, Chang L. "Irritable Bowel Syndrome". In: Talley NJ, Kane SV, Wallace MB, editors. Practical Gastroenterology and Hepatology: Small and Large Intestine and Pancreas. 1 ed. Oxford, United Kingdom: Wiley-Blackwell; Copyright 2010 with permission from John Wiley and Sons.

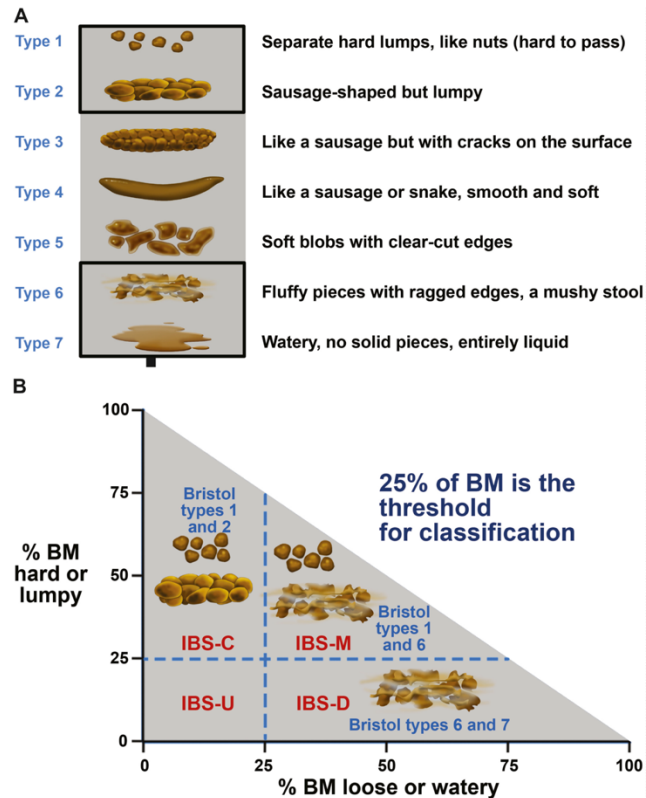


Figure 1: Bowel habit subtypes

This figure demonstrates the types of stool based on consistency and can be used as a guide for subtyping by predominant bowel habit.

BM; bowel movement

Reprinted from *Gastroenterology*:150(6), Lacy BE et al., *Bowel Disorders*, page 1396; Copyright 2016, with permission from Elsevier and the Rome Foundation.

Pathogenesis

Gut sensation and function are influenced by the state and activity of multiple domains, including the gut lumen, mucosa, enteric nervous system (ENS) and central nervous system (CNS), as well as communication between these domains. A unifying theory in the pathogenesis of IBS is that there is a dysregulation in the complex interplay among these domains, which collectively constitute the “gut-brain axis.”(41) There is not a consensus on the underlying etiology of this dysregulation, and the disorder may represent a combination of factors involving different mechanisms.

This section starts with an overview of visceral sensation, and a discussion of enhanced visceral perception in IBS. This is followed by a more detailed discussion of peripheral putative pathogenic mechanisms in IBS, including how peripheral mechanisms may affect the gut-brain axis. There is significant overlap and interrelatedness of the discussed concepts, which is summarized in Figure 2. The section concludes with a summary of evidence for a potential role of genetic factors. The contribution of central nervous system (CNS) mechanisms to enhanced visceral perception is only briefly discussed, as the focus of this dissertation is on peripheral mechanisms.

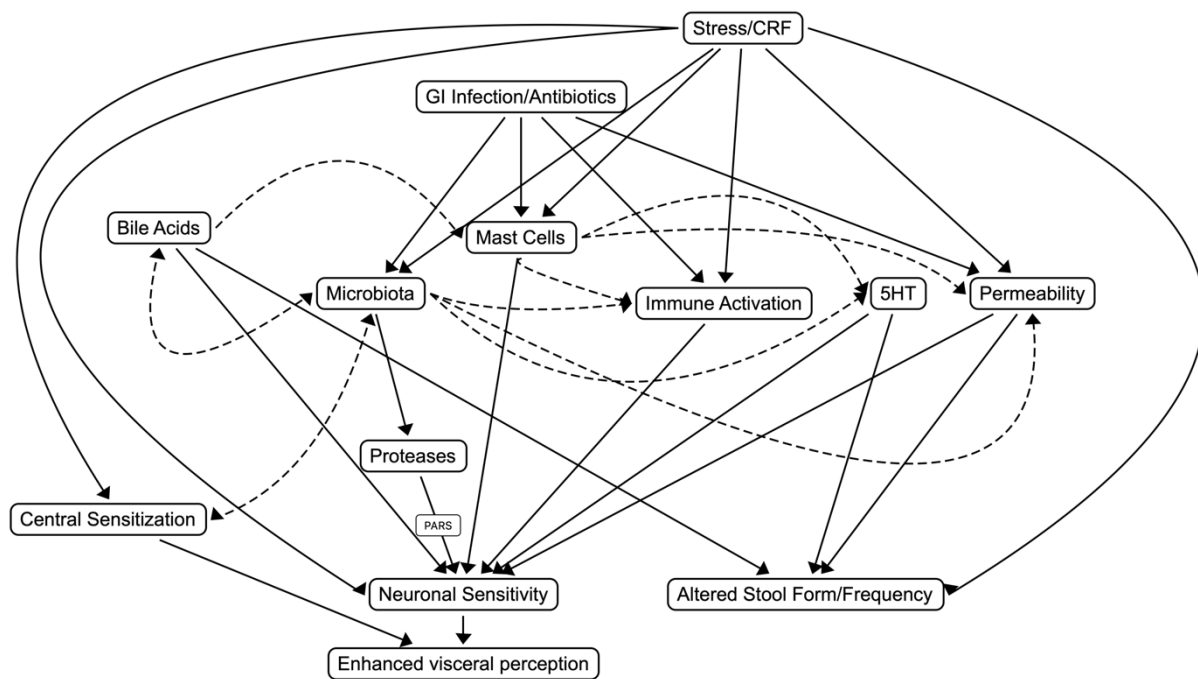


Figure 2: Interrelatedness of IBS pathophysiology concepts

Introduction to visceral sensation

Mechanical, chemical and other stimuli are detected in the gut by primary afferent neurons, which are extrinsic, intrinsic or intestinofugal. The primary nociceptors in the colon and rectum are the spinal primary afferent neurons.(42) Spinal afferents have variable thresholds for activation, which can change in response to injury and inflammation, and there are also silent

nociceptors, responding only in the presence of injury or inflammation.(43) The process of increased sensitivity of primary afferent neurons is referred to as peripheral sensitization, and it can occur via mediators released by various cell types including platelets, lymphocytes, mast cells and epithelial cells, and may also occur via alterations in second messenger systems or changes in gene expression.(44,45)

Spinal primary afferent neurons have cell bodies in the dorsal root ganglion (DRG).(46) Their central processes synapse in the dorsal horn, and input is conveyed along the ipsilateral dorsal column to the contralateral ventroposterolateral nucleus of the thalamus. The CNS can modulate the response to visceral input and can change the responsiveness of reflexes controlling gut function in response to input from visceral afferents.(41) The degree to which input results in conscious sensation can also be moderated. For example, injury or inflammation, in addition to increasing sensitivity of primary afferent neurons, can change the sensitivity of central circuits,(47) leading to decreased thresholds for sensation and increased responsiveness.(48)

Enhanced visceral perception

Enhanced perception of visceral stimuli (Figure 3), a reproducible finding in a significant subset of IBS patients,(49) can occur either as a result of greater sensitivity of visceral afferent pathways or central amplification of visceral afferent input.(50) There is strong evidence for a role of the CNS in enhanced visceral perception through central sensitization. The important brain regions can be grouped into those associated with visceral sensation (in healthy controls as well), those involved in endogenous pain modulation, and those involved in emotional circuits. A quantitative meta-analysis of brain regions activated in response to rectal distention in IBS patients and controls, revealed that IBS patients have increased activation of areas associated with emotional arousal (pregenual anterior cingulate cortex, amygdala) as well as a midbrain cluster associated with endogenous pain modulation, while controls showed more

consistent activation of the medial and lateral prefrontal cortex.(51) Areas associated with endogenous pain modulation, such as the prefrontal cortex, have also consistently shown altered activation in response to rectal distention in IBS.(52,53) There is also evidence for peripheral factors in enhanced visceral perception. Sensitivity can be reduced by intrarectal lidocaine (54) and can be increased by rectal glycerol.(55)

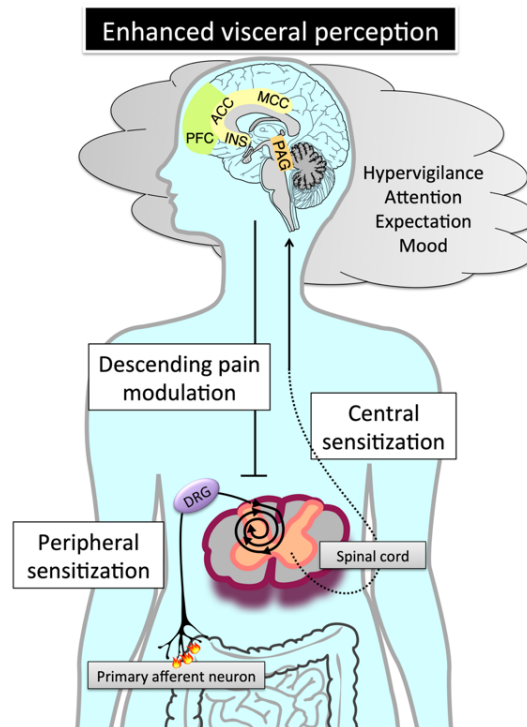


Figure 3: Enhanced visceral perception

Reprinted from Videlock EJ, Chang L. "Irritable Bowel Syndrome". In: Camilleri M, Podolsky D, editors. Yamada's Textbook of Gastroenterology. 6 ed. Oxford, United Kingdom: Wiley-Blackwell; Copyright 2014, with permission from John Wiley & Sons, Ltd.

Peripheral changes in IBS

Changes in neuron density and receptor expression

Some studies have shown increased nerve fibers in IBS patients compared with controls, in particular those expressing receptors for substance P and transient receptor potential vanilloid type 1 (TRPV-1),(56) but this has not been replicated in all studies.(57) Other studies have implicated cannabinoid receptors (58) and protease-activated receptors (PARs).(59)

Luminal and tissue mediators

There is strong evidence for the presence of sensitizing mediators in the tissue or intestinal lumen of IBS patients.(60-64) Multiple studies have demonstrated that biopsy supernatants from IBS patients increase firing of afferent neurons in animal models (64,65) and submucosal enteric neurons in human biopsies.(66,67) In one study, a patient's symptom severity was correlated with the degree to which their biopsy supernatant increased sensitivity in an animal model.(68) Protease activity is increased in IBS supernatants and may mediate the effect of these supernatants on neuronal sensitivity.(64) The increased visceral sensitivity following intracolonic administration of IBS biopsy supernatants in mice was decreased by protease inhibitors and was not present in mice lacking protease activated receptor 2 (PAR2).(64) Sensitivity of TRPV1 neurons was also increased by histamine, an effect mediated by the histamine 1 receptor.(69)

Increased gut permeability

Permeability has been linked to abdominal pain (70) and enhanced visceral perception (71) in patients with IBS. Gut permeability is frequently measured in experimental settings by ingestion of molecular probes which are excreted and can be measured in the urine.(72) Several studies have found increased small intestinal (71,73,74) and colonic permeability in IBS patients,(75) and increased paracellular permeability in colonic biopsies.(70) This has been corroborated by studies showing altered expression and cellular distribution of tight junction proteins in the colon (76) and jejunum of patients with IBS-D.(77) The association between increased permeability and IBS is strongest in PI-IBS (78) and IBS-D.(71,73,79)

Altered immune reactivity

Increased immune activation is hypothesized to play a role in the pathophysiology of IBS through effects on visceral sensitivity and gut epithelial function. The strongest evidence comes from studies in patients with PI-IBS,(27) which have shown increased numbers of T

lymphocytes in the lamina propria compared to controls,(32) and increased expression of the pro-inflammatory cytokine interleukin 1 beta (IL-1 β). (80)

Importantly, in one study, while sensitization of TRPV1 neurons persisted two years following infection, there were no longer differences in lymphocytes or tissue cytokine mRNA expression.(81) A potentially transient nature of immune alterations may help to explain why studies in unselected IBS patients have yielded less consistent results. The results of 22 studies measuring immune cells in intestinal biopsies (706 IBS patients, 401 controls) were summarized and pooled in a systematic review and meta-analysis published in 2018.(82) The results of this meta-analysis were increased mast cells in the rectosigmoid and descending colon (both in IBS-C and IBS-D) and increased CD3+ T cells in the rectosigmoid and the descending colon, which was likely due to increased CD4+ cells as CD8+ cell numbers were similar.(82)

Mast cells have been the most thoroughly studied immune cell population in IBS.(56,83,84) The putative role of mast cells in IBS is increased peripheral sensitization through release of histamine. In one study, the proximity of mast cells to sensory neurons was associated with increased severity and frequency of abdominal pain and discomfort.(83) In another study, mast cells were increased in IBS and were associated with increased colonic permeability and symptoms of diarrhea, but were not related to disease severity.(85) However, other studies have found similar numbers of mucosal immune cells in IBS and controls,(64,86-88) and one study found decreased mast cells, T cells and macrophages in IBS.(89)

Studies have also shown alterations in cytokines in IBS patients, but there has been more evidence for blood levels than colonic mucosal levels.(90) The most frequent findings in studies of serum cytokines are increased tumor necrosis factor alpha (TNF- α) (91-94) and increased interleukin 6 (IL-6),(91,93,94) although these findings have not been replicated in all studies.(87,95,96) Secretion of TNF- α , and IL-6 was also increased following stimulation of peripheral blood mononuclear cells (PBMCs) from IBS patients in comparison to controls.(95,97) Tissue cytokine results are more variable, though several studies have shown

decreased interleukin 10 (IL-10).(87,98,99) In most studies, there was no association of cytokines or cell counts with symptoms.(87,91,94,95)

Bile acid processing

Bile acids are absorbed in the terminal ileum via active transport. Bile acids in the colon and rectum increase secretion, motility,(100) and visceral sensitivity in response to rectal distention.(101) Animal studies suggest that the visceral sensitivity induced by bile acids involves activation of the farnesoid X receptor (FXR) on mast cells leading to increased expression of nerve growth factor and subsequent increased TRPV1 expression on dorsal root ganglion neurons.(102) Bile acid diarrhea is present in around 25% of IBS-D patients.(103-105) Colonic bile acids correlate with stool weight,(106) even in patients without overt bile acid malabsorption.(107) Bile acids may also be relevant for a subset of patients with IBS-C who have reduced total bile acids and reduced fecal deoxycholic acid.(106,108)

Serotonin

The majority of serotonin (5-HT) is found within the gut. It is primarily stored in enterochromaffin (EC) cells, but is also contained in serotonergic neurons of the myenteric plexus of the ENS.(109) EC cells release 5-HT in response to a variety of stimuli such as mechanical distention. Serotonin is a neurotransmitter that binds to receptors on primary afferent neurons in the lamina propria in a paracrine fashion. These neurons may include extrinsic nerves and mucosal projections of intrinsic primary afferent neurons, which initiate gut reflexes affecting motility and secretion. Serotonin does not diffuse back into cells and it is not degraded enzymatically. The effects are mitigated by reuptake into cells via transporter proteins, such as the serotonin reuptake transporter, SERT.

Evidence supports a role for serotonin processing and signaling in motility and transit alterations associated with IBS. Studies have shown increased postprandial concentrations of 5-HT in the platelet-depleted plasma of IBS patients compared to controls, particularly in IBS-D.(110,111) In

another study, colonic transit times correlated with the level of 5-HT in platelet-depleted plasma in IBS patients, with elevated levels predicting accelerated transit.(111) Studies also indicate changes in serotonin signaling in IBS. For example, while IBS-C patients had similar numbers of EC cells to controls, EC cells had increased serotonin content, which might suggest impaired release.(110)

Gut microbiota

There is evidence that the microbiota may play a role in IBS.(112) In a meta-analysis of 24 studies, the microbiome of IBS patients in comparison to controls had increased abundance of family Enterobacteriaceae (phylum Proteobacteria), family Lactobacillaceae, and genus Bacteroides, and decreased abundance of uncultured Clostridiales I, genus Faecalibacterium (including Faecalibacterium prausnitzii), and genus Bifidobacterium.(113) A putative role for alterations in the microbiome on IBS symptoms is supported by studies of probiotics,(114) antibiotics,(115) and diets known to alter the microbiome (116) in the treatment of IBS, as well as the observation that IBS has been associated with antibiotic use.(117) Based on either animal or human models, probiotics have been shown to affect visceral sensitivity,(118) motility,(119) intestinal permeability,(120) and immune activity, although the mechanisms governing these responses are not known.(121,122) The microbiota has been linked to motility and transit in animal models. For example, germ-free rats have alterations in gastric emptying, intestinal transit and motility that resolve with introduction of normal gut flora.(123) It is hypothesized that these effects relate to neuroactive microbial metabolites, but direct evidence supporting this hypothesis is lacking.(124,125) Gut microbes have even been shown to regulate serotonin synthesis from enteroendocrine cells,(126) supporting a bacterial-neuropeptide dependent pathway in regulation of GI motility.

Alterations in the gut-brain axis contributing to sensitization

Changes in the stress response system

There is both preclinical and clinical evidence to support the link between chronic stress and IBS. Chronic stressful life events affect clinical outcomes in IBS,(127) and PI-IBS is predicted by psychological symptoms and stress.(32,128) Furthermore, experimental stress increases visceral sensitivity (129,130) and gut permeability.(131) In addition, patients with IBS have an exaggerated increase in motility in response to stress.(132) The effect of stress on gut function and sensation likely results from alterations in the brain-gut axis, or the bidirectional communication between the enteric, autonomic and central nervous systems.(41) The nature of this communication is likely complex and multidimensional. In addition to communication via neurons, communication occurs via humoral mediators of the HPA axis, ANS and immune system. The microbiota may also participate in brain-gut interactions.

The HPA axis is the primary neuroendocrine system in the stress response. Activation of the HPA axis results in synthesis and release of corticotropin-releasing factor (CRF) in the paraventricular nucleus of the hypothalamus. CRF stimulates release of adrenocorticotropin hormone (ACTH) from the anterior pituitary, which leads to the release of cortisol from the adrenal glands. Cortisol in turn attenuates the response via negative feedback at the level of the hypothalamus and the pituitary. CRF is released into the circulation and has direct effects on many tissues. A large body of research in animal models has established an important role for central and peripheral effects of CRF on gut transit, motility and secretion, the local gut immune response, and changes in visceral sensitivity that occur in the context of tissue injury and inflammation.(133)

Early adverse life events (EALs)

Many chronic disorders are associated with EALs including chronic obstructive pulmonary disease, autoimmune disorders, obesity and mood disorders.(134,135) There is also a large

body of evidence linking early life stress to the development of anxiety behaviors and visceral hypersensitivity in animal models such that this paradigm is often used as an animal model of IBS.(136) Most studies use maternal separation (removal of pups from their mothers during the early post-natal period) as an animal model of EALs.(137-139) As adults, these rats show evidence of visceral hypersensitivity following water avoidance stress. Maternally separated animals also show alterations in the HPA axis and the microbiota,(140) as well as increased intestinal permeability.(141) The long-term effects of EALs may result from epigenetic programming. Epigenetics refers to mechanisms that affect gene expression (e.g. DNA methylation, histone modification) and may change phenotype without a change to the genetic sequence (genotype). These changes can be passed down meiotically and can permanently alter the expression of genes in somatic cells, an effect referred to as epigenetic programming.(142) Meaney and colleagues showed that methylation of the glucocorticoid receptor (GR) promoter in rodents was associated with perinatal stress, a phenomenon that they subsequently linked to a prolonged HPA axis response via decreased negative feedback in the adult animals.(143) Likewise, in humans, fetal exposure to cortisol resulted in reduced expression of the glucocorticoid receptor (GR) in the hippocampus, and a prolonged HPA axis response.(144) In addition, a study of human autopsy specimens revealed decreased hippocampal GR expression and increased methylation of the neuron-specific GR gene promoter in suicide victims with a history of childhood abuse when compared with suicide victims without abuse as well as controls without abuse who died from unrelated causes.(145) These findings have since been replicated in another study.(146) A role for epigenetics in animal models of IBS is supported by the finding that a histone deacetylase inhibitor prevented anxiety and visceral sensitivity in maternally separated animals.(147)

Gut-brain axis effects of peripheral changes

Disciplines such as neuroimmunology and neuroendocrinology have emerged as a result of the growing body of knowledge describing the complex network in which the nervous system, immune system and endocrine system, particularly the HPA axis, interact in both health and disease. Many stress-induced changes in GI physiology are mediated by CRF and include: increased colonic motility in humans(148), increased motility and visceromotor response in animal models,(149) increased permeability and mast cell degranulation in human biopsies,(150) and stress-induced increased small intestinal permeability in humans.(131) The following paragraphs describe how peripheral factors discussed above relate to the gut-brain axis.

Microbiota: Microbes contribute to the bidirectional communication that forms the brain-gut axis.(151) Both prenatal stress and postnatal stress in the form of maternal separation have been shown to affect gut microbiota in animal models.(140,152,153) Stress may affect the gut microbiota via changes in immune activation, permeability and autonomic nervous system-mediated changes in mucus secretion.(154) Changes in the microbiota, in turn, have influence on the central nervous system. For example, prenatal stress resulted in HPA hyperresponsiveness in rats, a phenotype which was associated with the microbiota.(155) In another study, changes in the microbiota were associated with altered levels of expression of brain-derived neurotrophic factor (BDNF).(156) There is preliminary evidence that mood affects gut microbiota and *vice versa* in humans.(157)

Immune activation: PI-IBS provides a good example of the interactions between stress and the immune system. PI-IBS is associated with increased gut lymphocytes and EC cells (78) and increased expression of IL-1 β .(80) PI-IBS is also associated with increased gut permeability.(74,78) The most important predictors of PI-IBS are chronic life stress and anxiety and depression.(32,128) In a likely analogous manner, psychological symptoms are related to GI symptoms and disease activity in IBD.(158) In both PI-IBS and IBD, there is evidence that

the influence of stress on disease is mediated by the HPA axis via action of CRF in the gut.(159)

Ibeakanma and colleagues studied the effect of stress on visceral sensitivity in mice following infection with *C. rodentium*, which produces a self-limited colitis.(160) Twenty-one days following the infection, a group of mice was subjected to 9 days of water avoidance stress. These mice had serum corticosterone and epinephrine levels that were 2- and 3- fold higher than mice exposed to infection alone without chronic stress. The mice exposed to stress also showed visceral hypersensitivity and increased visceromotor response to colorectal distention; patch-clamp studies revealed greater excitability of colonic primary afferent neurons and greater activity in response to balloon distention. While the levels of proteases were increased in post-colitis mice regardless of exposure to stress, protease inhibition did not affect the changes in neuron excitability in the stressed mice. This suggests that protease signaling may play a role in increased sensitivity following infection, but not the further augmented response seen with chronic stress following infection. Additional studies showed that epinephrine and cortisol increased the excitability of DRG neurons from the gut.(160) Potentially more interesting was the fact that while a protease activated receptor (PAR)-activating peptide did not affect the excitability of neurons in control mice, it markedly increased excitability in neurons from mice exposed to chronic stress. These findings suggest that prior infection and chronic stress lead to alterations in the brain-gut axis characterized by increased sensitivity of peripheral neurons to mediators associated with both, stress and inflammation.(160)

Permeability: Increased gut permeability may play a role in the interaction between stress, immune activation, dysbiosis and hypersensitivity in IBS. Permeability is increased by experimental stress in humans and animal models.(131,161) Stress-induced permeability is associated with changes in mucosal cellularity including increased mast cells counts and activation.(162) This process is also mediated by CRF,(163) particularly in maternally separated rat models of visceral hypersensitivity.(164)

Genetics

As discussed earlier, genetic factors in IBS is supported by familial clustering and twin studies.(19,165) A small percentage (2-3%) of IBS patients have rare functional variants of the voltage-gated channel $\text{Na}_v1.5$ (*SCN5A*) causing a channelopathy or the sucrose isomaltase (*SI*) gene causing carbohydrate malabsorption.(166,167)

In 2014, Ek et al. published the first large genome-wide association study (GWAS) in IBS.(168) The cohort was a Swedish twin study. IBS cases ($n = 534$) were selected based on responses to a bowel symptom questionnaire that had a high concordance with Rome II criteria and exclusion of those with reported organic GI disease. Controls ($n = 4,932$) answered “no” to ‘ever had recurrent abdominal problems.’ There were no single nucleotide polymorphisms (SNPs) meeting the p value threshold for genome-wide association ($p < 5 \times 10^{-8}$). Of 42 genes with $p < 10^{-4}$, fourteen were selected for validation in several population-based case-control cohorts (total $n = 3,511$). One locus (7p22.1) was significant in the replication. It includes the genes KDEL endoplasmic reticulum protein retention receptor 2 (*KDELR2*) and glutamate receptor, ionotropic, delta 2 (Grid2) interacting protein (*GRID2IP*). (168)

The same group of investigators performed a larger GWAS published in 2018.(169) The study used a population-based cohort in the United Kingdom and included 9,576 cases with self-reported IBS and 336,499 controls. The replication cohort included 2,045 IBS patients from multiple centers and 7,955 ancestry-matched healthy controls. There was high quality data for over 7 million SNPs in comparison to about 1 million in the 2014 study. There was one locus, rs10512344 on chromosome 9q31.2, which reached genome-wide significance. There were an additional fourteen loci with “suggestive” significance ($p < 5 \times 10^{-6}$). A gene set enrichment analysis showed that IBS risk genes were enriched for intracellular calcium activated chloride channel activity, ion gated channel activity, and anion channel activity, as well as for targets of the miR-15 family of microRNAs. The investigators noted the presence of another SNP at the chromosome 9 locus associated with age at menarche. A sex-specific analysis revealed the that

the IBS-risk SNP conferred no risk of IBS in men and the overall association was due to a strong association in women. Additional analyses conducted by the authors, suggest that association of rs10512344 with IBS is independent of allele associated with age at menarche.(169)

Also in 2018, the same group of authors used meta-analysis techniques to determine IBS-associated loci by pooling GWAS results from five independent European genotyped cohorts with Rome II compatible questionnaire data (total n: 1,335 IBS and 9,768 controls).(170) While no locus met genome-wide significance, seven regions (64 genes) met “suggestive” significance ($p \leq 5.0 \times 10^{-6}$). Functional annotation highlighted regulation of ion channel activity.(170)

The association of common variants with IBS has also been studied in a large number of case-control studies, which are summarized in the review by Gazouli et al.(171) One of the best-studied polymorphisms that has been linked to IBS is in a region of the serotonin transporter gene known as 5-HTTLPR. Overall, results have been inconclusive, with high quality studies and a meta-analysis that do not support the association of this gene with IBS,(172,173) but several subsequent studies which did find an association.(174,175) In one study, the polymorphism was shown to result in altered expression of SERT mRNA and protein, with higher expression in the colonic mucosa for the L/L genotype, which is more common in the IBS-C subtype.(174) Polymorphisms that have been associated with IBS have also been found in the genes for CRF receptor 1 (CRF-1R),(176,177) cannabinoid receptors,(178) catechol-O-methyltransferase (COMT), interleukins and TNF- α .(165,179-182)

In addition to genetic risk factors, researchers have identified alterations in a variety of epigenetic factors in IBS. These include alterations in gene methylation (183) and expression of non-coding microRNAs.(77,79,184,185) While overall, the results of these studies support the relevance of genetic and epigenetic factors, additional studies are needed for validation, and to determine the functional relevance of findings.

SUMMARY

In summary, IBS is a common disorder with an incompletely understood pathophysiology. There is increasing evidence for molecular changes in colon tissue of IBS patients. Stress and its effect on gut function and sensation plays a key role in symptom generation. While central factors such as stress and hypervigilance are undoubtedly linked to IBS, central processes manifest in peripheral changes, such as changes in permeability as a result of stress in human and animal models. Furthermore, changes that may originate in the periphery (diet, microbiota, infection) may affect the CNS, which could account for their role in IBS pathophysiology. While our understanding of the pathogenesis of IBS has evolved over recent years, it is still far from complete. Understanding peripheral mechanisms in IBS may lead to the development of gut-targeted therapies to treat symptoms in IBS.

INTRODUCTION TO DISSERTATION RESEARCH

The goal of my doctoral research was to investigate molecular changes in colon biopsies of IBS patients in comparison to healthy controls using a translational approach. The primary objective was to identify genes differentially expressed in IBS and study one of these genes in cellular models to understand how its dysregulation could be related to IBS symptoms or pathophysiology. Our group has previously found differences in IBS related to epigenetic regulation of gene expression (methylation, microRNA). In addition, researchers in the UCLA Center for Inflammatory Bowel Disease have implicated long non-coding RNAs (lncRNAs), another set of non-coding RNAs important in epigenetic regulation of gene expression, in IBD. The field of lncRNA biology is expanding rapidly and many lncRNAs play key roles in disease pathogenesis. Additionally, the lncRNA field is new, and based on important roles for these epigenetic regulators, it offers many opportunities to identify novel mechanisms. For these reasons, I set out to identify and study lncRNAs that may be involved in IBS pathogenesis.

The first phase of this research involved gene expression profiling of colonic mucosal biopsies from IBS patients and healthy controls using a microarray platform that identified protein-coding mRNAs as well as lncRNAs. Chapter 2 details the analysis of the mRNAs on the microarray and is a version of a manuscript published in *The American Journal of Physiology-Gastrointestinal and Liver Physiology*.⁽¹⁸⁶⁾ Chapter 3 describes analysis of the lncRNAs on the microarray and the process of selecting the lncRNA afadin divergent transcript (*AFDN-DT*) for further study in intestinal epithelial cell culture models. *AFDN-DT* was up-regulated in IBS and I hypothesized that dysregulation of *AFDN-DT* may affect epithelial permeability. In the second phase of this research, described in Chapter 4, I developed cell lines capable of inducible overexpression of *AFDN-DT*, and I tested the effect of *AFDN-DT* overexpression on transepithelial electrical resistance and transcellular permeability. I also used RNA sequencing to determine the downstream effectors of *AFDN-DT*.

An additional aim of my doctoral training was to develop proficiency in the analysis of high-throughput gene expression profiling data. I performed all analyses on the microarray data in Chapters 2 and 3 and the RNA sequencing data in Chapter 4. I also applied these techniques to publicly-available RNA sequencing data from brain and colon tissue of healthy deceased donors. I hypothesized that pathways important in the gut-brain axis could be identified through a systems biology approach to this data. These analyses are described in Chapter 5, a version of which will be submitted for publication. The code used for all analyses, which were performed using the R software,⁽¹⁸⁷⁾ is included in the Appendix.

REFERENCES:

1. Longstreth GF, Thompson WG, Chey WD, Houghton LA, Mearin F, Spiller RC. Functional bowel disorders. *Gastroenterology*. 2006;130(5):1480-1491.
2. Drossman DA, Hasler WL. Rome IV-Functional GI Disorders: Disorders of Gut-Brain Interaction. *Gastroenterology*. 2016;150(6):1257-1261.
3. Camilleri M. Peripheral mechanisms in irritable bowel syndrome. *N Engl J Med*. 2012;367(17):1626-1635.
4. Lovell RM, Ford AC. Global prevalence of and risk factors for irritable bowel syndrome: a meta-analysis. *Clin Gastroenterol Hepatol*. 2012;10(7):712-721.
5. El-Serag HB, Olden K, Bjorkman D. Health-related quality of life among persons with irritable bowel syndrome: a systematic review. *Aliment Pharmacol Ther*. 2002;16(6):1171-1185.
6. Gralnek IM, Hays RD, Kilbourne A, Naliboff B, Mayer EA. The impact of irritable bowel syndrome on health-related quality of life. *Gastroenterology*. 2000;119(3):654-660.
7. Saito YA, Schoenfeld P, Locke GR, 3rd. The epidemiology of irritable bowel syndrome in North America: a systematic review. *Am J Gastroenterol*. 2002;97(8):1910-1915.
8. Miller V, Hopkins L, Whorwell PJ. Suicidal ideation in patients with irritable bowel syndrome. *Clin Gastroenterol Hepatol*. 2004;2(12):1064-1068.
9. Spiegel B, Schoenfeld P, Naliboff B. Systematic review: the prevalence of suicidal behaviour in patients with chronic abdominal pain and irritable bowel syndrome. *Aliment Pharmacol Ther*. 2007;26(2):183-193.
10. Canavan C, West J, Card T. Review article: the economic impact of the irritable bowel syndrome. *Aliment Pharmacol Ther*. 2014;40(9):1023-1034.
11. Ladabaum U, Boyd E, Zhao WK, Mannalithara A, Sharabidze A, Singh G, Chung E, Levin TR. Diagnosis, comorbidities, and management of irritable bowel syndrome in patients in a large health maintenance organization. *Clin Gastroenterol Hepatol*. 2012;10(1):37-45.
12. Lovell RM, Ford AC. Effect of gender on prevalence of irritable bowel syndrome in the community: systematic review and meta-analysis. *Am J Gastroenterol*. 2012;107(7):991-1000.
13. Andrews EB, Eaton SC, Hollis KA, Hopkins JS, Ameen V, Hamm LR, Cook SF, Tennis P, Mangel AW. Prevalence and demographics of irritable bowel syndrome: results from a large web-based survey. *Aliment Pharmacol Ther*. 2005;22(10):935-942.
14. Hungin AP, Chang L, Locke GR, Dennis EH, Barghout V. Irritable bowel syndrome in the United States: prevalence, symptom patterns and impact. *Aliment Pharmacol Ther*. 2005;21(11):1365-1375.

15. Minocha A, Johnson WD, Abell TL, Wigington WC. Prevalence, sociodemography, and quality of life of older versus younger patients with irritable bowel syndrome: a population-based study. *Dig Dis Sci*. 2006;51(3):446-453.
16. Saito YA, Petersen GM, Larson JJ, Atkinson EJ, Fridley BL, de Andrade M, Locke GR, 3rd, Zimmerman JM, Almazar-Elder AE, Talley NJ. Familial aggregation of irritable bowel syndrome: a family case-control study. *Am J Gastroenterol*. 2010;105(4):833-841.
17. Saito YA, Zimmerman JM, Harmsen WS, De Andrade M, Locke GR, 3rd, Petersen GM, Talley NJ. Irritable bowel syndrome aggregates strongly in families: a family-based case-control study. *Neurogastroenterol Motil*. 2008;20(7):790-797.
18. Locke GR, 3rd, Zinsmeister AR, Talley NJ, Fett SL, Melton LJ, 3rd. Familial association in adults with functional gastrointestinal disorders. *Mayo Clin Proc*. 2000;75(9):907-912.
19. Waehrens R, Ohlsson H, Sundquist J, Sundquist K, Zoller B. Risk of irritable bowel syndrome in first-degree, second-degree and third-degree relatives of affected individuals: a nationwide family study in Sweden. *Gut*. 2015;64(2):215-221.
20. Morris-Yates A, Talley NJ, Boyce PM, Nandurkar S, Andrews G. Evidence of a genetic contribution to functional bowel disorder. *Am J Gastroenterol*. 1998;93(8):1311-1317.
21. Bengtson MB, Ronning T, Vatn MH, Harris JR. Irritable bowel syndrome in twins: genes and environment. *Gut*. 2006;55(12):1754-1759.
22. Levy RL, Jones KR, Whitehead WE, Feld SI, Talley NJ, Corey LA. Irritable bowel syndrome in twins: heredity and social learning both contribute to etiology. *Gastroenterology*. 2001;121(4):799-804.
23. Levy RL. Exploring the intergenerational transmission of illness behavior: from observations to experimental intervention. *Ann Behav Med*. 2011;41(2):174-182.
24. Levy RL, Whitehead WE, Von Korff MR, Feld AD. Intergenerational transmission of gastrointestinal illness behavior. *Am J Gastroenterol*. 2000;95(2):451-456.
25. Levy RL, Whitehead WE, Walker LS, Von Korff M, Feld AD, Garner M, Christie D. Increased somatic complaints and health-care utilization in children: effects of parent IBS status and parent response to gastrointestinal symptoms. *Am J Gastroenterol*. 2004;99(12):2442-2451.
26. Whitehead WE, Winget C, Fedoravicius AS, Wooley S, Blackwell B. Learned illness behavior in patients with irritable bowel syndrome and peptic ulcer. *Dig Dis Sci*. 1982;27(3):202-208.
27. Barbara G, Grover M, Bercik P, Corsetti M, Ghoshal UC, Ohman L, Rajilic-Stojanovic M. Rome Foundation Working Team Report on Post-Infection Irritable Bowel Syndrome. *Gastroenterology*. 2019;156(1):46-58 e47.
28. Klem F, Wadhwa A, Prokop LJ, Sundt WJ, Farrugia G, Camilleri M, Singh S, Grover M. Prevalence, Risk Factors, and Outcomes of Irritable Bowel Syndrome After Infectious

- Enteritis: A Systematic Review and Meta-analysis. *Gastroenterology*. 2017;152(5):1042-1054 e1041.
29. Whitehead WE, Crowell MD, Robinson JC, Heller BR, Schuster MM. Effects of stressful life events on bowel symptoms: subjects with irritable bowel syndrome compared with subjects without bowel dysfunction. *Gut*. 1992;33(6):825-830.
 30. Chitkara DK, van Tilburg MA, Blois-Martin N, Whitehead WE. Early life risk factors that contribute to irritable bowel syndrome in adults: a systematic review. *Am J Gastroenterol*. 2008;103(3):765-774.
 31. Bradford K, Shih W, Videlock EJ, Presson AP, Naliboff BD, Mayer EA, Chang L. Association between early adverse life events and irritable bowel syndrome. *Clin Gastroenterol Hepatol*. 2012;10(4):385-390.e381-383.
 32. Dunlop SP, Jenkins D, Neal KR, Spiller RC. Relative importance of enterochromaffin cell hyperplasia, anxiety, and depression in postinfectious IBS. *Gastroenterology*. 2003;125(6):1651-1659.
 33. Talley NJ, Boyce P, Jones M. Identification of distinct upper and lower gastrointestinal symptom groupings in an urban population. *Gut*. 1998;42(5):690-695.
 34. Kellow JE. The 'pro' case. The Rome III criteria. *Neurogastroenterol Motil*. 2007;19(10):787-792.
 35. Tytgat GN. Foreword II. In: Drossman D, Corazziari E, Delvaux M, Spiller R, Talley N, Thompson WG, editors. Rome III The Functional Gastrointestinal Disorders. McLean, VA: Degnon Associates, Inc; 2006. p. 639-685.
 36. Whitehead WE, Drossman DA. Validation of symptom-based diagnostic criteria for irritable bowel syndrome: a critical review. *Am J Gastroenterol*. 2010;105(4):814-820; quiz 813, 821.
 37. Lacy BE, Mearin F, Chang L, Chey WD, Lembo AJ, Simren M, Spiller R. Bowel Disorders. *Gastroenterology*. 2016, 10.1053/j.gastro.2016.02.031.
 38. Whitehead WE, Palsson OS, Simren M. Irritable bowel syndrome: what do the new Rome IV diagnostic guidelines mean for patient management? *Expert Rev Gastroenterol Hepatol*. 2017;11(4):281-283.
 39. Pimentel M, Hwang L, Melmed GY, Low K, Vasiliauskas E, Ippoliti A, Yang J, Lezcano S, Conklin JL, Sahakian A. New clinical method for distinguishing D-IBS from other gastrointestinal conditions causing diarrhea: the LA/IBS diagnostic strategy. *Dig Dis Sci*. 2010;55(1):145-149.
 40. Palsson OS, Baggish J, Whitehead WE. Episodic nature of symptoms in irritable bowel syndrome. *Am J Gastroenterol*. 2014;109(9):1450-1460.
 41. Mayer EA, Tillisch K. The brain-gut axis in abdominal pain syndromes. *Annu Rev Med*. 2011;62:381-396.

42. Blackshaw LA, Brookes SJ, Grundy D, Schemann M. Sensory transmission in the gastrointestinal tract. *Neurogastroenterol Motil.* 2007;19(1 Suppl):1-19.
43. Liu CN, Devor M, Waxman SG, Kocsis JD. Subthreshold oscillations induced by spinal nerve injury in dissociated muscle and cutaneous afferents of mouse DRG. *J Neurophysiol.* 2002;87(4):2009-2017.
44. Grundy D, Al-Chaer ED, Aziz Q, Collins SM, Ke MY, Tache Y, Wood JD. Fundamentals of neurogastroenterology: Basic science. *Gastroenterology.* 2006;130(5):1391-1411.
45. Feng B, La JH, Schwartz ES, Gebhart GF. Irritable bowel syndrome: methods, mechanisms, and pathophysiology. Neural and neuro-immune mechanisms of visceral hypersensitivity in irritable bowel syndrome. *Am J Physiol Gastrointest Liver Physiol.* 2012;302(10):G1085-1098.
46. Berthoud HR, Blackshaw LA, Brookes SJ, Grundy D. Neuroanatomy of extrinsic afferents supplying the gastrointestinal tract. *Neurogastroenterol Motil.* 2004;16 Suppl 1:28-33.
47. Watkins LR, Maier SF. Immune regulation of central nervous system functions: from sickness responses to pathological pain. *J Intern Med.* 2005;257(2):139-155.
48. Al-Chaer ED, Westlund KN, Willis WD. Sensitization of postsynaptic dorsal column neuronal responses by colon inflammation. *Neuroreport.* 1997;8(15):3267-3273.
49. Ludidi S, Conchillo JM, Keszthelyi D, Van Avesaat M, Kruimel JW, Jonkers DM, Masclee AA. Rectal hypersensitivity as hallmark for irritable bowel syndrome: defining the optimal cutoff. *Neurogastroenterol Motil.* 2012;24(8):729-733, e345-726.
50. Mayer EA. The neurobiology of stress and gastrointestinal disease. *Gut.* 2000;47(6):861-869.
51. Tillisch K, Mayer EA, Labus JS. Quantitative meta-analysis identifies brain regions activated during rectal distension in irritable bowel syndrome. *Gastroenterology.* 2011;140(1):91-100.
52. Mayer EA, Berman S, Suyenobu B, Labus J, Mandelkern MA, Naliboff BD, Chang L. Differences in brain responses to visceral pain between patients with irritable bowel syndrome and ulcerative colitis. *Pain.* 2005;115(3):398-409.
53. Wilder-Smith CH, Schindler D, Lovblad K, Redmond SM, Nirkko A. Brain functional magnetic resonance imaging of rectal pain and activation of endogenous inhibitory mechanisms in irritable bowel syndrome patient subgroups and healthy controls. *Gut.* 2004;53(11):1595-1601.
54. Verne GN, Robinson ME, Vase L, Price DD. Reversal of visceral and cutaneous hyperalgesia by local rectal anesthesia in irritable bowel syndrome (IBS) patients. *Pain.* 2003;105(1-2):223-230.
55. Bouin M, Delvaux M, Blanc C, Lagier E, Delisle MB, Fioramonti J, Bueno L, Frexinos J. Intrarectal injection of glycerol induces hypersensitivity to rectal distension in healthy

- subjects without modifying rectal compliance. *Eur J Gastroenterol Hepatol*. 2001;13(5):573-580.
56. Akbar A, Yiangou Y, Facer P, Walters JR, Anand P, Ghosh S. Increased capsaicin receptor TRPV1-expressing sensory fibres in irritable bowel syndrome and their correlation with abdominal pain. *Gut*. 2008;57(7):923-929.
 57. van Wanrooij SJ, Wouters MM, Van Oudenhove L, Vanbrabant W, Mondelaers S, Kollmann P, Kreutz F, Schemann M, Boeckxstaens GE. Sensitivity testing in irritable bowel syndrome with rectal capsaicin stimulations: role of TRPV1 upregulation and sensitization in visceral hypersensitivity? *Am J Gastroenterol*. 2014;109(1):99-109.
 58. Cenac N, Altier C, Motta JP, d'Aldebert E, Galeano S, Zamponi GW, Vergnolle N. Potentiation of TRPV4 signalling by histamine and serotonin: an important mechanism for visceral hypersensitivity. *Gut*. 2010;59(4):481-488.
 59. Zhao JH, Dong L, Shi HT, Wang ZY, Shi HY, Ding H. The expression of protease-activated receptor 2 and 4 in the colon of irritable bowel syndrome patients. *Dig Dis Sci*. 2012;57(1):58-64.
 60. Balemans D, Aguilera-Lizarraga J, Florens MV, Jain P, Denadai-Souza A, Viola MF, Alpizar YA, Van Der Merwe S, Vanden Berghe P, Talavera K, Vanner S, Wouters MM, Boeckxstaens GE. Histamine-mediated potentiation of transient receptor potential (TRP) ankyrin 1 and TRP vanilloid 4 signaling in submucosal neurons in patients with irritable bowel syndrome. *Am J Physiol Gastrointest Liver Physiol*. 2019;316(3):G338-G349.
 61. Desormeaux C, Bautzova T, Garcia-Caraballo S, Rolland C, Barbaro MR, Brierley SM, Barbara G, Vergnolle N, Cenac N. Protease-activated receptor 1 is implicated in irritable bowel syndrome mediators-induced signaling to thoracic human sensory neurons. *Pain*. 2018;159(7):1257-1267.
 62. Guarino MP, Barbara G, Cicenía A, Altomare A, Barbaro MR, Cocca S, Scirocco A, Cremon C, Emerenziani S, Stanghellini V, Cicala M, Severi C. Supernatants of irritable bowel syndrome mucosal biopsies impair human colonic smooth muscle contractility. *Neurogastroenterol Motil*. 2017;29(2).
 63. Cenac N, Bautzova T, Le Faouder P, Veldhuis NA, Poole DP, Rolland C, Bertrand J, Liedtke W, Dubourdeau M, Bertrand-Michel J, Zecchi L, Stanghellini V, Bunnett NW, Barbara G, Vergnolle N. Quantification and Potential Functions of Endogenous Agonists of Transient Receptor Potential Channels in Patients With Irritable Bowel Syndrome. *Gastroenterology*. 2015;149(2):433-444 e437.
 64. Cenac N, Andrews CN, Holzhausen M, Chapman K, Cottrell G, Andrade-Gordon P, Steinhoff M, Barbara G, Beck P, Bunnett NW, Sharkey KA, Ferraz JG, Shaffer E, Vergnolle N. Role for protease activity in visceral pain in irritable bowel syndrome. *J Clin Invest*. 2007;117(3):636-647.
 65. Barbara G, Wang B, Stanghellini V, de Giorgio R, Cremon C, Di Nardo G, Trevisani M, Campi B, Geppetti P, Tonini M, Bunnett NW, Grundy D, Corinaldesi R. Mast cell-dependent excitation of visceral-nociceptive sensory neurons in irritable bowel syndrome. *Gastroenterology*. 2007;132(1):26-37.

66. Buhner S, Li Q, Berger T, Vignali S, Barbara G, De Giorgio R, Stanghellini V, Schemann M. Submucous rather than myenteric neurons are activated by mucosal biopsy supernatants from irritable bowel syndrome patients. *Neurogastroenterol Motil*. 2012;24(12):1134-e1572.
67. Buhner S, Li Q, Vignali S, Barbara G, De Giorgio R, Stanghellini V, Cremon C, Zeller F, Langer R, Daniel H, Michel K, Schemann M. Activation of human enteric neurons by supernatants of colonic biopsy specimens from patients with irritable bowel syndrome. *Gastroenterology*. 2009;137(4):1425-1434.
68. Buhner S, Braak B, Li Q, Kugler EM, Klooker T, Wouters M, Donovan J, Vignali S, Mazzuoli-Weber G, Grundy D, Boeckxstaens G, Schemann M. Neuronal activation by mucosal biopsy supernatants from irritable bowel syndrome patients is linked to visceral sensitivity. *Exp Physiol*. 2014;99(10):1299-1311.
69. Wouters MM, Balemans D, Van Wanrooy S, Dooley J, Cibert-Goton V, Alpizar YA, Valdez-Morales EE, Nasser Y, Van Veldhoven PP, Vanbrabant W, Van der Merwe S, Mols R, Ghesquiere B, Cirillo C, Kortekaas I, Carmeliet P, Peetermans WE, Vermeire S, Rutgeerts P, Augustijns P, Hellings PW, Belmans A, Vanner S, Bulmer DC, Talavera K, Vanden Berghe P, Liston A, Boeckxstaens GE. Histamine Receptor H1-Mediated Sensitization of TRPV1 Mediates Visceral Hypersensitivity and Symptoms in Patients With Irritable Bowel Syndrome. *Gastroenterology*. 2016;150(4):875-887.e879.
70. Piche T, Barbara G, Aubert P, Bruley des Varannes S, Dainese R, Nano JL, Cremon C, Stanghellini V, De Giorgio R, Galmiche JP, Neunlist M. Impaired intestinal barrier integrity in the colon of patients with irritable bowel syndrome: involvement of soluble mediators. *Gut*. 2009;58(2):196-201.
71. Zhou Q, Zhang B, Verne GN. Intestinal membrane permeability and hypersensitivity in the irritable bowel syndrome. *Pain*. 2009;146(1-2):41-46.
72. Camilleri M, Madsen K, Spiller R, Greenwood-Van Meerveld B, Verne GN. Intestinal barrier function in health and gastrointestinal disease. *Neurogastroenterol Motil*. 2012;24(6):503-512.
73. Dunlop SP, Hebden J, Campbell E, Naesdal J, Olbe L, Perkins AC, Spiller RC. Abnormal intestinal permeability in subgroups of diarrhea-predominant irritable bowel syndromes. *Am J Gastroenterol*. 2006;101(6):1288-1294.
74. Marshall JK, Thabane M, Garg AX, Clark W, Meddings J, Collins SM, Investigators WEL. Intestinal permeability in patients with irritable bowel syndrome after a waterborne outbreak of acute gastroenteritis in Walkerton, Ontario. *Aliment Pharmacol Ther*. 2004;20(11-12):1317-1322.
75. Gecse K, Roka R, Sera T, Rosztoczy A, Annahazi A, Izbeki F, Nagy F, Molnar T, Szepes Z, Pavics L, Bueno L, Wittmann T. Leaky gut in patients with diarrhea-predominant irritable bowel syndrome and inactive ulcerative colitis. *Digestion*. 2012;85(1):40-46.
76. Bertiaux-Vandaele N, Youmba SB, Belmonte L, Lecleire S, Antonietti M, Gourcerol G, Leroi AM, Dechelotte P, Menard JF, Ducrotte P, Coeffier M. The expression and the

cellular distribution of the tight junction proteins are altered in irritable bowel syndrome patients with differences according to the disease subtype. *Am J Gastroenterol*. 2011;106(12):2165-2173.

77. Martinez C, Rodino-Janeiro BK, Lobo B, Stanifer ML, Klaus B, Granzow M, Gonzalez-Castro AM, Salvo-Romero E, Alonso-Cotner C, Pigrau M, Roeth R, Rappold G, Huber W, Gonzalez-Silos R, Lorenzo J, de Torres I, Azpiroz F, Boulant S, Vicario M, Niesler B, Santos J. miR-16 and miR-125b are involved in barrier function dysregulation through the modulation of claudin-2 and cingulin expression in the jejunum in IBS with diarrhoea. *Gut*. 2017;66(9):1537-1538.
78. Spiller RC, Jenkins D, Thornley JP, Hebden JM, Wright T, Skinner M, Neal KR. Increased rectal mucosal enteroendocrine cells, T lymphocytes, and increased gut permeability following acute *Campylobacter* enteritis and in post-dysenteric irritable bowel syndrome. *Gut*. 2000;47(6):804-811.
79. Zhou Q, Souba WW, Croce CM, Verne GN. MicroRNA-29a regulates intestinal membrane permeability in patients with irritable bowel syndrome. *Gut*. 2010;59(6):775-784.
80. Gwee KA, Collins SM, Read NW, Rajnakova A, Deng Y, Graham JC, McKendrick MW, Moolhala SM. Increased rectal mucosal expression of interleukin 1beta in recently acquired post-infectious irritable bowel syndrome. *Gut*. 2003;52(4):523-526.
81. Balemans D, Mondelaers SU, Cibert-Goton V, Stakenborg N, Aguilera-Lizarraga J, Dooley J, Liston A, Bulmer DC, Vanden Berghe P, Boeckstaens GE, Wouters MM. Evidence for long-term sensitization of the bowel in patients with post-infectious-IBS. *Sci Rep*. 2017;7(1):13606.
82. Bashashati M, Moossavi S, Cremon C, Barbaro MR, Moraveji S, Talmon G, Rezaei N, Hughes PA, Bian ZX, Choi CH, Lee OY, Coeffier M, Chang L, Ohman L, Schmulson MJ, McCallum RW, Simren M, Sharkey KA, Barbara G. Colonic immune cells in irritable bowel syndrome: A systematic review and meta-analysis. *Neurogastroenterol Motil*. 2018;30(1).
83. Barbara G, Stanghellini V, De Giorgio R, Cremon C, Cottrell GS, Santini D, Pasquinelli G, Morselli-Labate AM, Grady EF, Bunnett NW, Collins SM, Corinaldesi R. Activated mast cells in proximity to colonic nerves correlate with abdominal pain in irritable bowel syndrome. *Gastroenterology*. 2004;126(3):693-702.
84. Lee KJ, Kim YB, Kim JH, Kwon HC, Kim DK, Cho SW. The alteration of enterochromaffin cell, mast cell, and lamina propria T lymphocyte numbers in irritable bowel syndrome and its relationship with psychological factors. *J Gastroenterol Hepatol*. 2008;23(11):1689-1694.
85. Vivinus-Nebot M, Dainese R, Anty R, Saint-Paul MC, Nano JL, Gonthier N, Marjoux S, Frin-Mathy G, Bernard G, Hebuterne X, Tran A, Theodorou V, Piche T. Combination of allergic factors can worsen diarrheic irritable bowel syndrome: role of barrier defects and mast cells. *Am J Gastroenterol*. 2012;107(1):75-81.

86. Klooker TK, Braak B, Koopman KE, Welting O, Wouters MM, van der Heide S, Schemann M, Bischoff SC, van den Wijngaard RM, Boeckxstaens GE. The mast cell stabiliser ketotifen decreases visceral hypersensitivity and improves intestinal symptoms in patients with irritable bowel syndrome. *Gut*. 2010;59(9):1213-1221.
87. Chang L, Adeyemo M, Karagiannides I, Videlock EJ, Bowe C, Shih W, Presson AP, Yuan PQ, Cortina G, Gong H, Singh S, Licudine A, Mayer M, Tache Y, Pothoulakis C, Mayer EA. Serum and colonic mucosal immune markers in irritable bowel syndrome. *Am J Gastroenterol*. 2012;107(2):262-272.
88. Holmen N, Isaksson S, Simren M, Sjøvall H, Ohman L. CD4+CD25+ regulatory T cells in irritable bowel syndrome patients. *Neurogastroenterol Motil*. 2007;19(2):119-125.
89. Braak B, Klooker TK, Wouters MM, Welting O, van der Loos CM, Stanisor OI, van Diest S, van den Wijngaard RM, Boeckxstaens GE. Mucosal immune cell numbers and visceral sensitivity in patients with irritable bowel syndrome: is there any relationship? *Am J Gastroenterol*. 2012;107(5):715-726.
90. Ohman L, Isaksson S, Lindmark AC, Posserud I, Stotzer PO, Strid H, Sjøvall H, Simren M. T-cell activation in patients with irritable bowel syndrome. *Am J Gastroenterol*. 2009;104(5):1205-1212.
91. Bennet SMP, Palsson O, Whitehead WE, Barrow DA, Tornblom H, Ohman L, Simren M, van Tilburg MAL. Systemic cytokines are elevated in a subset of patients with irritable bowel syndrome but largely unrelated to symptom characteristics. *Neurogastroenterol Motil*. 2018;30(10):e13378.
92. Darkoh C, Comer L, Zewdie G, Harold S, Snyder N, Dupont HL. Chemotactic chemokines are important in the pathogenesis of irritable bowel syndrome. *PLoS One*. 2014;9(3):e93144.
93. Seyedmirzaee S, Hayatbakhsh MM, Ahmadi B, Baniasadi N, Bagheri Rafsanjani AM, Nikpoor AR, Mohammadi M. Serum immune biomarkers in irritable bowel syndrome. *Clin Res Hepatol Gastroenterol*. 2016;40(5):631-637.
94. Vara EJ, Brokstad KA, Hausken T, Lied GA. Altered levels of cytokines in patients with irritable bowel syndrome are not correlated with fatigue. *Int J Gen Med*. 2018;11:285-291.
95. Nasser Y, Petes C, Simmers C, Basso L, Altier C, Gee K, Vanner SJ. Activation of Peripheral Blood CD4+ T-Cells in IBS is not Associated with Gastrointestinal or Psychological Symptoms. *Sci Rep*. 2019;9(1):3710.
96. Ohman L, Lindmark AC, Isaksson S, Posserud I, Strid H, Sjøvall H, Simren M. Increased TLR2 expression on blood monocytes in irritable bowel syndrome patients. *Eur J Gastroenterol Hepatol*. 2012;24(4):398-405.
97. Liebrechts T, Adam B, Bredack C, Roth A, Heinzl S, Lester S, Downie-Doyle S, Smith E, Drew P, Talley NJ, Holtmann G. Immune activation in patients with irritable bowel syndrome. *Gastroenterology*. 2007;132(3):913-920.

98. Macsharry J, O'Mahony L, Fanning A, Bairead E, Sherlock G, Tiesman J, Fulmer A, Kiely B, Dinan TG, Shanahan F, Quigley EM. Mucosal cytokine imbalance in irritable bowel syndrome. *Scand J Gastroenterol*. 2008;43(12):1467-1476.
99. Gao J. Correlation between anxiety-depression status and cytokines in diarrhea-predominant irritable bowel syndrome. *Exp Ther Med*. 2013;6(1):93-96.
100. Kirwan WO, Smith AN, Mitchell WD, Falconer JD, Eastwood MA. Bile acids and colonic motility in the rabbit and the human. *Gut*. 1975;16(11):894-902.
101. Edwards CA, Brown S, Baxter AJ, Bannister JJ, Read NW. Effect of bile acid on anorectal function in man. *Gut*. 1989;30(3):383-386.
102. Li WT, Luo QQ, Wang B, Chen X, Yan XJ, Qiu HY, Chen SL. Bile acids induce visceral hypersensitivity via mucosal mast cell-to-nociceptor signaling that involves the farnesoid X receptor/nerve growth factor/transient receptor potential vanilloid 1 axis. *FASEB J*. 2019;33(2):2435-2450.
103. Valentin N, Camilleri M, Altayar O, Vijayvargiya P, Acosta A, Nelson AD, Murad MH. Biomarkers for bile acid diarrhoea in functional bowel disorder with diarrhoea: a systematic review and meta-analysis. *Gut*. 2016;65(12):1951-1959.
104. Aziz I, Mumtaz S, Bholah H, Chowdhury FU, Sanders DS, Ford AC. High Prevalence of Idiopathic Bile Acid Diarrhea Among Patients With Diarrhea-Predominant Irritable Bowel Syndrome Based on Rome III Criteria. *Clin Gastroenterol Hepatol*. 2015;13(9):1650-1655 e1652.
105. Slattery SA, Niaz O, Aziz Q, Ford AC, Farmer AD. Systematic review with meta-analysis: the prevalence of bile acid malabsorption in the irritable bowel syndrome with diarrhoea. *Aliment Pharmacol Ther*. 2015;42(1):3-11.
106. Vijayvargiya P, Camilleri M, Burton D, Busciglio I, Lueke A, Donato LJ. Bile and fat excretion are biomarkers of clinically significant diarrhoea and constipation in irritable bowel syndrome. *Aliment Pharmacol Ther*. 2019;49(6):744-758.
107. Peleman C, Camilleri M, Busciglio I, Burton D, Donato L, Zinsmeister AR. Colonic Transit and Bile Acid Synthesis or Excretion in Patients With Irritable Bowel Syndrome-Diarrhea Without Bile Acid Malabsorption. *Clin Gastroenterol Hepatol*. 2017;15(5):720-727 e721.
108. Vijayvargiya P, Busciglio I, Burton D, Donato L, Lueke A, Camilleri M. Bile Acid Deficiency in a Subgroup of Patients With Irritable Bowel Syndrome With Constipation Based on Biomarkers in Serum and Fecal Samples. *Clin Gastroenterol Hepatol*. 2018;16(4):522-527.
109. Gershon MD, Tack J. The serotonin signaling system: from basic understanding to drug development for functional GI disorders. *Gastroenterology*. 2007;132(1):397-414.
110. Atkinson W, Lockhart S, Whorwell PJ, Keevil B, Houghton LA. Altered 5-hydroxytryptamine signaling in patients with constipation- and diarrhea-predominant irritable bowel syndrome. *Gastroenterology*. 2006;130(1):34-43.

111. Houghton LA, Atkinson W, Lockhart C, Whorwell PJ, Keevil B. Sigmoid-colonic motility in health and irritable bowel syndrome: a role for 5-hydroxytryptamine. *Neurogastroenterol Motil.* 2007;19(9):724-731.
112. Simren M, Barbara G, Flint HJ, Spiegel BM, Spiller RC, Vanner S, Verdu EF, Whorwell PJ, Zoetendal EG, Rome Foundation C. Intestinal microbiota in functional bowel disorders: a Rome foundation report. *Gut.* 2013;62(1):159-176.
113. Pittayanon R, Lau JT, Yuan Y, Leontiadis GI, Tse F, Surette M, Moayyedi P. Gut Microbiota in Patients With Irritable Bowel Syndrome-a Systematic Review. *Gastroenterology.* 2019, 10.1053/j.gastro.2019.03.049.
114. Rousseaux C, Thuru X, Gelot A, Barnich N, Neut C, Dubuquoy L, Dubuquoy C, Merour E, Geboes K, Chamaillard M, Ouwehand A, Leyer G, Carcano D, Colombel JF, Ardid D, Desreumaux P. Lactobacillus acidophilus modulates intestinal pain and induces opioid and cannabinoid receptors. *Nat Med.* 2007;13(1):35-37.
115. Pimentel M, Park S, Mirocha J, Kane SV, Kong Y. The effect of a nonabsorbed oral antibiotic (rifaximin) on the symptoms of the irritable bowel syndrome: a randomized trial. *Ann Intern Med.* 2006;145(8):557-563.
116. Hustoft TN, Hausken T, Ystad SO, Valeur J, Brokstad K, Hatlebakk JG, Lied GA. Effects of varying dietary content of fermentable short-chain carbohydrates on symptoms, fecal microenvironment, and cytokine profiles in patients with irritable bowel syndrome. *Neurogastroenterol Motil.* 2016, 10.1111/nmo.12969.
117. Villarreal AA, Aberger FJ, Benrud R, Gundrum JD. Use of broad-spectrum antibiotics and the development of irritable bowel syndrome. *WMJ.* 2012;111(1):17-20.
118. Verdu EF, Bercik P, Verma-Gandhu M, Huang XX, Blennerhassett P, Jackson W, Mao Y, Wang L, Rochat F, Collins SM. Specific probiotic therapy attenuates antibiotic induced visceral hypersensitivity in mice. *Gut.* 2006;55(2):182-190.
119. Kim HJ, Camilleri M, McKinzie S, Lempke MB, Burton DD, Thomforde GM, Zinsmeister AR. A randomized controlled trial of a probiotic, VSL#3, on gut transit and symptoms in diarrhoea-predominant irritable bowel syndrome. *Aliment Pharmacol Ther.* 2003;17(7):895-904.
120. Zareie M, Johnson-Henry K, Jury J, Yang PC, Ngan BY, McKay DM, Soderholm JD, Perdue MH, Sherman PM. Probiotics prevent bacterial translocation and improve intestinal barrier function in rats following chronic psychological stress. *Gut.* 2006;55(11):1553-1560.
121. Round JL, Mazmanian SK. The gut microbiota shapes intestinal immune responses during health and disease. *Nat Rev Immunol.* 2009;9(5):313-323.
122. O'Mahony L, McCarthy J, Kelly P, Hurley G, Luo F, Chen K, O'Sullivan GC, Kiely B, Collins JK, Shanahan F, Quigley EM. Lactobacillus and bifidobacterium in irritable bowel syndrome: symptom responses and relationship to cytokine profiles. *Gastroenterology.* 2005;128(3):541-551.

123. Husebye E, Hellstrom PM, Sundler F, Chen J, Midtvedt T. Influence of microbial species on small intestinal myoelectric activity and transit in germ-free rats. *Am J Physiol Gastrointest Liver Physiol*. 2001;280(3):G368-380.
124. Martin CR, Osadchiy V, Kalani A, Mayer EA. The Brain-Gut-Microbiome Axis. *Cell Mol Gastroenterol Hepatol*. 2018;6(2):133-148.
125. Foster JA, Rinaman L, Cryan JF. Stress & the gut-brain axis: Regulation by the microbiome. *Neurobiol Stress*. 2017;7:124-136.
126. Yano JM, Yu K, Donaldson GP, Shastri GG, Ann P, Ma L, Nagler CR, Ismagilov RF, Mazmanian SK, Hsiao EY. Indigenous bacteria from the gut microbiota regulate host serotonin biosynthesis. *Cell*. 2015;161(2):264-276.
127. Bennett EJ, Tennant CC, Piesse C, Badcock CA, Kellow JE. Level of chronic life stress predicts clinical outcome in irritable bowel syndrome. *Gut*. 1998;43(2):256-261.
128. Gwee KA, Leong YL, Graham C, McKendrick MW, Collins SM, Walters SJ, Underwood JE, Read NW. The role of psychological and biological factors in postinfective gut dysfunction. *Gut*. 1999;44(3):400-406.
129. Posserud I, Agerforz P, Ekman R, Bjornsson ES, Abrahamsson H, Simren M. Altered visceral perceptual and neuroendocrine response in patients with irritable bowel syndrome during mental stress. *Gut*. 2004;53(8):1102-1108.
130. Dickhaus B, Mayer EA, Firooz N, Stains J, Conde F, Olivas TI, Fass R, Chang L, Mayer M, Naliboff BD. Irritable bowel syndrome patients show enhanced modulation of visceral perception by auditory stress. *Am J Gastroenterol*. 2003;98(1):135-143.
131. Vanuytsel T, van Wanrooy S, Vanheel H, Vanormelingen C, Verschueren S, Houben E, Salim Rasoel S, Tomicronth J, Holvoet L, Farre R, Van Oudenhove L, Boeckstaens G, Verbeke K, Tack J. Psychological stress and corticotropin-releasing hormone increase intestinal permeability in humans by a mast cell-dependent mechanism. *Gut*. 2014;63(8):1293-1299.
132. Welgan P, Meshkinpour H, Hoehler F. The effect of stress on colon motor and electrical activity in irritable bowel syndrome. *Psychosom Med*. 1985;47(2):139-149.
133. Fukudo S, Nomura T, Hongo M. Impact of corticotropin-releasing hormone on gastrointestinal motility and adrenocorticotrophic hormone in normal controls and patients with irritable bowel syndrome. *Gut*. 1998;42(6):845-849.
134. Anda RF, Felitti VJ, Bremner JD, Walker JD, Whitfield C, Perry BD, Dube SR, Giles WH. The enduring effects of abuse and related adverse experiences in childhood. A convergence of evidence from neurobiology and epidemiology. *Eur Arch Psychiatry Clin Neurosci*. 2006;256(3):174-186.
135. Anda RF, Brown DW, Dube SR, Bremner JD, Felitti VJ, Giles WH. Adverse childhood experiences and chronic obstructive pulmonary disease in adults. *Am J Prev Med*. 2008;34(5):396-403.

136. O'Mahony SM, Hyland NP, Dinan TG, Cryan JF. Maternal separation as a model of brain-gut axis dysfunction. *Psychopharmacology (Berl)*. 2011;214(1):71-88.
137. Heim C, Nemeroff CB. The role of childhood trauma in the neurobiology of mood and anxiety disorders: preclinical and clinical studies. *Biol Psychiatry*. 2001;49(12):1023-1039.
138. Coutinho SV, Plotsky PM, Sablad M, Miller JC, Zhou H, Bayati AI, McRoberts JA, Mayer EA. Neonatal maternal separation alters stress-induced responses to viscerosomatic nociceptive stimuli in rat. *Am J Physiol Gastrointest Liver Physiol*. 2002;282(2):G307-316.
139. Chung EK, Zhang X, Li Z, Zhang H, Xu H, Bian Z. Neonatal maternal separation enhances central sensitivity to noxious colorectal distention in rat. *Brain Res*. 2007;1153:68-77.
140. O'Mahony SM, Marchesi JR, Scully P, Codling C, Ceolho AM, Quigley EM, Cryan JF, Dinan TG. Early life stress alters behavior, immunity, and microbiota in rats: implications for irritable bowel syndrome and psychiatric illnesses. *Biol Psychiatry*. 2009;65(3):263-267.
141. Soderholm JD, Yates DA, Gareau MG, Yang PC, MacQueen G, Perdue MH. Neonatal maternal separation predisposes adult rats to colonic barrier dysfunction in response to mild stress. *Am J Physiol Gastrointest Liver Physiol*. 2002;283(6):G1257-1263.
142. Elzinga BM, Roelofs K, Tollenaar MS, Bakvis P, van Pelt J, Spinhoven P. Diminished cortisol responses to psychosocial stress associated with lifetime adverse events a study among healthy young subjects. *Psychoneuroendocrinology*. 2008;33(2):227-237.
143. Rinne T, de Kloet ER, Wouters L, Goekoop JG, DeRijk RH, van den Brink W. Hyperresponsiveness of hypothalamic-pituitary-adrenal axis to combined dexamethasone/corticotropin-releasing hormone challenge in female borderline personality disorder subjects with a history of sustained childhood abuse. *Biol Psychiatry*. 2002;52(11):1102-1112.
144. Carpenter LL, Carvalho JP, Tyrka AR, Wier LM, Mello AF, Mello MF, Anderson GM, Wilkinson CW, Price LH. Decreased adrenocorticotropic hormone and cortisol responses to stress in healthy adults reporting significant childhood maltreatment. *Biol Psychiatry*. 2007;62(10):1080-1087.
145. McGowan PO, Sasaki A, D'Alessio AC, Dymov S, Labonte B, Szyf M, Turecki G, Meaney MJ. Epigenetic regulation of the glucocorticoid receptor in human brain associates with childhood abuse. *Nat Neurosci*. 2009;12(3):342-348.
146. Labonte B, Yerko V, Gross J, Mechawar N, Meaney MJ, Szyf M, Turecki G. Differential glucocorticoid receptor exon 1(B), 1(C), and 1(H) expression and methylation in suicide completers with a history of childhood abuse. *Biol Psychiatry*. 2012;72(1):41-48.
147. Moloney RD, Stilling RM, Dinan TG, Cryan JF. Early-life stress-induced visceral hypersensitivity and anxiety behavior is reversed by histone deacetylase inhibition. *Neurogastroenterol Motil*. 2015;27(12):1831-1836.

148. Kano M, Muratsubaki T, Van Oudenhove L, Morishita J, Yoshizawa M, Kohno K, Yagihashi M, Tanaka Y, Mugikura S, Dupont P, Ly HG, Takase K, Kanazawa M, Fukudo S. Altered brain and gut responses to corticotropin-releasing hormone (CRH) in patients with irritable bowel syndrome. *Sci Rep.* 2017;7(1):12425.
149. Tache Y, Million M. Role of Corticotropin-releasing Factor Signaling in Stress-related Alterations of Colonic Motility and Hyperalgesia. *J Neurogastroenterol Motil.* 2015;21(1):8-24.
150. Wallon C, Yang PC, Keita AV, Ericson AC, McKay DM, Sherman PM, Perdue MH, Soderholm JD. Corticotropin-releasing hormone (CRH) regulates macromolecular permeability via mast cells in normal human colonic biopsies in vitro. *Gut.* 2008;57(1):50-58.
151. Mayer EA, Tillisch K, Gupta A. Gut/brain axis and the microbiota. *J Clin Invest.* 2015;125(3):926-938.
152. Bailey MT, Lubach GR, Coe CL. Prenatal stress alters bacterial colonization of the gut in infant monkeys. *J Pediatr Gastroenterol Nutr.* 2004;38(4):414-421.
153. Bailey MT, Coe CL. Maternal separation disrupts the integrity of the intestinal microflora in infant rhesus monkeys. *Dev Psychobiol.* 1999;35(2):146-155.
154. Rhee SH, Pothoulakis C, Mayer EA. Principles and clinical implications of the brain-gut-enteric microbiota axis. *Nat Rev Gastroenterol Hepatol.* 2009;6(5):306-314.
155. Golubeva AV, Crampton S, Desbonnet L, Edge D, O'Sullivan O, Lomasney KW, Zhdanov AV, Crispie F, Moloney RD, Borre YE, Cotter PD, Hyland NP, O'Halloran KD, Dinan TG, O'Keeffe GW, Cryan JF. Prenatal stress-induced alterations in major physiological systems correlate with gut microbiota composition in adulthood. *Psychoneuroendocrinology.* 2015;60:58-74.
156. Bercik P, Denou E, Collins J, Jackson W, Lu J, Jury J, Deng Y, Blennerhassett P, Macri J, McCoy KD, Verdu EF, Collins SM. The intestinal microbiota affect central levels of brain-derived neurotropic factor and behavior in mice. *Gastroenterology.* 2011;141(2):599-609, 609 e591-593.
157. Collins SM, Bercik P. The relationship between intestinal microbiota and the central nervous system in normal gastrointestinal function and disease. *Gastroenterology.* 2009;136(6):2003-2014.
158. Grover M, Herfarth H, Drossman DA. The functional-organic dichotomy: postinfectious irritable bowel syndrome and inflammatory bowel disease-irritable bowel syndrome. *Clin Gastroenterol Hepatol.* 2009;7(1):48-53.
159. Kiank C, Tache Y, Larauche M. Stress-related modulation of inflammation in experimental models of bowel disease and post-infectious irritable bowel syndrome: role of corticotropin-releasing factor receptors. *Brain Behav Immun.* 2010;24(1):41-48.
160. Ibeakanma C, Ochoa-Cortes F, Miranda-Morales M, McDonald T, Spreadbury I, Cenac N, Cattaruzza F, Hurlbut D, Vanner S, Bunnett N, Vergnolle N, Vanner S. Brain-gut

- interactions increase peripheral nociceptive signaling in mice with postinfectious irritable bowel syndrome. *Gastroenterology*. 2011;141(6):2098-2108 e2095.
161. Saunders PR, Kosecka U, McKay DM, Perdue MH. Acute stressors stimulate ion secretion and increase epithelial permeability in rat intestine. *Am J Physiol*. 1994;267(5 Pt 1):G794-799.
 162. Santos J, Yang PC, Soderholm JD, Benjamin M, Perdue MH. Role of mast cells in chronic stress induced colonic epithelial barrier dysfunction in the rat. *Gut*. 2001;48(5):630-636.
 163. Saunders PR, Santos J, Hanssen NP, Yates D, Groot JA, Perdue MH. Physical and psychological stress in rats enhances colonic epithelial permeability via peripheral CRH. *Dig Dis Sci*. 2002;47(1):208-215.
 164. Gareau MG, Jury J, Perdue MH. Neonatal maternal separation of rat pups results in abnormal cholinergic regulation of epithelial permeability. *Am J Physiol Gastrointest Liver Physiol*. 2007;293(1):G198-203.
 165. Saito YA. The role of genetics in IBS. *Gastroenterol Clin North Am*. 2011;40(1):45-67.
 166. Henstrom M, Diekmann L, Bonfiglio F, Hadizadeh F, Kuech EM, von Kockritz-Blickwede M, Thingholm LB, Zheng T, Assadi G, Dierks C, Heine M, Philipp U, Distl O, Money ME, Belheouane M, Heinsen FA, Rafter J, Nardone G, Cuomo R, Usai-Satta P, Galeazzi F, Neri M, Walter S, Simren M, Karling P, Ohlsson B, Schmidt PT, Lindberg G, Dlugosz A, Agreus L, Andreasson A, Mayer E, Baines JF, Engstrand L, Portincasa P, Bellini M, Stanghellini V, Barbara G, Chang L, Camilleri M, Franke A, Naim HY, D'Amato M. Functional variants in the sucrase-isomaltase gene associate with increased risk of irritable bowel syndrome. *Gut*. 2018;67(2):263-270.
 167. Beyder A, Mazzone A, Strege PR, Tester DJ, Saito YA, Bernard CE, Enders FT, Ek WE, Schmidt PT, Dlugosz A, Lindberg G, Karling P, Ohlsson B, Gazouli M, Nardone G, Cuomo R, Usai-Satta P, Galeazzi F, Neri M, Portincasa P, Bellini M, Barbara G, Camilleri M, Locke GR, Talley NJ, D'Amato M, Ackerman MJ, Farrugia G. Loss-of-function of the voltage-gated sodium channel NaV1.5 (channelopathies) in patients with irritable bowel syndrome. *Gastroenterology*. 2014;146(7):1659-1668.
 168. Ek WE, Reznichenko A, Ripke S, Niesler B, Zucchelli M, Rivera NV, Schmidt PT, Pedersen NL, Magnusson P, Talley NJ, Holliday EG, Houghton L, Gazouli M, Karamanolis G, Rappold G, Burwinkel B, Surowy H, Rafter J, Assadi G, Li L, Papadaki E, Gambaccini D, Marchi S, Colucci R, Blandizzi C, Barbaro R, Karling P, Walter S, Ohlsson B, Tornblom H, Bresso F, Andreasson A, Dlugosz A, Simren M, Agreus L, Lindberg G, Boeckxstaens G, Bellini M, Stanghellini V, Barbara G, Daly MJ, Camilleri M, Wouters MM, D'Amato M. Exploring the genetics of irritable bowel syndrome: a GWA study in the general population and replication in multinational case-control cohorts. *Gut*. 2015;64(11):1774-1782.
 169. Bonfiglio F, Zheng T, Garcia-Etxebarria K, Hadizadeh F, Bujanda L, Bresso F, Agreus L, Andreasson A, Dlugosz A, Lindberg G, Schmidt PT, Karling P, Ohlsson B, Simren M, Walter S, Nardone G, Cuomo R, Usai-Satta P, Galeazzi F, Neri M, Portincasa P, Bellini M, Barbara G, Latiano A, Hubenthal M, Thijs V, Netea MG, Jonkers D, Chang L, Mayer

- EA, Wouters MM, Boeckxstaens G, Camilleri M, Franke A, Zhernakova A, D'Amato M. Female-Specific Association Between Variants on Chromosome 9 and Self-Reported Diagnosis of Irritable Bowel Syndrome. *Gastroenterology*. 2018;155(1):168-179.
170. Bonfiglio F, Henstrom M, Nag A, Hadizadeh F, Zheng T, Cenit MC, Tigchelaar E, Williams F, Reznichenko A, Ek WE, Rivera NV, Homuth G, Aghdassi AA, Kacprowski T, Mannikko M, Karhunen V, Bujanda L, Rafter J, Wijmenga C, Ronkainen J, Hysi P, Zhernakova A, D'Amato M. A GWAS meta-analysis from 5 population-based cohorts implicates ion channel genes in the pathogenesis of irritable bowel syndrome. *Neurogastroenterol Motil*. 2018;30(9):e13358.
 171. Gazouli M, Wouters MM, Kapur-Pojksic L, Bengtson MB, Friedman E, Nikcevic G, Demetriou CA, Mulak A, Santos J, Niesler B. Lessons learned--resolving the enigma of genetic factors in IBS. *Nat Rev Gastroenterol Hepatol*. 2016;13(2):77-87.
 172. Saito YA, Larson JJ, Atkinson EJ, Ryu E, Almazan AE, Petersen GM, Talley NJ. The role of 5-HTT LPR and GNbeta3 825C>T polymorphisms and gene-environment interactions in irritable bowel syndrome (IBS). *Dig Dis Sci*. 2012;57(10):2650-2657.
 173. Van Kerkhoven LA, Laheij RJ, Jansen JB. Meta-analysis: a functional polymorphism in the gene encoding for activity of the serotonin transporter protein is not associated with the irritable bowel syndrome. *Aliment Pharmacol Ther*. 2007;26(7):979-986.
 174. Wang YM, Chang Y, Chang YY, Cheng J, Li J, Wang T, Zhang QY, Liang DC, Sun B, Wang BM. Serotonin transporter gene promoter region polymorphisms and serotonin transporter expression in the colonic mucosa of irritable bowel syndrome patients. *Neurogastroenterol Motil*. 2012;24(6):560-565, e254-565.
 175. Kohen R, Jarrett ME, Cain KC, Jun SE, Navaja GP, Symonds S, Heitkemper MM. The serotonin transporter polymorphism rs25531 is associated with irritable bowel syndrome. *Dig Dis Sci*. 2009;54(12):2663-2670.
 176. Sato N, Suzuki N, Sasaki A, Aizawa E, Obayashi T, Kanazawa M, Mizuno T, Kano M, Aoki M, Fukudo S. Corticotropin-releasing hormone receptor 1 gene variants in irritable bowel syndrome. *PLoS One*. 2012;7(9):e42450.
 177. Orand A, Naliboff B, Gadd M, Shih W, Ju T, Presson AP, Mayer EA, Chang L. Corticotropin-releasing hormone receptor 1 (CRH-R1) polymorphisms are associated with irritable bowel syndrome and acoustic startle response. *Psychoneuroendocrinology*. 2016;73:133-141.
 178. Park JM, Choi MG, Cho YK, Lee IS, Kim SW, Choi KY, Chung IS. Cannabinoid receptor 1 gene polymorphism and irritable bowel syndrome in the Korean population: a hypothesis-generating study. *J Clin Gastroenterol*. 2011;45(1):45-49.
 179. Karling P, Danielsson A, Wikgren M, Soderstrom I, Del-Favero J, Adolfsson R, Norrback KF. The relationship between the val158met catechol-O-methyltransferase (COMT) polymorphism and irritable bowel syndrome. *PLoS One*. 2011;6(3):e18035.

180. Romero-Valdovinos M, Gudino-Ramirez A, Reyes-Gordillo J, Martinez-Flores WA, Ramirez-Miranda ME, Maravilla P, Olivo-Diaz A. Interleukin-8 and -10 gene polymorphisms in irritable bowel syndrome. *Mol Biol Rep.* 2012;39(9):8837-8843.
181. van der Veek PP, van den Berg M, de Kroon YE, Verspaget HW, Masclee AA. Role of tumor necrosis factor-alpha and interleukin-10 gene polymorphisms in irritable bowel syndrome. *Am J Gastroenterol.* 2005;100(11):2510-2516.
182. Spiller R, Swan C, Campbell E. Identifying and testing candidate genes underlying the inflammatory basis of irritable bowel syndrome. *Gut.* 2011;60(Suppl 1):A164.
183. Mahurkar S, Polytarchou C, Iliopoulos D, Pothoulakis C, Mayer EA, Chang L. Genome-wide DNA methylation profiling of peripheral blood mononuclear cells in irritable bowel syndrome. *Neurogastroenterol Motil.* 2016;28(3):410-422.
184. Wohlfarth C, Schmitteckert S, Hartle JD, Houghton LA, Dweep H, Fortea M, Assadi G, Braun A, Mederer T, Pohner S, Becker PP, Fischer C, Granzow M, Monnikes H, Mayer EA, Sayuk G, Boeckxstaens G, Wouters MM, Simren M, Lindberg G, Ohlsson B, Schmidt PT, Dlugosz A, Agreus L, Andreasson A, D'Amato M, Burwinkel B, Bermejo JL, Roth R, Lasitschka F, Vicario M, Metzger M, Santos J, Rappold GA, Martinez C, Niesler B. miR-16 and miR-103 impact 5-HT4 receptor signalling and correlate with symptom profile in irritable bowel syndrome. *Sci Rep.* 2017;7(1):14680.
185. Zhou Q, Yang L, Larson S, Basra S, Merwat S, Tan A, Croce C, Verne GN. Decreased miR-199 augments visceral pain in patients with IBS through translational upregulation of TRPV1. *Gut.* 2016;65(5):797-805.
186. Videlock EJ, Mahurkar-Joshi S, Hoffman JM, Iliopoulos D, Pothoulakis C, Mayer EA, Chang L. Sigmoid colon mucosal gene expression supports alterations of neuronal signaling in irritable bowel syndrome with constipation. *Am J Physiol Gastrointest Liver Physiol.* 2018;315(1):G140-G157.
187. R Core Team. R: A language and environment for statistical computing. 2013. Vienna, Austria: R Foundation for Statistical Computing. Available from: <http://www.R-project.org/>.

Chapter 2:

**Sigmoid colon mucosal gene expression supports alterations
of neuronal signaling in irritable bowel syndrome with constipation**

ABSTRACT

Objective: Peripheral factors likely play a role in at least a subset of irritable bowel syndrome (IBS) patients. Few studies have investigated mucosal gene expression using an unbiased approach. Here, we performed mucosal gene profiling in a sex-balanced sample to identify relevant signaling pathways and gene networks and compare to publicly available profiling data from additional cohorts.

Design: 20 Rome III+ IBS patients (10 IBS with constipation (IBS-C), 10 IBS with diarrhea (IBS-D) 5 men/women each) and 10 age/sex-matched healthy controls (HCs) underwent sigmoidoscopy with biopsy for gene microarray analysis, including differential expression, weighted gene co-expression network analysis (WGCNA), gene set enrichment analysis and comparison with publicly available data. Expression levels of 67 genes were validated in an expanded cohort, including the above samples and 18 additional participants (6 each IBS-C, IBS-D, HCs) using the NanoString nCounter technology.

Results: There were 1270 differentially expressed genes (FDR<0.05) in IBS-C vs. HCs but none in IBS or IBS-D vs. HCs. WGCNA analysis identified activation of the cAMP/protein kinase A signaling pathway. Nine of 67 genes were validated by Nanostring nCounter technology (FDR<0.05) in the expanded sample. Comparison with publicly available microarray data from the Mayo Clinic and University of Nottingham supports the reproducibility of 17 genes from the microarray analysis and 3 of 9 genes validated by nCounter in IBS-C vs. HCs.

Conclusion: This study supports the involvement of peripheral mechanisms in IBS-C, particularly pathways mediating neuronal signaling.

INTRODUCTION

Irritable bowel syndrome (IBS) is prevalent and costly condition with limited therapies and poorly understood pathophysiology. Affecting around 11% of the general population in the U.S.,⁽⁸⁴⁾ IBS is associated with a significant healthcare and economic burden, and a lack of treatment satisfaction.^(25, 49) Improved understanding of the molecular basis for the dysregulated gastrointestinal function and sensation in IBS may lead to development of targeted therapies and improved quality of life.

There is clear and growing evidence that peripheral mechanisms play a role in IBS pathophysiology, for at least a subset of IBS patients.⁽²⁴⁾ Peripheral mechanisms manifest in a range of phenotypes, including, but not limited to, alterations in intestinal permeability^(41, 132), serotonin signaling^(7, 36, 38, 45, 54, 63, 122), immune milieu, ^(2, 4, 12, 13, 20, 30, 39, 94) bile acids,⁽²¹⁾ microbiota and neuronal sensitivity^(9, 16-18, 27, 57, 95, 118, 124). It is not known whether these phenotypes result from universally deregulated molecular pathways or whether IBS represents a syndrome caused by heterogeneous pathophysiological mechanisms. In order to identify molecular pathways associated with IBS, we conducted a microarray analysis of sigmoid mucosal gene expression in a pilot study of Rome III+ IBS patients, and age and sex-matched healthy controls (HCs) who were free of active psychiatric disease, and had limited use of medications.

The National Institutes of Health recently launched an initiative to improve experimental reproducibility.⁽³⁷⁾ This has notoriously been a challenge in IBS. This likely results from the combination of heterogeneity in both population and study methodology, and pathophysiology resulting from subtle changes – a large sample is needed to detect small differences when standard deviations are large. Alternatively, IBS may not be caused by a universal pathophysiological mechanism. This difficulty with reproducibility may actually be underrepresented in the literature due to publication bias for positive results since very few studies used unbiased approaches (e.g. microarray or RNA-sequencing evaluation of mucosal

gene expression). In order to identify reproducible results within our own cohort, we compared our findings with publicly available data. Our comparison data comes from two microarray studies comparing well-phenotyped IBS patients with HCs. The first study, referred to as “Mayo”, evaluated gene expression in sigmoid biopsies from 36 Rome II+ IBS patients (15 IBS with constipation, IBS-C) and 25 HCs.(2) The second study, “Nottingham”, evaluated rectal biopsies in 56 Rome II+(115) IBS patients (19 IBS-C) and 25 HCs in addition to a group of post-infection IBS patients.(112) The only currently published RNA-sequencing study in IBS,(22) as well as the comprehensive follow-up studies(19, 23) have focused on IBS-D.

Here, we report on the 1) differential expression of transcripts found by microarray; 2) comparison with results from two prior microarray studies (Mayo and Nottingham); 3) weighted gene co-expression network analysis (WGCNA) and gene set enrichment analysis (GSEA) to identify pathways and networks relevant to IBS; and 4) validation of selected genes in a larger cohort. We have also created a website (<http://uclacns.org/ibs-microarray-comparison/>) which display tabular and graphical results of gene expression data from this study, Mayo and Nottingham for any gene present on the arrays.

MATERIALS AND METHODS

Participants and protocol

Rome III+(80) IBS patients and HCs (men and women, aged 18-55) were recruited primarily by community advertisement. Bowel habit subtypes were based on Rome III criteria. The diagnosis was confirmed by a clinician with expertise in IBS. HCs had no personal or family history of IBS or other chronic pain conditions. Additional exclusion criteria for all subjects included: infectious or inflammatory disorders, active psychiatric illness over the past six months assessed by structured clinical interview for the DSM-IV (MINI),(105) use of corticosteroids in the past six months, use of narcotics, antidepressants or other medications that could affect neuroendocrine function in the past two months, or current tobacco or alcohol abuse. Participants were

compensated. The study was approved by the UCLA Institutional Review Board, and all subjects signed a written informed consent prior to starting the study. Overall IBS symptom, abdominal pain, and bloating severity over the prior week were each assessed with a numeric rating scale (0-20).⁽⁹⁰⁾ Current anxiety and depression symptoms were measured with the Hospital Anxiety and Depression (HAD) scale.⁽¹³³⁾ Sigmoid biopsies (30cm) were obtained by flexible sigmoidoscopy following tap water enemas. Specimens were flash frozen in liquid nitrogen and RNA was extracted with TRIzol reagent (ThermoFisher Scientific, Waltham, MA). Participants for the microarray study as well as the additional participants included in the nCounter validation were selected from the same biorepository at UCLA. The only criteria used to select participants for inclusion in both the microarray and validation cohorts was optimal matching of sex and age within each bowel habit subtype.

Microarray

Gene expression was measured using the Arraystar Human LncRNA Microarray V3.0 (Arraystar Inc, Rockville, MD), which profiles over 30,500 LncRNAs and 26,100 mRNAs. RNA quantity and quality were first measured using the NanoDrop ND-1000 (ThermoFisher Scientific) and RNA integrity was assessed by standard denaturing agarose gel electrophoresis. Sample labeling and array hybridization were performed according to the Agilent One-Color Microarray-Based Gene Expression Analysis protocol (Agilent Technologies, Santa Clara, CA) with minor modifications. Briefly, mRNA was purified from total RNA after removal of rRNA (mRNA-ONLY™ Eukaryotic mRNA Isolation Kit, Epicentre, Madison, WI). Then, each sample was amplified and transcribed into fluorescent cRNA along the entire length of the transcripts without 3' bias utilizing a mixture of oligo(dT) and random priming method (Flash RNA Labeling Kit, Arraystar). The labeled cRNAs were purified using the RNeasy Mini Kit (Qiagen, Valencia, CA). The concentration and specific activity of the labeled cRNAs (pmol Cy3/μg cRNA) were measured with the NanoDrop ND-1000. 1 μg of each labeled cRNA was fragmented by adding

5 µl 10 × Blocking Agent and 1 µl of 25 × Fragmentation Buffer, heated to 60 °C for 30 min, and 25 µl 2 × GE Hybridization buffer was added to dilute the labeled cRNA. 50 µl of hybridization solution was dispensed into the gasket slide and assembled to the LncRNA expression microarray slide. The slides were incubated for 17 hours at 65°C in an Agilent Hybridization Oven. The hybridized arrays were washed, fixed and scanned using the Agilent DNA Microarray Scanner (part number G2505C). Agilent Feature Extraction software (version 11.0.1.1) was used to analyze acquired array images. Quantile normalization and subsequent data processing were performed with using the GeneSpring GX v12.0 software package (Agilent Technologies). After quantile normalization of the raw data, LncRNAs and mRNAs that had flags in Present or Marginal (“All Targets Value”) in at least 10 out of 30 samples were chosen for further data analysis. Targets with mean raw signal intensity of <200 in both IBS or HCs were excluded. Data have been deposited in the ArrayExpress database at EMBL-EBI under accession number E-MTAB-5811 (<https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-5811>).

Validation of selected targets

RNA (100ng) from the 30 samples analyzed by microarray and 18 additional samples (identical recruitment and protocol) was analyzed using a custom nCounter panel including 67 targets and three reference genes (*GAPDH*, *BACT*, *PPIA*; NanoString Technologies, Seattle, WA).

Statistical Analysis

Code used for analyses in R is provided in the online Supplement.

Sample Size

Based on the proportion of differentially expressed transcripts (DETs) in the Mayo study,(2) (23/47,000) and an approximate standard deviation of 0.5, we estimated that 8 per group would provide 80% power to detect DETs with a fold change (FC) of ≥2.(78)

Microarray differential expression

DETs were determined using the Linear Models for Microarray Analysis (limma)(107) package in R version 3.3.1(99) following log₂ transformation. An adjusted-p value threshold (Benjamini-Hochberg) of 0.05 determined statistical significance.

Acquisition and analysis of external data

External data (Mayo:E-TABM-176,(2) Nottingham:E-GEOD-36701(112)) were retrieved using the ArrayExpress package(61) and Affymetrix probe annotations were updated (annotate package(47)). Data were normalized using the robust multiarray average function (affy package(46)). The limma package was used to account for technical replicates (duplicateCorrelation function) and to determine DETs.(107)

Determination of microarray probe transcript specificity

The “ReAnnotator” pipeline(5) was used to obtain accurate alignments of Arraystar probe sequences from the current study to the exome. Updated annotations of Affymetrix probes (2015) were downloaded from ThermoFisher. Among non-specific probes in the current study, the probe with the maximum variance across samples was retained. Common results were defined as probes targeting the same transcript or gene and differentially expressed in the same direction (i.e. up- or down-regulated in both) with a threshold of $p < 0.05$.

WGCNA

WGCNA (67) was performed in R on targets with mean raw signal intensity >200 using a soft-thresholding power 14 and minimum module size 30. Highly-correlated modules ($r > 0.75$) were merged.

Functional annotation

Gene Set Enrichment Analysis (GSEA)(111) was performed using the Java GSEA Desktop Application (v2.2.4) on all annotated mRNAs (log fold change and p value were entered) with mean signal intensity >200 in either group (n=14,775) using default parameters against the

following genes sets obtained from Molecular Signatures Database (v6.0)(89): Reactome,(43) KEGG,(60) and Gene Ontology (GO) Biological Process (BP).(6) Overrepresented GO terms for Mayo/Nottingham comparison and WGCNA were determined using the hypergeometric function in the GOstats(44) package in R. These overrepresented terms as well as those yielded from analysis in DAVID(55, 56) were used to select descriptive module names. Pathway analysis was performed using Ingenuity Pathway Analysis (Qiagen, Valencia, CA) using experimentally-observed direct relationships in the Ingenuity Knowledge Base.

NanoString nCounter expression and association with symptoms

nCounter expression data was analyzed with the NanoStringNorm and limma packages in R.(107, 120) Counts were normalized to the geometric mean of all samples to adjust for technical variation between the four cartridges. Due to high variation of the housekeeping genes included in the code set (*GAPDH*, *ACTB*, *PPIA*), RNA quantity was normalized to the geometric mean of genes with the lowest coefficients of variation (*GABBR1*, *RBM5*, *MUC4*, *IL17RE*, *RNF19A*, *ELMOD3*, *USP34*, *CDC42EP5*, *OPLAH*, *SGSM2*) which was determined using NanoStringNorm (results were unchanged when *MUC4*, which was differentially expressed, was removed from the normalization set). Three outliers (mean count z-scores >3) were excluded from analysis (an additional with z-score 3.03 was retained) yielding 45 samples with nCounter data. Using the voom function(69) in the limma package, count data was prepared for analysis with linear models and DETs were determined with limma. An adjusted-p-value threshold of 0.05 was used to determine statistical significance. Association with symptoms was determined using linear models. If sex, age or BMI was a predictor of the outcome variable, it was included in the model (sex for overall symptoms and all 3 for abdominal pain).

RESULTS

Study Population

Characteristics of the study participants are shown in Table 1. The total cohort with available NanoString nCounter gene expression data (excluding outliers, $n=3$) included 31 IBS patients and 14 HCs (51.6% and 42.9% women, respectively). These included a total of 15 IBS-C patients (10 women, 5 men), 16 IBS-D patients (6 women, 10 men) and 14 HCs (6 women, 8 men).

Only 24 participants (all IBS) reported any medication use and most used laxatives/antidiarrheals and ibuprofen only on an as-needed basis. All participants were instructed to hold aspirin and non-steroidal anti-inflammatory drugs (NSAIDs) for 72 hours prior to the procedure. No participants in the microarray cohort reported use of antidepressants/neuromodulators. In the expanded cohort, one participant took sertraline (100 mg/day) and bupropion (150 mg/day) and another took trazodone (50 mg/day). Participants were instructed not to take unnecessary medications 24-48 hours prior to the procedure.

Identification of IBS gene signature by microarray profiling analysis

Differentially expressed transcripts

For IBS vs. HCs and IBS-D vs. HCs, there were no DETs meeting the threshold of $FDR < 0.05$. However, there were 120 and 66 DETs using an unadjusted p value threshold of 0.001. For IBS-C, there were 1,270 DETs with $FDR < 0.05$ (790 up-regulated and 480 down-regulated), mapping to 1,246 specific transcripts (retaining the probe with highest variance across samples among non-specific probes, Supplement 1). A heat map of the top 100 genes is shown in Figure 1A. Since our sample was age- and sex-matched, our primary analysis did not include sex. There were 6 genes for which there was a significant effect of sex ($FDR < 0.05$), and the effect of IBS-C vs. HC remained significant ($FDR < 0.05$) in all six after controlling for sex.

Table 1: Participant characteristics

Variable (range)	Group		Microarray		nCounter	
	Group		IBS	HC	IBS	HC
<i>n</i>			20	10	31	14
%Women			50	50	51.6	42.9
Age			33.8 (10.8)	31.7 (8.5)	34 (12)	35.4 (8.4)
BMI			24.4 (6.4)	24.7 (2.5)	24.5 (5.5)	25.8 (3)
Ethnicity						
Hispanic			1 (5%)	2 (20%)	4 (12.9%)	1 (7.1%)
Race						
Asian			5 (25%)	3 (30%)	8 (25.8%)	3 (21.4%)
Black			5 (10%)	3 (10%)	8 (9.7%)	3 (7.1%)
White			5 (50%)	3 (40%)	8 (48.4%)	3 (57.1%)
Other/mixed			3 (15%)	2 (20%)	5 (16.1%)	2 (14.3%)
Bowel habit subtype						
IBS-C			10 (50%)		15 (48.4%)	
IBS-D			10 (50%)		16 (51.6%)	
HAD anxiety score (0-21)			5.6 (6)	3.3 (2.8)	6.4 (6)*	2.9 (2.4)
HAD depression score (0-21)			3 (3.6)	1.9 (2.1)	3 (3.1)	1.4 (1.8)
Overall IBS severity (0-20)			10.4 (4.4)		10.1 (4)	
Abdominal pain (0-20)			9.8 (4.6)		9.7 (4.2)	

Values are means (SD) or frequency (%). nCounter, counter cohort is inclusive of microarray cohort. HC, healthy control; IBS, irritable bowel syndrome; BMI, body mass index; HAD, Hospital Anxiety and Depression; IBS-C, IBS with constipation; IBS-D, IBS with diarrhea. * $P < 0.05$ between IBS and HCs.

Table 2: External sample and study characteristics

Sample	Mayo			Nottingham		
	IBS-D	IBS-C	HC	IBS-D	IBS-C	HC
<i>n</i>	21	15	25	27	18	26
%Women	86%	100%	92%	67%	89%	62%
Recruitment Protocol (prep)	By mail from database, Rome II Sigmoidoscopy (Fleet's enema)			Outpatient GI clinic, Rome II Sigmoidoscopy (none)		
Biopsy site	Sigmoid colon			Rectum		
Microarray	Human Genome U133-Plus2.0 (Affymetrix)					

Comparison with external gene expression data

Demographics of the samples and study protocols from Mayo Clinic and the University of Nottingham are shown in Table 2. Probes from the Arraystar (Agilent) array used in our study and the Affymetrix array used in Mayo and Nottingham were matched using the following algorithm. First, we identified transcripts or genes matched by only a single probe in each array. There were 452 probes uniquely matched to a specific transcript, and one probe uniquely matched to a specific gene. There were 90 Affymetrix probes that were non-specific and overlapped with more than one Arraystar probe (e.g. two Arraystar probes were transcript-specific with multiple Affymetrix probes matching). In these cases, the affected Arraystar probes were reduced to retain the probe with the highest variance across samples (48 Arraystar probes excluded). This yielded a set of Affymetrix probes that each only matched a single Arraystar probe, though in many cases there were more than one Affymetrix probes matched with each Arraystar probe. The final matches were determined separately for Mayo and Nottingham by reducing the multiple matching Affymetrix probes to the one with the highest variance across samples. Among the 1,246 DETs (IBS-C vs. HCs) in the current study, 859 were matched with probes from the Mayo and Nottingham data. Among these, the number of DETs with the same direction of FC ($p < 0.05$) in IBS-C and HC samples from Mayo and Nottingham were 57 and 185, respectively with 17 matched in all three data sets.

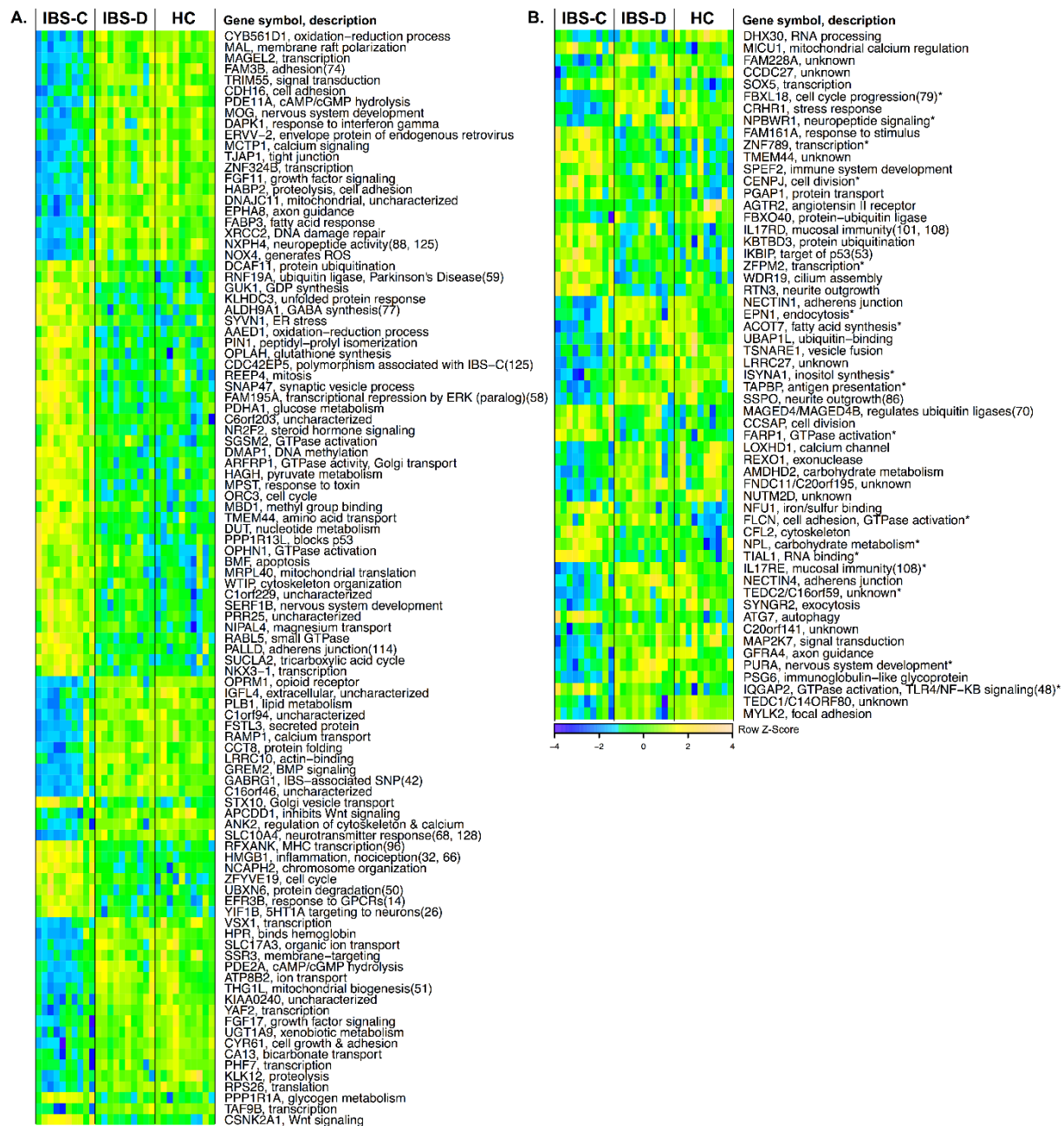


Figure 1: Heatmaps of differentially expressed genes

A. Top 100 differentially expressed genes in IBS-C vs. HC. B. Differentially expressed genes in IBS-C vs. HC that are replicated in both external cohorts (Mayo Clinic and University of Nottingham). Expression values are scaled. Descriptions are from gene ontology or KEGG pathways except where referenced. Genes and samples (within bowel habit subtypes) are clustered by Euclidean distance. Sample clustering is based on top 500 probes after ordering by variance across all 30 samples.

IBS-C, IBS with constipation; IBS-D, IBS with diarrhea; HC, healthy control; cAMP/cGMP, cyclic adenosine/guanosine monophosphate; GTP, guanosine triphosphate; GTPase, enzymatic function catalyzing formation of GDP from GTP; BMP, bone-morphogenic protein; SNP, single nucleotide polymorphism; MHC, major histocompatibility complex; ATP, adenosine triphosphate; GPCR, G protein-coupled receptor; PKA, protein kinase A.

Table 3: Genes differentially expressed in IBS-C vs. HC concordant in Mayo

Name (Symbol)	AveExpr	FC (FDR)
<i>Upregulated</i>		
Autophagy related 7 (ATG7)	10.99	1.2 (0.031)
Centriole, cilia and spindle associated protein (CCSAP)	8.90	1.4 (0.04)
Centromere protein J (CENPJ)*	7.63	1.4 (0.032)
Cofilin 2 (CFL2)	10.05	1.5 (0.015)
Family with sequence similarity 161 member A (FAM161A)	7.47	1.4 (0.022)
FERM, ARH/RhoGEF and pleckstrin domain protein 1 (FARP1)*	8.97	1.3 (0.026)
Folliculin (FLCN)*	9.81	1.3 (0.041)
IKBKB interacting protein (IKBIP)	8.38	1.5 (0.035)
Interleukin 17 receptor D (IL17RD)	8.26	1.4 (0.029)
IQ motif containing GTPase activating protein 2 (IQGAP2)*	12.94	1.3 (0.048)
Kelch repeat and BTB domain containing 3 (KBTBD3)	8.24	1.3 (0.025)
MAGE family member D4B (MAGED4/MAGED4B)	9.07	1.4 (0.024)
Mitochondrial calcium uptake 1 (MICU1)	8.00	1.3 (0.046)
NFU1 iron-sulfur cluster scaffold (NFU1)	9.91	1.4 (0.032)
N-acetylneuraminase pyruvate lyase (NPL)*	9.61	1.7 (0.022)
PostGPI attachment to proteins 1 (PGAP1)	7.54	1.6 (0.031)
Reticulon 3 (RTN3)	8.57	1.3 (0.049)
SRY-box 5 (SOX5)	8.02	1.4 (0.03)
Sperm flagellar 2 (SPEF2)	7.31	1.6 (0.033)
TIA1 cytotoxic granule associated RNA binding protein like 1 (TIAL1)*	9.36	1.3 (0.027)
Transmembrane protein 44 (TMEM44)	7.44	1.5 (0.003)
WD repeat domain 19 (WDR19)	8.66	1.3 (0.047)
Zinc finger protein, FOG family member 2 (ZFPM2)*	8.43	1.7 (0.01)
Zinc finger protein 789 (ZNF789)*	7.44	1.3 (0.032)
<i>Downregulated</i>		
Acyl-CoA thioesterase 7 (ACOT7)*	8.58	0.6 (0.019)
Angiotensin II receptor type 2 (AGTR2)	7.35	0.6 (0.019)
Amidohydrolase domain containing 2 (AMDHD2)	9.54	0.8 (0.028)
Chromosome 20 open reading frame 141 (c20orf141)	11.70	0.8 (0.029)
Coiled-coil domain containing 27 (CCDC27)	7.81	0.8 (0.046)
Corticotropin releasing hormone receptor 1 (CRHR1)	7.53	0.7 (0.018)
DExH-box helicase 30 (DHX30)	8.16	0.8 (0.012)
Epsin 1 (EPN1)*	8.23	0.7 (0.041)
Family with sequence similarity 228 member A (FAM228A)	7.90	0.8 (0.038)
F-box and leucine rich repeat protein 18 (FBXL18)*	7.55	0.7 (0.013)
F-box protein 40 (FBXO40)	7.81	0.4 (0.032)
Fibronectin type III domain containing 11 (FNDC11/C20orf195)	9.56	0.8 (0.03)
GDNF family receptor alpha 4 (GFRA4)	11.93	0.7 (0.045)
Interleukin 17 receptor E (IL17RE)*	10.46	0.7 (0.025)
Inositol-3-phosphate synthase 1 (ISYNA1)*	8.84	0.8 (0.01)
Lipoxygenase homology domains 1 (LOXHD1)	9.48	0.7 (0.021)
Leucine rich repeat containing 27 (LRRC27)	8.87	0.8 (0.029)
Mitogen-activated protein kinase kinase 7 (MAP2K7)	11.66	0.7 (0.036)
Myosin light chain kinase 2 (MYLK2)	14.59	0.5 (0.008)
Nectin cell adhesion molecule 1 (NECTIN1)	8.27	0.6 (0.01)
Nectin cell adhesion molecule 4 (NECTIN4)	10.16	0.7 (0.029)
Neuropeptides B and W receptor 1 (NPBWR1)*	7.73	0.6 (0.017)
NUT family member 2D (NUTM2D)	9.59	0.8 (0.046)
Pregnancy specific beta-1-glycoprotein 6 (PSG6)	12.64	0.6 (0.01)
Purine rich element binding protein A (PURA)*	12.90	0.7 (0.026)
RNA exonuclease 1 homolog (REXO1)	9.43	0.7 (0.039)
SCO-spondin (SSPO)	9.15	0.8 (0.029)
Synaptogyrin 2 (SYNGR2)	11.12	0.8 (0.034)
TAP binding protein (TAPBP)*	9.20	0.7 (0.013)
Chromosome 14 open reading frame 80 (TEDC1)	13.37	0.5 (0.014)
Chromosome 16 open reading frame 59 (TEDC2)*	10.55	0.5 (0.008)
T-SNARE domain containing 1 (TSNARE1)	8.89	0.7 (0.008)
Ubiquitin associated protein 1 like (UBAP1L)	8.95	0.8 (0.041)

AveExpr, average expression (log2 transformed); FC, fold change, FDR, Benjamini-Hochberg adjusted *P* value. *Also differentially expressed in this direction in Nottingham.

To address a possible limitation of comparing rectal (Nottingham) expression with sigmoid, we focused further investigation on the 57 common genes between our study and the Mayo study (Figure 1B, Table 3). Up-regulated genes included reticulon 3 which is important in neurite outgrowth.(29) Down-regulated genes included corticotropin releasing factor receptor 1 (CRHR1), SCO-Spondin (SSPO) which we previously found to be hyper-methylated in peripheral blood mononuclear cells from IBS vs. HCs,(82) and the glial cell line-derived neurotrophic factor receptor, GRFA4.

Identification of IBS gene networks by performing WGCNA analysis

WGCNA yielded 45 gene modules of which 26 significantly correlated with IBS status ($p < 0.05$, Figure 2). None of the modules were significantly associated with sex. One module, subsequently named “cAMP-PKA”, was associated with IBS, IBS-C, overall symptom and bloating severity. Top GO terms associated with this module are shown in Table 4 and the other IBS-associated modules are described in Table 5. Results of the pathway analysis for the “cAMP-PKA” module are shown in Figure 3, which is displayed as a subset of two networks. The IPA analysis highlights up-regulation of cAMP/protein kinase A signaling. Interestingly, following review of the “hub” genes, those with the highest “module membership” (correlations with module “eigengene”), revealed a role for cAMP signaling in other modules that were associated with IBS, and the function of the gene products were consistent with increased cAMP (Figure 4, Table 6). For example, PDE9A, which hydrolyzes cAMP, is in a module that is down-regulated in IBS-C. IPA Pathway analysis settings, main results and references for the relationships in the network are provided as a supplementary file.

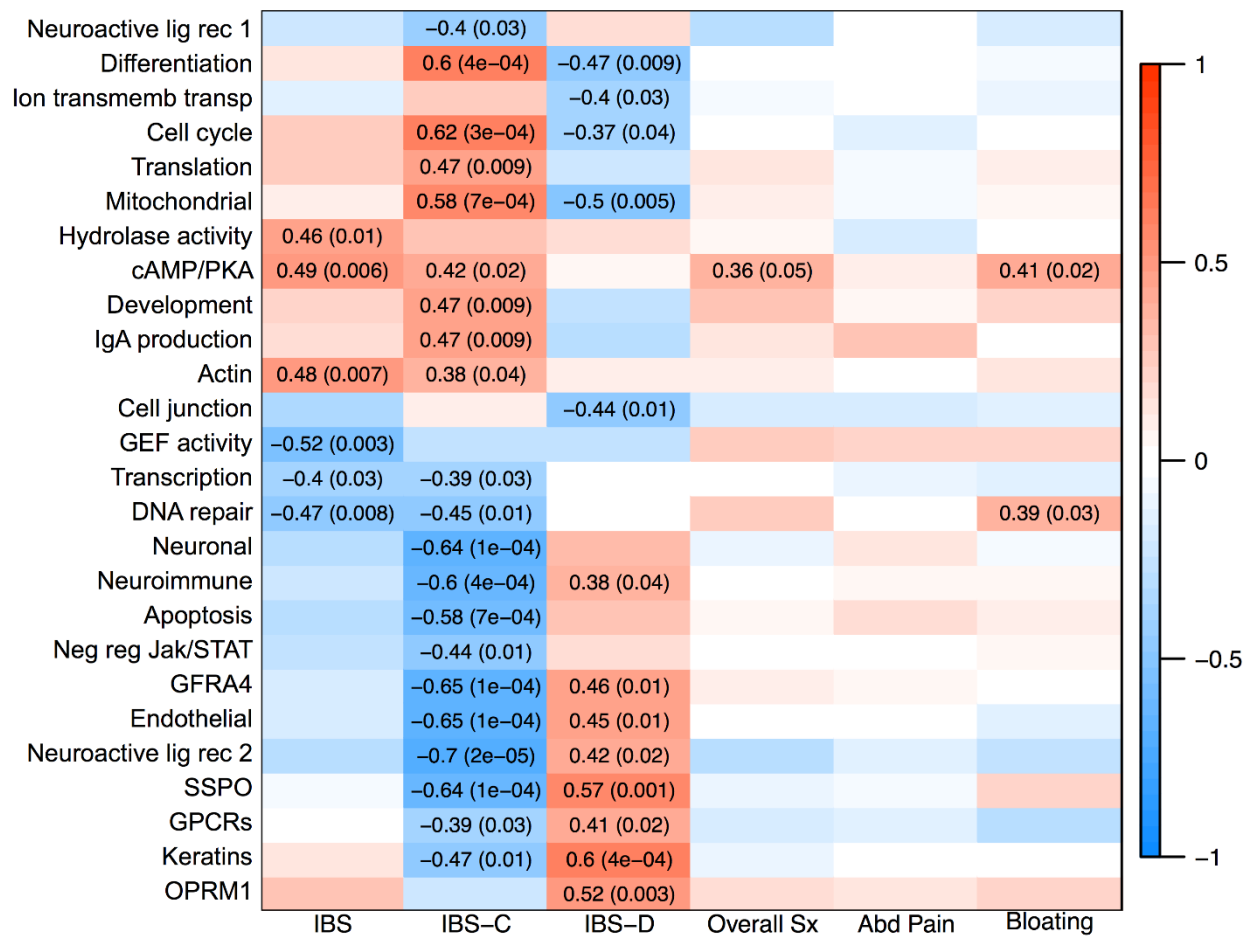


Figure 2: WGCNA modules associated with IBS

Values are correlations between module eigengene and IBS status or overall symptom severity (0-20) with associated p value. Red and blue indicate positive and negative correlations, respectively

Table 4: Gene ontology terms overrepresented among genes in the cAMP/PKA module

GO ID	P Value	Term	Genes
0035589	0.001	G-protein coupled purinergic nucleotide receptor signaling pathway	<i>PTAFR, P2RY14</i>
0051058	0.002	Negative regulation of small GTPase mediated signal transduction	<i>ADRA1A, MYOC, RASAL2</i>
1902692	0.004	Regulation of neuroblast proliferation	<i>CX3CRI, SIX3</i>
0045907	0.006	Positive regulation of vasoconstriction	<i>ADRA1A, PTAFR</i>
0048246	0.008	Macrophage chemotaxis	<i>AZU1, CX3CRI</i>
0071560	0.010	Cellular response to transforming growth factor beta stimulus	<i>AMHR2, LDLRAD4, CDK9, CX3CRI, SOX5, HIPK2, CHST11</i>
0042063	0.012	Gliogenesis	<i>AZU1, ERCC2, ETV5, MBD1, MYOC, SOX5</i>
0046578	0.013	Regulation of Ras protein signal transduction	<i>ADRA1A, MYOC, RASAL2</i>
0050808	0.016	Synapse organization	<i>ACHE, ETV5, SNGG, PDLIM5</i>
0061351	0.018	Neural precursor cell proliferation	<i>CX3CRI, SIX3, SOX5, TACC2</i>
0030182	0.020	Neuron differentiation	<i>CNGA3, ERCC2, ETV5, FES, MBD1, MYOC, ROBO1, SIX3, SOX5, SPTAN1, LAMTOR3, CACNA1G, VAPA, RASAL2, PDLIM5, NSMF, HIPK2, MDGA2</i>
1903307	0.020	Positive regulation of regulated secretory pathway	<i>PTAFR, CACNA1G</i>
0051591	0.022	Response to cAMP	<i>CNGA3, CREM, HNF4A, PTAFR, PYGM, SLC26A6</i>
0007155	0.024	Cell adhesion	<i>ACHE, AZU1, DDR1, CX3CRI, FES, MYOC, PTAFR, ROBO1, ADAM12, EFS, FBLIM1, PCDHA11, ANTXR1</i>
0007600	0.030	Sensory perception	<i>AOC2, CNGA3, PTAFR, SCNN1D, SIX3, OR10H2, CNGB3</i>
0019048	0.030	Modulation by virus of host morphology or physiology	<i>VAPA, HIPK2</i>
0034220	0.030	Ion transmembrane transport	<i>AQP5, CNGA3, PTAFR, S100A1, SCNN1D, VDAC2, CACNA1G, ATP5J2, CNGB3, SLC26A6, SLC26A9</i>
0099536	0.031	Synaptic signaling	<i>ACHE, ADRA1A, CX3CRI, ETV5, SNGG, CACNA1G, NSMF, PANX2, SHISA7</i>
0060395	0.042	SMAD protein signal transduction	<i>HNF4A, HIPK2</i>

Overrepresented terms were determined using the hypergeometric function in the GOstats package for R (44).

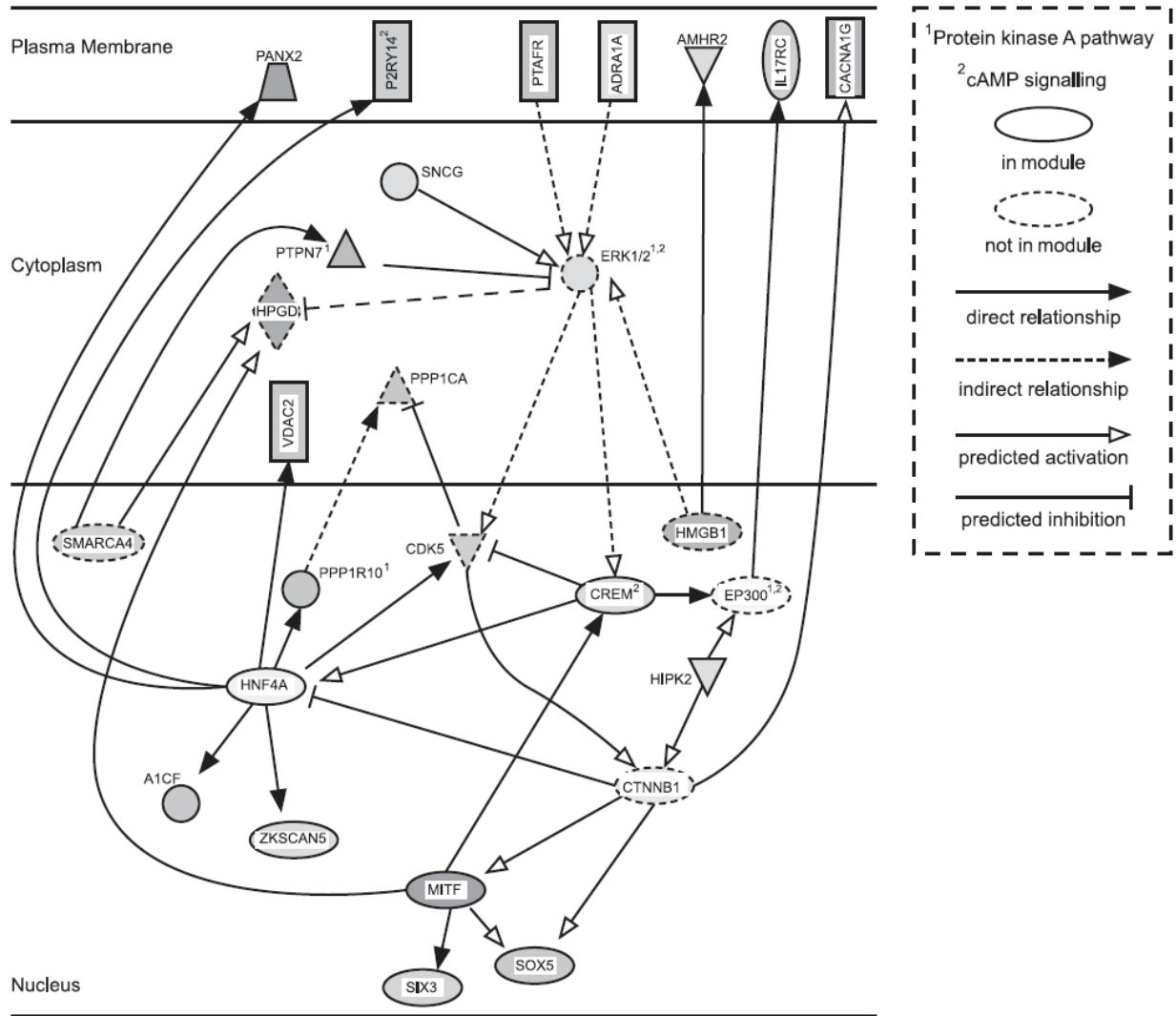


Figure 3: IPA pathway analysis network for the cAMP/PKA module

Genes not in the cAMP/PKA module have shapes outlined in dashed lines. Fill is proportional to fold change from the microarray data (darker = greater fold change) and all are up-regulated except for P2RY14. Figure generated using IPA (Qiagen, Valencia, CA).

Table 5: Description of IBS-associated modules

Module (size)	Description
Neuroactive ligand receptors 1 (758)	Hub genes enriched (0.01) for KEGG pathway “Neuroactive ligand-receptor interaction” and include 5-hydroxytryptamine receptor 4 (<i>HTR4</i>), cholinergic receptor muscarinic 1 (<i>CHRM1</i>), dopamine receptor D4 (<i>DRD4</i>), and glutamate metabotropic receptor 2 (<i>GRM2</i>). Full module also enriched for GO-BP “negative regulation of cytokine production” (0.002), and “negative regulation of tumor necrosis factor production” (0.0005)
Differentiation (435)	Hub genes enriched (FDR <0.05) for cell differentiation and nervous system development. Representative genes for the latter included glial cell derived neurotrophic factor (<i>GDNF</i>) and beta tubulin (<i>TUBB2B</i>). The module overall was enriched for RNA transcription (0.005).
Ion transmembrane transport (248)	Enriched for GO:BP “regulation of transmembrane ion transport” (0.005)
Cell cycle (273)	Very heavily enriched for cell-cycle genes (GO:BP, FDR = 1.5e-49)
Translation (2685)	Highly enriched for terms related to protein translation. Top two are GO:BP “RNA splicing via transesterification reactions” and “ribonucleoprotein complex biogenesis” (1.8e-07 for each)
Mitochondrial (1915)	Module and hub genes heavily enriched for terms and pathways related to cellular respiration and oxidative phosphorylation (FDR <1e-15 for several)
Hydrolase activity (165)	No terms or pathways enriched $P < 0.005$ or FDR <0.05. Nine of 45 hub genes map to GO:BP “regulation of hydrolase activity” (0.02)
cAMP/PKA (184)	Based on Ingenuity Pathway Analysis.
Development (234)	Hub genes marginally (0.006) enriched for “anatomical structure development” and “single organism developmental process.” Relevant genes include the transcriptional regulators GATA binding protein 2 (<i>GATA2</i>), Sp3 transcription factor (<i>SP3</i>) and homeobox B3 (<i>HOXB3</i>).
IgA production (127)	Heavily enriched for GO terms related to adaptive immune response (<0.005) as well as KEGG pathway “Intestinal immune network for IgA production” (FDR <0.0005). Hub genes include three HLA class molecules (DMA, DRA, DMB) integrin subunit alpha and beta (<i>ITGA4</i> , <i>ITGB7</i>).
Actin (77)	Enriched (<0.005) for several terms related to regulation of actin polymerization.
Cell junction (889)	Enriched (FDR <0.05) for GO:MF “cadherin-binding in cell adhesion” and CC “adherens junction.” Relevant genes include vinculin (<i>VCL</i>), tight junction protein 1 (<i>TJP1</i>) and catenin delta 1 (<i>CTNND1</i>)
GEF activity (224)	Hub genes enriched for GO:BP “guanyl-nucleotide exchange factor activity” (<0.005).
Transcription (143)	GO:BP “DNA-templated transcription” is most representative (19 genes, 0.04)
DNA repair (168)	Enriched for several terms related to DNA repair (i.e., GO:BP “double-strand break repair via nonhomologous end joining”, 0.001)
Neuronal (578)	GO:BP: “protein localization to synapse” is most enriched (0.0001) and there are 25 genes mapping to GO:BP “nervous system development” (0.03). Relevant genes include neuronal cell adhesion molecule (<i>NRCAM</i>) and neuropeptide Y (<i>NPY</i>).
Neuroimmune (1074)	Hub genes are enriched for sensory perception (FDR <0.05) and include 6 olfactory receptors. Hub genes also contain multiple genes related to the inflammatory response and innate immune system. These include tumor necrosis factor (<i>TNF</i>), <i>CD70</i> , <i>CD80</i> , macrophage expressed 1/Perforin-2 (<i>MPEG1</i>), toll like receptor 9 (<i>TLR9</i>), mucin 22 (<i>MUC22</i>), defensin beta 113 (<i>DEFB113</i>), cysteine rich secretory protein 3 (<i>CRISP3</i>), and the serine protease inhibitors serpin family E members 1&2 (<i>SERPINE1/2</i>).
Apoptosis (275)	Enriched for GO:BP “positive regulation of apoptotic process” (<0.005). Relevant genes include caspase 2 (<i>CASP2</i>) and slit guidance ligand 2 (<i>SLIT2</i>).
Negative regulation of JAK/STAT (139)	The top GO:BP term is “negative regulation of JAK-STAT cascade (0.003) based on caveolin 1 (<i>CAV1</i>) and inositol polyphosphate-5-phosphatase F (<i>INPP5F</i>).
GFRA4 (161)	No significantly enriched terms or pathways. GDNF family receptor alpha is one of top genes (MM = 0.93). Nonsignificantly enriched terms include GO:BP “regulation of GTPase activity” (0.05), “cellular response to BMP stimulus” (0.02), “glial cell-derived neurotrophic factor receptor signaling pathway” (0.04), and “catechol-containing compound metabolic process” (0.01).
Endothelial (107)	Enriched (0.001) for GO:BP “regulation of tube size” based on expression of endothelial nitric oxide synthase (eNOS, <i>NOS3</i>) and adrenoceptor alpha 1A (<i>ADRA1A</i>).
Neuroactive ligand receptors 2 (892)	Hub genes enriched for membrane proteins (GO:CC, FDR <0.05). Top 100 genes include receptors for NPY (<i>NPY1R</i>) and neuropeptides B and W (<i>NPBWRI</i>), melanocortin (<i>MC5R</i>) and GABA (<i>GABRG1</i>).
SSPO (186)	SCO-Spondin (<i>SSPO</i>) is among top two genes in module (MM = 0.95). GO:BP “ion transport” is also enriched (<0.005)
GPCRs (328)	Enriched (<0.005) for GO:BP “G-protein coupled receptor signaling pathway”. Relevant genes include 5-hydroxytryptamine receptors 6 and 3D (<i>HTR6</i> , <i>HTR3D</i>), urocortin 2 (<i>UCN2</i>), and four olfactory receptors.
Keratins (827)	Hub genes enriched for GO:CC “Intermediate filament cytoskeleton” (FDR = 0.05). The majority of these are keratins or keratin associated proteins. The module overall (but not hub genes exclusively) is enriched for GO:BP “response to steroid hormone” (0.009). Relevant genes include hepatocyte nuclear factor 4 alpha (<i>HNF4A</i>), KRAS proto-oncogene, GTPase (<i>KRAS</i>) and estrogen related receptor alpha (<i>ESRRA</i>).
OPRM1 (198)	Not significantly enriched for pathways or terms. Opioid receptor mu 1 (<i>OPRM1</i>) is a hub gene (MM = 0.84). Marginally enriched for GO:BP “sensory response to pain” ($P = 0.02$) due to <i>OPRM1</i> and purinergic receptor P2X 4 (<i>P2RX4</i>).

Module membership (MM) is the correlation with module eigengene. Hub genes are those with $MM > 0.8$. MM P values < 1e-7 for all hub genes. Overrepresented terms were determined using the hypergeometric function in the GOstats package for R(44) or DAVID. (55, 56). Numbers in parentheses are unadjusted P values except where noted. GO, gene ontology; BP, biological process; MF, molecular function; CC, cellular component; cAMP, cyclic adenosine monophosphate; PKA, protein kinase A; GEF, guanine nucleotide exchange factor; GPCRs, G-protein coupled receptors.

Table 6: Correlation of cAMP-related genes with module eigengenes

Gene	MM*	P	Module
<i>Upregulated modules</i>			
<i>ADRA1A</i>	0.711	1.1E-05	cAMP/PKA
<i>CREB3L1</i>	-0.682	3.3E-05	Hydrolase activity
<i>GABRB3</i>	0.714	9.6E-06	Actin
<i>GEM</i>	0.772	5.9E-07	Actin
<i>PANX2</i>	0.830	1.4E-08	cAMP/PKA
<i>P2RY14</i>	-0.660	7.3E-05	cAMP/PKA
<i>RGS20</i>	0.500	4.9E-03	Actin
<i>Downregulated modules</i>			
<i>ARHGEF15</i>	0.564	1.2E-03	GEF activity
<i>CACNA1G**</i>	0.934	4.6E-14	GEF activity
<i>CACNB4</i>	0.776	4.6E-07	GEF activity
<i>CALHM2</i>	0.818	3.3E-08	Transcription
<i>KCNJ11</i>	0.928	1.6E-13	Transcription
<i>KCNMB3</i>	0.896	2.3E-11	DNA repair
<i>PDE9A</i>	0.669	5.2E-05	DNA repair
<i>RAPGEF6</i>	0.918	8.9E-13	GEF activity

*MM (module membership) is the correlation between the gene and the module eigengene; **a separate isoform from that in the orangered4 module.

Modules up-regulated in IBS

Modules down-regulated in IBS

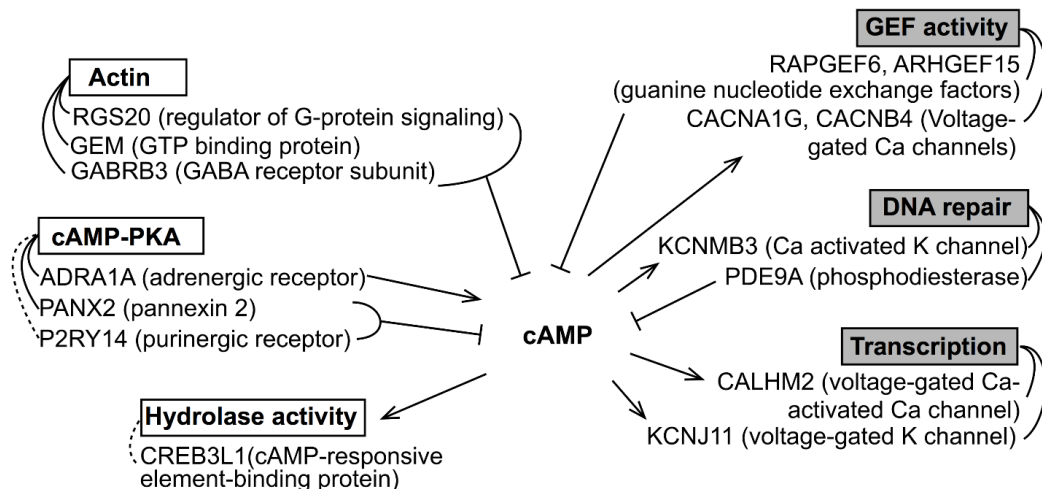


Figure 4: cAMP signaling roles for hub genes in IBS-associated modules

Directionality of arrows indicates upstream/downstream of cAMP. Dotted lines indicate a negative correlation with the module eigengene. Note that CACNA1G in the GEF activity module is a different isoform than in the cAMP/PKA (Figure 3). The module memberships and p values are shown in Table 6.

cAMP, cyclic adenosine monophosphate; *PKA*, protein kinase A; *GEF*, guanine nucleotide exchange factor

Identification of IBS signaling pathways based on the gene expression data

GSEA results are provided in a supplementary file. There were no gene sets enriched in IBS-C or HCs meeting a threshold of $FDR < 0.25$ (recommended GSEA threshold). Most gene sets that were enriched among HCs with nominal (unadjusted) $p < 0.05$ were thematically related to synaptic transmission or G-protein coupled receptor (GPCR) signaling. These included Reactome pathways “Activation of kainite receptors upon glutamate binding” ($p=0.024$), “G alpha(s) signaling” ($p=0.026$), “GPCR downstream signaling” ($p=0.036$), KEGG pathway “neuroactive ligand receptor interaction” ($p=0.014$), and GO:BP gene-sets “neuronal synaptic plasticity” (0.004), and “regulation of synaptic transmission glutaminergic” ($p=0.021$). Despite the absence of DETs in IBS vs. HCs, the Reactome pathway “signaling by FGFR1 mutants” was enriched in HCs vs. IBS (p/FDR 0.012/0.19). The 20 genes in this gene set include growth factors and downstream effectors (Jak/STAT pathway). The top enriched Reactome gene set was smooth muscle contraction (p/FDR 0.004/0.29); however, these genes (actins, myosin kinases) likely represent expression at cell-junctions rather than muscle. No gene sets were enriched at $FDR < 0.25$ in IBS-D vs. HC, IBS-C vs. IBS-D or in men vs. women. Within IBS, the following KEGG gene sets were enriched in men vs. women (p/FDR): “allograft rejection” (0.002/0.12), “graft versus host disease” (0.006/.086), and “autoimmune thyroid disease” (0.002/0.139). These gene sets overlap and include (human leukocyte antigen) HLA complex proteins and tumor necrosis factor (TNF).

Validation of selected targets

Selection of genes to validate ($n=67$) was based on the following: a) visualization of dot plots to ensure consistent group difference, b) programmatic search (annotate package in R) of abstracts linked to each Entrez ID for the words “pain” or “gastrointestinal” and manual search of retrieved abstracts, c) comparison with external datasets, d) comparison with results of IBS GWA studies (*HTR3A*, *CDC42EP5*(125)), and e) results of other studies.

Differential Expression

In the expanded cohort (15 IBS-C vs. 14 HC), differential expression of 9/67 DETs was validated with nCounter (FDR<0.05, same direction of FC as in microarray, Figure 5, Table 7). There were an additional six (*ALDH3A1*, *ELMOD3*, *GH1*, *RAMP1*, *SSPO*, *TMEM80*) with unadjusted p value<0.05 and five (*ALDAH9A1*, *CLDN15*, *EZH2*, *RNF19A*, *YIF1B*) with significant positive correlations (p<0.05) between microarray and nCounter values. Associations between gene expression and IBS symptoms were found though these did not retain significance after adjusting for multiple comparisons (n=67 genes).

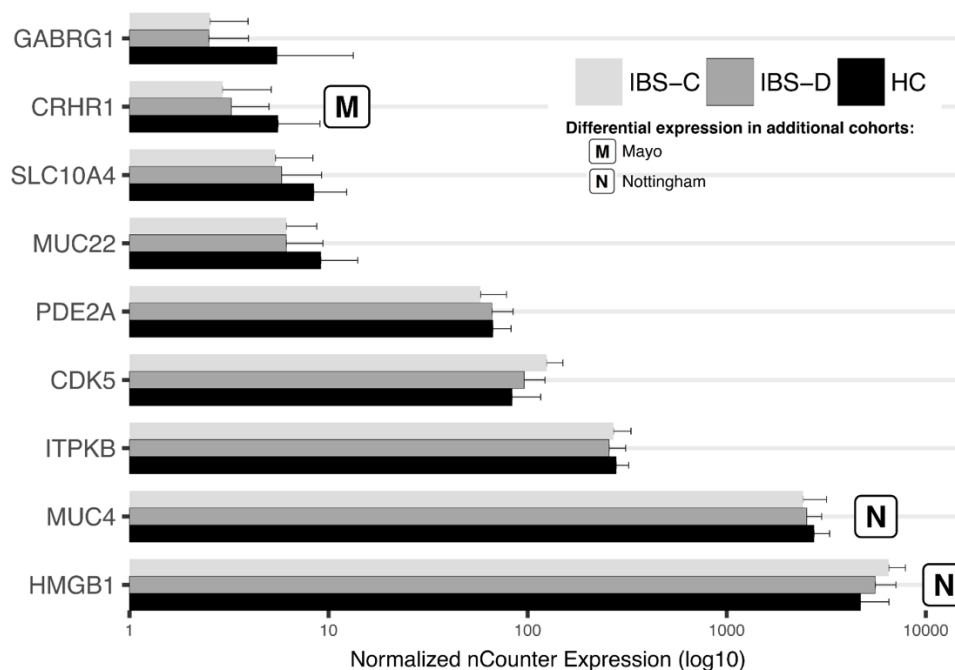


Figure 5: nCounter expression of 9 validated genes in IBS-C vs. HC

Plotted results are mean (SD) for nCounter expression (n=45). Short description of genes not described in paper: ITPKB is expressed in cells with a brush border,(104) but is also important in immune cells(92) and neurons.(110) Its molecular function is related to intracellular calcium regulation.(113) SLC10A4 is a bile acid transporter, about which relatively little is known including whether it is able to transport bile acids or neurotransmitters.(1, 103, 109) Slc10a4 knockout mice show altered dopamine homeostasis.(87, 128). MUC4 is a major component of intestinal mucus. Roles include barrier function and host-microbe interactions. Expression of MUC22, a susceptibility locus for panbronchiolitis and childhood asthma, is up-regulated in bronchial epithelial cells after exposure to a virus-like molecule or lipopolysaccharide.(31, 52)

Table 7: Results of nCounter validation

Gene	AveExpr (all $n = 45$)	IBS-C vs. HC	IBS-D vs. HC	IBS vs. HC
		FC (FDR)	FC (FDR)	FC (FDR)
<i>AK9</i>	11.7	1.1 (0.74)	1.1 (0.63)	1.1 (0.61)
<i>ALDH3A1</i>	11.1	0.7 (0.07)*	0.8 (0.3)	0.7 (0.06)*
<i>ALDH9A1</i>	14.8	1.1 (0.5)	1.0 (0.79)	1.0 (0.86)
<i>AQP9</i>	8.5	0.6 (0.17)	0.5 (0.21)	0.6 (0.06)*
<i>CCDC134</i>	11.9	1.0 (0.76)	0.9 (0.54)	0.9 (0.48)
<i>CDC42EP5</i>	16.7	1.1 (0.3)	1.0 (0.67)	1.1 (0.44)
<i>CDK5</i>	11.2	1.3 (0.05)	1.1 (0.57)	1.2 (0.09)*
<i>CLDN15</i>	12.5	1.0 (0.91)	0.9 (0.61)	1 (0.85)
<i>CRHR1</i>	6.4	0.5 (0.05)	0.6 (0.41)	0.5 (0.06)*
<i>DBI</i>	16.2	1.1 (0.33)	1.0 (1.0)	1.0 (0.59)
<i>DHRS12</i>	12.4	0.9 (0.48)	0.9 (0.61)	0.9 (0.48)
<i>EFR3B</i>	9.6	1.0 (0.87)	0.9 (0.79)	1.0 (0.93)
<i>ELMOD3</i>	12.3	0.8 (0.11)*	0.9 (0.61)	0.9 (0.18)
<i>EZH2</i>	13.0	1.0 (0.74)	1.0 (0.83)	1.0 (0.94)
<i>FABP3</i>	11.5	1.0 (0.94)	1.0 (0.96)	1.0 (0.97)
<i>GABBR1</i>	12.8	0.6 (0.01)	0.8 (0.54)	0.7 (0.04)
<i>GABRA1</i>	8.6	0.9 (0.76)	1.0 (1.0)	1.0 (0.88)
<i>GABRG1</i>	6.2	0.5 (0.05)	0.6 (0.21)	0.6 (0.04)
<i>GH1</i>	8.0	0.7 (0.12)*	0.6 (0.21)*	0.7 (0.06)*
<i>HMGAI</i>	14.3	1.0 (0.87)	1.1 (0.57)	1.1 (0.59)
<i>HMGB1</i>	17.0	1.2 (0.05)	1.1 (0.3)	1.2 (0.06)*
<i>HTR3A</i>	7.9	0.8 (0.33)	0.7 (0.21)	0.7 (0.14)*
<i>IL17RD</i>	9.9	0.8 (0.12)	0.9 (0.54)	0.8 (0.16)
<i>IL17RE</i>	14.9	0.9 (0.29)	1.0 (0.82)	0.9 (0.48)
<i>ITPKB</i>	12.6	0.8 (0.05)	0.8 (0.21)*	0.8 (0.04)
<i>JRK</i>	11.4	0.9 (0.63)	0.9 (0.61)	0.9 (0.48)
<i>KAT5</i>	14.2	1.1 (0.59)	1.1 (0.61)	1.1 (0.48)
<i>MLN</i>	6.5	1.1 (0.87)	0.7 (0.45)	0.9 (0.61)
<i>MUC22</i>	7.3	0.6 (0.05)	0.7 (0.21)	0.6 (0.04)
<i>MUC4</i>	15.9	0.7 (0.02)	0.8 (0.41)	0.8 (0.04)
<i>MUCL1</i>	6.4	0.8 (0.62)	0.8 (0.61)	0.8 (0.48)
<i>NECTIN1</i>	13.3	1.0 (0.76)	1.0 (0.83)	1.0 (0.97)
<i>NKX2-1</i>	7.1	0.6 (0.08)	0.6 (0.21)	0.6 (0.04)
<i>NKX3-1</i>	8.1	0.9 (0.48)	0.9 (0.61)	0.9 (0.48)
<i>NOTCH4</i>	10.8	0.9 (0.64)	1.0 (0.99)	1.0 (0.85)
<i>NPBWR1</i>	7.5	1.1 (0.74)	0.9 (0.82)	1.0 (0.97)
<i>NRG3</i>	8.3	0.9 (0.74)	1.0 (0.82)	0.9 (0.77)
<i>OPHN1</i>	12.5	1.1 (0.51)	1.1 (0.61)	1.1 (0.48)
<i>OPLAH</i>	12.2	0.7 (0.11)	0.8 (0.3)	0.8 (0.09)
<i>OPRM1</i>	6.7	0.8 (0.62)	0.7 (0.59)	0.8 (0.48)
<i>PALLD</i>	15.3	1.1 (0.62)	1.1 (0.61)	1.1 (0.51)
<i>PCDHA3</i>	8.5	0.6 (0.02)	0.8 (0.29)	0.7 (0.04)
<i>PDE11A</i>	9.5	0.8 (0.32)	0.9 (0.61)	0.9 (0.36)
<i>PDE2A</i>	10.5	0.7 (0.02)	0.9 (0.61)	0.8 (0.06)*
<i>PPP1CA</i>	17.0	1.1 (0.18)	1.2 (0.3)	1.2 (0.13)*
<i>PPP1R1A</i>	8.9	0.7 (0.16)	0.8 (0.54)	0.8 (0.18)
<i>RAMP1</i>	9.4	0.7 (0.09)*	0.7 (0.21)*	0.7 (0.04)
<i>RBM5</i>	14.9	0.9 (0.33)	1.0 (0.61)	0.9 (0.44)
<i>RNF19A</i>	14.9	1.0 (0.76)	1.0 (0.99)	1.0 (0.86)
<i>SI00A1</i>	7.7	0.8 (0.37)	1.2 (0.61)	1.0 (0.86)
<i>SGSM2</i>	13.0	0.8 (0.05)	0.9 (0.62)	0.8 (0.15)
<i>SIGMAR1</i>	13.5	1.0 (0.76)	1.0 (0.82)	1.0 (0.78)
<i>SLC10A4</i>	7.2	0.5 (0.02)	0.6 (0.21)	0.6 (0.03)
<i>SLC17A3</i>	8.4	1.0 (0.91)	0.9 (0.61)	1.0 (0.85)
<i>SMARCA4</i>	14.7	1.1 (0.24)	1.1 (0.54)	1.1 (0.22)
<i>SOX5</i>	9.9	0.7 (0.01)	0.8 (0.21)	0.7 (0.03)
<i>SPATS2</i>	12.6	1.2 (0.16)	1.1 (0.57)	1.1 (0.21)
<i>SSPO</i>	6.2	0.5 (0.08)*	0.5 (0.21)	0.5 (0.04)
<i>ST3GAL3</i>	12.2	0.9 (0.32)	0.9 (0.56)	0.9 (0.29)
<i>THBD</i>	11.7	0.7 (0.12)	0.8 (0.54)	0.8 (0.15)
<i>TMEM44</i>	13.7	0.9 (0.76)	0.9 (0.72)	0.9 (0.73)
<i>TMEM80</i>	13.6	1.2 (0.12)*	1.0 (0.98)	1.1 (0.44)
<i>TRIM37</i>	13.4	1.1 (0.51)	1.1 (0.61)	1.1 (0.48)
<i>TXLNG</i>	13.3	0.9 (0.65)	0.9 (0.61)	0.9 (0.53)
<i>UGT1A9</i>	15.2	1.0 (0.78)	1.0 (0.98)	1.0 (0.87)
<i>USP34</i>	15.1	0.9 (0.09)	1.0 (0.61)	0.9 (0.2)
<i>YIF1B</i>	12.9	1.1 (0.33)	0.9 (0.61)	1.0 (0.87)

Boldface indicates FDR < 0.05. Unmarked values < 0.05 indicate direction did not match microarray. AveExpr, average expression (log2 transformed); FC, fold change, FDR, Benjamini-Hochberg adjusted *P* value. *Unadjusted *P* value < 0.05.

DISCUSSION

The most important findings from this study are 1) sigmoid mucosal gene expression profiles differentiate IBS-C from HCs and 2) results of multiple analyses and in several cohorts support the relevance of neurally-mediated mechanisms in IBS overall and specifically in IBS-C.

1) Defining a molecular signature for IBS-C

Very few studies have reported DETs in IBS-C vs. HCs. A survey of the literature revealed ~90 DETs in IBS (those included in our microarray are shown in Table 8). Recently, Peters et al. compared expression of tight junction proteins in sigmoid colonic mucosal biopsies from women with IBS-C vs. HC women and none were reported as differentially expressed.(97) As a demonstration of the heterogeneity in findings with different samples and methodology, some of the genes in this study were DETs in our study (FDR<0.05, Table 9). An important strength of our study is that our microarray sample included no individuals using SSRIs and there were only two with antidepressant use in the validation sample. Medication use was not reported in the Nottingham comparison study, but in the Mayo study, 42% of IBS participants took medications (SSRI, SNRI, TCA, or dopaminergic agent) vs. 1% of HCs. It is possible that this difference in medication use contributed to our findings. Divergent results by bowel habit subtype is supported by other studies,(28, 57) but notably, there is evidence for the absence of subtype-specific effects as well.(38, 93) We caution the reader against over-interpretation of these findings. While gene networks may differ in IBS-C and IBS-D, one cannot conclude that these differences are causative and therefore represent different underlying pathophysiology. These differences also are not evidence of bowel habit differences at the phenotype level (i.e. nerve sensitivity). Furthermore, in accordance with our sample size calculation, one can only conclude that there are unlikely to be genes with greater than 2-fold difference between IBS-D and HC. We are underpowered to make a negative conclusion for smaller differences that could still be biologically relevant.

Table 8: Comparison of findings from current study for genes found to be differentially expressed in the literature

Symbol	Published Change & Subgroup (Ref.)	Fold Change, P Value			Product
		IBS	IBS-C	IBS-D	
<i>ALDOB</i> ***	↑ D (22)	0.9 (0.3)	0.8 (0.06)	1.0 (0.88)	Fructose-bisphosphate aldolase B
<i>ALDOC</i> ***	↑ D (22)	1.3 (0.05)	1.6 (0.01)	1.1 (0.45)	Fructose-bisphosphate aldolase C
<i>AQP8</i>	↓ D (121)	1.3 (0.25)	1.7 (0.06)	1.0 (0.94)	Aquaporin-8
<i>BIRC5</i>	↑ D (22)	1.1 (0.53)	1.5 (0.02)	0.8 (0.16)	Baculoviral IAP repeat-containing protein 5 isof. 3
<i>C4BPA</i> ***	↑ D (21)	0.9 (0.29)	0.7 (0.002)	1.2 (0.13)	Complement component 4 binding protein, alpha
<i>CACNA1H</i> ***	↑ (102)	1.1 (0.57)	1.2 (0.07)	0.9 (0.36)	Voltage-dep. T-type calcium channel sub. Alpha-1h isof. A
<i>CARD16</i> ***	↑ (2)	1.1 (0.44)	1.1 (0.25)	1.0 (0.86)	Caspase recruitment domain-containing protein 16 isof. 1
<i>CASP1</i> ***	↑ (2)	0.9 (0.55)	1.0 (0.7)	0.8 (0.16)	Caspase-1 isof. Gamma
<i>CCL20</i> ***	↑ D (21, 22)	0.8 (0.12)	0.6 (0.004)	1.1 (0.7)	C-C motif chemokine 20 isof.1 precursor
<i>CD163LI</i> ***	↑ (2)	1.2 (0.25)	1.6 (0.03)	1.0 (0.84)	Scavenger recep. Cysteine-rich type 1 protein m160 precursor
<i>CKB</i>	↑ (2)	0.9 (0.33)	0.9 (0.41)	0.9 (0.38)	Creatine kinase B-type
<i>DTL</i>	↑ (2)	0.8 (0.03)	0.9 (0.12)	0.8 (0.03)	Denticleless protein homolog
<i>DUOX2</i> **	↑ (2)	1.3 (0.06)	1.3 (0.05)	1.2 (0.18)	Dual oxidase 2 precursor
<i>DUSP27</i>	↑ D (22)	1.0 (0.95)	0.9 (0.81)	1.0 (0.9)	Dual specificity phosphatase 27 (putative)
<i>ERAP2</i>	↑ (2)	0.9 (0.8)	1.4 (0.5)	0.6 (0.27)	Endoplasmic reticulum aminopeptidase 2
<i>F2RL1</i>	↑ (75)	0.9 (0.63)	0.9 (0.29)	1.0 (0.82)	Proteinase-activated receptor 2 precursor
<i>F2RL3</i> ***	↑ (130)	1.0 (0.58)	1.0 (0.99)	0.9 (0.34)	PAR4 coagulation factor II (thrombin) receptor-like 3
<i>FAAH</i>	↓ C (31)	1.1 (0.09)	1.2 (0.05)	1.1 (0.31)	Fatty-acid amide hydrolase 1
<i>FCGR2A</i>	↑ (2)	0.5 (0.09)	0.7 (0.3)	0.5 (0.05)	Low affinity Ig gamma Fc region recep. II-a isof.1 precursor
<i>FGFR4</i> *	↓ D (21)	0.9 (0.12)	1.0 (0.86)	0.8 (0.01)	Fibroblast growth factor receptor 4 isof. 2 precursor
<i>FN1</i>	↓ D (21, 22)	0.8 (0.07)	0.8 (0.12)	0.8 (0.12)	Fibronectin isof. 5 preproprotein
<i>FOXP3</i>	↑ (13)	0.6 (0.001)	0.5 (0.001)	0.7 (0.02)	Forkhead box P3
<i>GPBAR1</i>	↑ C (21)	0.9 (0.2)	1.0 (0.46)	0.9 (0.14)	G protein-coupled bile acid receptor 1
<i>GPRI</i>	↑ D (98)	1.3 (0.01)	1.5 (0.01)	1.2 (0.11)	G protein-coupled estrogen receptor 1
<i>GUCA2B</i> ***	↑ D (22)	1.2 (0.4)	1.4 (0.25)	1.1 (0.75)	Guanylate cyclase activator 2B (uroguanylin)
<i>HELLS</i> ***	↑ (2)	1.2 (0.4)	1.6 (0.01)	1.1 (0.75)	Helicase, lymphoid-specific
<i>HLA-DRB5</i>	↓ D (22)	1.2 (0.55)	1.0 (0.73)	1.3 (0.49)	MHC, class II, DR beta 5 precursor
<i>HSPA1A</i> ***	↓ D (22)	0.9 (0.48)	0.9 (0.57)	0.9 (0.51)	Heat shock 70 kda protein 1A/1B
<i>HSPA1B</i>	↓ D (22)	1.0 (0.92)	1.0 (0.77)	1.0 (0.9)	Heat shock 70 kda protein 1B
<i>IFIT3</i> **	↓ D (21, 22)	0.9 (0.07)	0.9 (0.19)	0.9 (0.07)	Interferon-induced protein with tetratricopeptide repeats 3
<i>IGFBP2</i> ***	↑ D (22)	1.6 (0.07)	2.1 (0.01)	1.2 (0.61)	Insulin-like growth factor-binding protein 2 precursor
<i>IL10</i> ***	↑ (13, 30)	0.8 (0.28)	0.9 (0.62)	0.7 (0.18)	Interleukin-10 precursor
<i>IL13</i>	↑ (125)	0.9 (0.14)	0.8 (0.02)	1.0 (0.85)	Interleukin-13 precursor
<i>IL1B</i>	↑ (4)	0.8 (0.31)	0.8 (0.31)	0.8 (0.45)	Interleukin-1 beta proprotein
<i>KCNK3</i>	↑ (2)	1.0 (0.81)	0.8 (0.26)	1.1 (0.48)	Potassium voltage-gated channel subfamily S member 3
<i>KDELRL2</i>	↑ (42)	1.1 (0.45)	1.2 (0.1)	1.0 (0.73)	ER lumen protein retaining receptor 2 isof. 1
<i>LCN15</i> ***	↑ D (22)	1.0 (0.9)	0.8 (0.16)	1.2 (0.23)	Lipocalin 15
<i>LYZ</i>	↓ (2)	1.0 (0.96)	1.3 (0.32)	0.8 (0.36)	Lysozyme C precursor
<i>MCM5</i> ***	↑ (2)	1.0 (0.7)	1.2 (0.1)	0.9 (0.32)	DNA replication licensing factor MCM5
<i>MS4A4A</i> **	↑ (2)	1.4 (0.07)	1.8 (0.01)	1.1 (0.71)	Membrane-spanning 4-domains subfam. A member 4a isof. 1
<i>MUC20</i> *	↑ (2)	1.2 (0.04)	1.3 (0.01)	1.1 (0.49)	Mucin-20 isof. S
<i>MX1</i> ***	↓ D (22)	1.0 (0.94)	1.1 (0.76)	0.9 (0.67)	Interferon-induced GTP-binding protein Mx1
<i>NCF4</i> *	↑ (2)	1.3 (0.04)	1.6 (0.001)	1.0 (0.99)	Neutrophil cytosol factor 4 isof. 2
<i>NOS2</i>	↑ (4)	0.9 (0.07)	1.0 (0.48)	0.8 (0.02)	Nitric oxide synthase, inducible
<i>NR1H4</i>	↑ D (21)	1.1 (0.88)	1.9 (0.11)	0.6 (0.18)	Bile acid receptor isof. 2
<i>OCLN</i> **	↑ C (21)	1.2 (0.06)	1.2 (0.06)	1.1 (0.17)	Occludin isof. B precursor
<i>P2RY1</i> ***	↑ (81)	1.1 (0.51)	1.1 (0.37)	1.0 (0.79)	P2Y purinoceptor 1
<i>P2RY2</i>	↑ (81)	1.0 (0.83)	1.0 (0.31)	0.9 (0.17)	P2Y purinoceptor 2
<i>P2RY4</i> ***	↑ D (21, 22)	1.0 (0.89)	0.9 (0.13)	1.1 (0.2)	Pyrimidinergic receptor P2Y, G-protein coupled, 4
<i>PDZD3</i> ***	↑ D (21, 22)	1.1 (0.25)	1.2 (0.25)	1.1 (0.4)	Na(+)/H(+) exchange regulatory cofactor NHE-RF4 isof. 2
<i>PSME2</i> ***	↑ (2)	1.0 (0.62)	1.2 (0.01)	0.9 (0.09)	Proteasome activator complex subunit 2
<i>RBP2</i>	↑ D (22)	1.2 (0.33)	1.4 (0.08)	1.0 (0.89)	Retinol-binding protein 2
<i>RFC4</i> ***	↑ (2)	1.1 (0.27)	1.3 (0.04)	1.0 (0.81)	Replication factor C subunit 4
<i>S100A10</i>	↑ (20, 106)	0.9 (0.11)	0.9 (0.34)	0.8 (0.07)	Protein S100-A10
<i>SLC10A2</i>	↑ D (21)	1.0 (0.86)	1.1 (0.33)	0.9 (0.51)	Ileal sodium/bile acid cotransporter
<i>SLC6A4</i>	↑ (36, 45, 63, 76, 122)	1.9 (0.01)	1.3 (0.38)	2.9 (0.001)	Sodium-dependent serotonin transporter
<i>TAP2</i> ***	↓ (2)	1.1 (0.79)	1.9 (0.05)	0.6 (0.12)	Antigen peptide transporter 2 isof. 2
<i>TFF1</i>	↑ D (22)	1.0 (0.92)	1.1 (0.62)	0.9 (0.76)	Trefoil factor 1 precursor
<i>TH</i>	↑ D (129)	0.9 (0.25)	0.9 (0.37)	0.9 (0.26)	Tyrosine 3-monooxygenase isof. B
<i>TJP1</i> ***	↓ (21, 119)	0.9 (0.21)	0.9 (0.5)	0.9 (0.13)	Tight junction protein ZO-1 isof. A
<i>TLR2</i> ***	↑ (12)	1.0 (0.86)	1.1 (0.64)	1.0 (0.86)	Toll-like receptor 2 precursor
<i>TLR4</i> ***	↑ (12)	1.0 (0.63)	1.0 (0.92)	1.1 (0.36)	Toll-like receptor 4 isof. A precursor
<i>TNFSF15</i>	↓ (21, 112)	1.0 (0.88)	1.0 (0.68)	1.0 (0.88)	TNF ligand superfam. member 15 isof. VEGF-251 precursor
<i>TPH1</i>	↓ (63)	1.3 (0.17)	1.5 (0.12)	1.2 (0.4)	Tryptophan 5-hydroxylase 1
<i>TPSAB1</i>	↑ (4, 75)	0.8 (0.03)	0.9 (0.17)	0.8 (0.02)	Tryptase alpha/beta-1 precursor
<i>TRPV1</i>	↑ (64, 81, 131)	0.8 (0.05)	0.7 (0.08)	0.8 (0.1)	TRP cation channel subfamily V member 1
<i>VIP</i> *	↑ D (21, 22)	0.7 (0.06)	0.6 (0.03)	0.8 (0.25)	VIP peptides isoform 2 preproprotein

Continued

Table 8.—Continued

Symbol	Published Change & Subgroup (Ref.)	Fold Change, <i>P</i> Value			Product
		IBS	IBS-C	IBS-D	
<i>VSIG2</i>	↑ (2)	0.8 (0.07)	0.8 (0.16)	0.8 (0.08)	V-set and Ig domain-containing protein 2 precursor
<i>VSIG4*</i>	↑ (2)	1.4 (0.03)	1.7 (0)	1.1 (0.5)	V-set and Ig domain-containing protein 4 isof. 1 precursor
<i>WDR72***</i>	↓ D (14)	0.8 (0.02)	0.8 (0.01)	0.9 (0.15)	WD repeat-containing protein 72

↑, increased vs. healthy controls (HCs); ↓ decreased vs. HCs; D, IBS with diarrhea; C, IBS with constipation; isof., isoform; dep., dependent; sub., subunit; Ig, immunoglobulin; recep., receptor; subfam., subfamily; MHC, major histocompatibility complex; TNF, tumor necrosis factor; TRP, transient receptor potential. *Our results agree with reported findings at $P < 0.05$, ** $P < 0.1$ and ***not significant but same direction, respectively.

Table 9: Differentially expressed tight junction genes

Symbol	Peters	UCLA	Mayo/Aerssens
	FC (<i>P</i>)	FC (FDR)	FC (<i>P</i> , FDR)
<i>MARK2</i>	1.2 (0.02)	1.4 (0.03)	0.9 (0.49, 0.78)
<i>LLGL2</i>	1.2 (0.03)	1.3 (0.02)	1.0 (0.34, 0.69)
<i>TJAP1</i>	1.2 (0.10)	0.6 (0.01)	1.0 (0.35, 0.69)
<i>PRKCZ</i>	1.2 (0.11)	1.8 (0.01)	1.0 (1.00, 1.00)
<i>CLDN15</i>	1.2 (0.22)	1.4 (0.02)	0.9 (0.01, 0.32)
<i>MLLT4</i>	1.2 (0.36)	1.5 (0.01)	1.1 (0.02, 0.34)
<i>PRKCI</i>	1.1 (0.46)	1.3 (0.04)	1.1 (0.10, 0.47)
<i>CSNK2A1</i>	2.0 (0.50)	1.7 (0.00)	0.9 (0.01, 0.32)
<i>CLDN6</i>	1.1 (0.56)	0.7 (0.01)	0.9 (0.08, 0.45)
<i>SPTA1</i>	0.9 (0.69)	1.4 (0.03)	1.0 (0.20, 0.58)
<i>CSDA</i>	1.0 (0.75)	0.7 (0.02)	0.8 (0.00, 0.30)
<i>MPDZ</i>	1.1 (0.77)	0.7 (0.02)	1.1 (0.15, 0.52)
<i>B2M</i>	3.1 (0.83)	1.4 (0.03)	1.0 (0.04, 0.39)

Tight junction genes measured in IBS-C vs. HC by Peters et al. (97); 13 of 88 were differentially expressed (FDR < 0.05). FC, fold change, FDR, Benjamini-Hochberg adjusted *P* value.

2) Evidence for neurally-mediated mechanisms in IBS

cAMP-PKA signaling

WGCNA and IPA analysis identified up-regulation of cAMP-PKA signaling in IBS and IBS-C biopsies compared to HCs. The down-regulation of phosphodiesterase 2A (*PDE2A*) (validated by nCounter), which hydrolyzes cAMP is also supportive. PKA signaling activates T-type calcium channels, and in our dataset, expression of the T-type calcium channel Cav3.1 (*CACNA1G*) was increased in IBS (both IBS-C and IBS-D). The potential importance of this channel in IBS is supported by a study in rats, in which knockdown of Cav3.1 attenuated butyrate-induced visceral hypersensitivity.(83) In addition, Scanzi et al. found increased expression of an alternate isoform with similar properties(127) (Cav3.2, *CACNA1H*) in biopsies from IBS patients vs. HCs and in a mouse model of visceral pain.

While cAMP-PKA signaling is important in nociceptive pain signaling,(11) this pathway regulates many aspects of cellular physiology including chloride secretion(15) and serotonin release by

enterochromaffin cells.(34) Furthermore, while evidence exists for altered neuronal sensitivity in IBS (9, 16-18, 27, 38, 95, 100, 118, 124) both desensitization(95) and sensitization(124) may occur. For these reasons as well as the increased relative abundance of mucosal projections of enteric neurons as compared to extrinsic afferents in the colonic mucosa, our findings are more likely to represent signaling mechanisms unrelated to pain. It is also possible that the differences in cAMP-PKA signaling may involve other cell types. Interestingly one of the hub genes in the cAMP-PKA module is an epithelial cell apical sodium channel (*SCNN1D*, correlation with module eigengene 0.91), which may influence fluid transport and contribute to the constipation phenotype.(62)

Neurotransmitter and neuropeptide signaling

CRHR1 was down-regulated (Mayo study and current study microarray and nCounter), perhaps involving a response to increased stress-induced release of CRH. In a previous study,(30) *CRHR1* mRNA was not detectable in 81% of samples by quantitative real-time PCR, but microarray and nCounter are both more sensitive detection methods.

The gamma-aminobutyric acid type A receptor gamma1 subunit (*GABRG1*) was also down-regulated in IBS. In genome wide association study, a variant in this gene was associated with IBS within the discovery (but not replication) sample.(42) In the enteric nervous system, activation of GABA A receptors exerts an excitatory effect.(33) These receptors are also located on epithelial cells and agonists stimulate luminal chloride secretion.(73) Thus, potential relevance of GABA receptor down-regulation in IBS-C could be decreased peristalsis and/or secretion. An additional mechanism may be related to GABA receptor subunit stoichiometry. There are 19 GABA receptor subunit genes in the human genome and several dozen subunit combinations for the pentameric receptor have been described.(35) Differential composition of GABA receptors can affect membrane targeting and ligand sensitivity, which may be relevant in

IBS.(35, 65) It is also possible that the relevance to IBS is related to neuroimmune mechanisms as GABA receptors are present on immune cells.(8)

Neurite outgrowth

Several genes related to neurite outgrowth were dysregulated in IBS-C vs. HCs both in our cohort and Mayo (reticulon 3, *SSPO*, *GRFA4*). *SSPO*, which encodes the secreted glycoprotein SCO-spondin that may participate in neurite outgrowth and neuron/glia interaction (85) was previously found to be hyper-methylated in peripheral blood mononuclear cells in IBS.(82) Interestingly, changes in neuron density have been reported in full-thickness samples from animal models of IBS,(72, 116) and patients with slow-transit constipation.(123)

Additional notable findings

We also found that high mobility group box 1 (HMGB1) was up-regulated in IBS-C. HMGB1 is a chromatin-binding protein as well as a pro-inflammatory cytokine capable of increasing permeability of intestinal epithelial cell monolayers and ileal permeability and bacterial translocation in mice. While HMGB1 has well-defined roles in inflammation, it has recently emerged as a potential mediator of nociceptive pain signaling.(3, 40, 66) For example, in a mouse model of cyclophosphamide-induced cystitis it mediated mechanical bladder sensitivity induced by the protease activated receptor PAR4.(66) HMGB1 is a target of cyclin-dependent kinase 5 (*CDK5*),(117) which was also up-regulated and validated in our cohort. In a colon cancer derived cell line, CDK5 phosphorylated HMGB1 at serine-181, a post-translational modification that has been associated with increased secretion of HMGB1.(71) CDK5 activation is also pro-nociceptive through interactions with TRPV and purinergic receptors.(91, 126)

We would also like to note the enrichment of a few gene sets that met the threshold of $FDR < 0.25$. Specifically, enrichment of “signaling by FGFR1 mutants” in HCs vs. IBS overall may warrant further investigation into this pathway. In addition, enrichment of the gene sets “allograft

rejection”, “graft versus host disease” and “autoimmune thyroid disease” in men vs. women with IBS suggests that immune regulation may play a larger role in IBS among men.

Table 10: Expression of neuronal and glial/astrocytic markers

Gene Name, Symbol (alias, Entrez ID)	Mean	IBS-C vs. HC	IBS vs. HC	IBS-D vs. HC
		FC (p)	FC (p)	FC (p)
Neuronal				
Enolase2, ENO2 (neuron speific enolase, 2026)	6.8	1.0 (0.62)	1.0 (0.74)	1.1 (0.29)
Ubiquitin C-terminal hydrolase L1, UCHL1 (PGP9.5, 7345)	9.0	0.8 (0.2)	1.0 (0.9)	1.3 (0.1)
Microtubule associated protein 2, MAP2 (4133)	5.0	1.2 (0.12)	1.1 (0.42)	1.2 (0.31)
	8.2	0.9 (0.35)	1.0 (0.96)	1.0 (0.84)
Synaptophysin, SYP (6855)	6.5	1.1 (0.15)	1.1 (0.43)	1.0 (0.93)
Neurofilament, heavy, NEFH (4744)	15.2	1.0 (0.74)	1.0 (0.39)	0.9 (0.24)
Calbindin 2, CALB2 (calretinin, 794)	9.2	0.8 (0.14)	0.9 (0.3)	1.0 (0.77)
	7.5	0.6 (0.04)	0.7 (0.04)	0.7 (0.13)
Astrocyte/glial				
Proteolipid protein 1, PLP1 (5354)	7.3	0.8 (0.08)	0.9 (0.49)	1.1 (0.54)
	6.3	0.9 (0.46)	1.0 (0.72)	1.0 (0.9)
S100 calcium binding protein B, S100B (6285)	6.1	0.8 (0.14)	1.0 (0.8)	0.9 (0.28)
Glial fibrillary acidic protein, GFAP (2670)	7.9	0.9 (0.56)	0.9 (0.4)	0.9 (0.38)

Mean is mean log2 normalized signal; FC, fold change.

Limitations

The most important limitation to consider in the interpretation of these results is that the results are not cell-type specific and furthermore, there may be group differences in cell-type composition of the biopsies with no reliable methods to measure or account for these differences. The up-regulated gene sets in HCs vs. IBS-C suggests that either there may be more neurons and/or increased transcriptional activity in neurons. However, decreased enteric neurons in the submucosal plexus has previously been associated with constipation.(10) We attempted to address this question by looking at expression of traditional neuronal and astrocyte/glial markers (Table 10), which did not reveal significant differences. Additionally, there are limitations to drawing conclusions on neuronal physiology based on findings from mucosal biopsies. While the mucosa is innervated by sensory nerve fibers, and biopsies frequently include some submucosa, an alternative explanation is that the expression changes may reflect alterations in transcriptional activity of glial cells or enteroendocrine cells.

Another significant limitation is the small sample size of the microarray study. However, the composition (equal men and women, medication use) and rigor of phenotyping increases the probability that significant findings are not due to sampling bias. Despite the larger sample size

included in the IBS vs. HC comparison (n=20 vs. n=10 for IBS-C), we did not find DETs meeting a $FDR < 0.05$ threshold. This may be due to bowel habit-specific differences in peripheral mechanisms as has been suggested by others.(57) Importantly, our study was only powered to detect larger changes ($FC \geq 2$) at an $FDR \leq 0.05$. Therefore, caution should be used in the interpretation of negative results (absence of changes).

Reproducibility between IBS studies remains a challenge. There were 85 DETs in the literature which were present on our microarray, and we found the same direction of change (up or down in IBS vs. HCs) for 44 of these (Table 8). We did not find differences in SERT, which has been reported in several(36, 63, 122) but not all(20, 22) studies. While we have attempted to highlight findings that are reproducible through comparison with publicly available data, this approach has computational limitations as well as limitations due to differences in protocol (e.g. different bowel preparation) and patient populations. In particular, the current study used Rome III(80) criteria while Mayo and Nottingham used Rome II(115) which are less specific and may identify a less severe population. Given that inclusion of IBS patients in all three cohorts was supervised by experts in the field, the differences in diagnostic criteria are less likely to result in sample differences than for survey-based studies. We recommend pooling resources and increasing collaboration with the goal of larger samples that are uniformly phenotyped and processed in order to yield robust findings.

CONCLUSIONS

In conclusion, using a small but well-phenotyped and age- and sex-balanced sample, we have found additional support for peripheral mechanisms in IBS and demonstrate that these findings were more significant in IBS-C patients within our cohort, suggesting that pathophysiologic processes differ between IBS-C and IBS-D. Based on some of the observed changes, future investigations should focus on alterations in neuronal physiology.

REFERENCES:

1. Abe T, Kanemitsu Y, Nakasone M, Kawahata I, Yamakuni T, Nakajima A, Suzuki N, Nishikawa M, Hishinuma T, and Tomioka Y. SLC10A4 is a protease-activated transporter that transports bile acids. *J Biochem* 154: 93-101, 2013.
2. Aerssens J, Camilleri M, Talloen W, Thielemans L, Gohlmann HW, van dW, I, Thielemans T, De HR, Andrews CN, Bharucha AE, Carlson PJ, Busciglio I, Burton DD, Smyrk T, Urrutia R, and Coulie B. Alterations in mucosal immunity identified in the colon of patients with irritable bowel syndrome. *Clin Gastroenterol Hepatol* 6: 194-205, 2008.
3. Agalave NM, and Svensson CI. Extracellular high-mobility group box 1 protein (HMGB1) as a mediator of persistent pain. *Mol Med* 20: 569-578, 2015.
4. An S, Zong G, Wang Z, Shi J, Du H, and Hu J. Expression of inducible nitric oxide synthase in mast cells contributes to the regulation of inflammatory cytokines in irritable bowel syndrome with diarrhea. *Neurogastroenterol Motil* 28: 1083-1093, 2016.
5. Arloth J, Bader DM, Roh S, and Altmann A. Re-Annotator: Annotation Pipeline for Microarray Probe Sequences. *PLoS One* 10: e0139516, 2015.
6. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, and Sherlock G. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 25: 25-29, 2000.
7. Atkinson W, Lockhart S, Whorwell PJ, Keevil B, and Houghton LA. Altered 5-hydroxytryptamine signaling in patients with constipation- and diarrhea-predominant irritable bowel syndrome. *Gastroenterology* 130: 34-43, 2006.
8. Auteri M, Zizzo MG, and Serio R. GABA and GABA receptors in the gastrointestinal tract: from motility to inflammation. *Pharmacol Res* 93: 11-21, 2015.
9. Barbara G, Wang B, Stanghellini V, de Giorgio R, Cremon C, Di Nardo G, Trevisani M, Campi B, Geppetti P, Tonini M, Bunnett NW, Grundy D, and Corinaldesi R. Mast cell-dependent excitation of visceral-nociceptive sensory neurons in irritable bowel syndrome. *Gastroenterology* 132: 26-37, 2007.
10. Bassotti G, Villanacci V, Nascimbeni R, Asteria CR, Fisogni S, Nesi G, Legrenzi L, Mariano M, Tonelli F, Morelli A, and Salerni B. Colonic neuropathological aspects in patients with intractable constipation due to obstructed defecation. *Modern Pathology* 20: 367-374, 2007.
11. Bavencoffe A, Li Y, Wu Z, Yang Q, Herrera J, Kennedy EJ, Walters ET, and Dessauer CW. Persistent electrical activity in primary nociceptors after spinal cord injury is maintained by scaffolded adenylyl cyclase and protein kinase A and is associated with altered adenylyl cyclase regulation. *J Neurosci* 36: 1660-1668, 2016.
12. Belmonte L, Beutheu Youmba S, Bertiaux-Vandaele N, Antonietti M, Lecleire S, Zalar A, Gourcerol G, Leroi AM, Dechelotte P, Coeffier M, and Ducrotte P. Role of toll like

- receptors in irritable bowel syndrome: differential mucosal immune activation according to the disease subtype. *PLoS One* 7: e42777, 2012.
13. Bennet SM, Polster A, Tornblom H, Isaksson S, Capronnier S, Tessier A, Le Neve B, Simren M, and Ohman L. Global cytokine profiles and association with clinical characteristics in patients with irritable bowel syndrome. *Am J Gastroenterol* 111: 1165-1176, 2016.
 14. Bojjireddy N, Guzman-Hernandez ML, Reinhard NR, Jovic M, and Balla T. EFR3s are palmitoylated plasma membrane proteins that control responsiveness to G-protein-coupled receptors. *J Cell Sci* 128: 118-128, 2015.
 15. Borthwick LA, McGaw J, Conner G, Taylor CJ, Gerke V, Mehta A, Robson L, and Muimo R. The formation of the cAMP/protein kinase A-dependent annexin 2-S100A10 complex with cystic fibrosis conductance regulator protein (CFTR) regulates CFTR channel function. *Mol Biol Cell* 18: 3388-3397, 2007.
 16. Buhner S, Braak B, Li Q, Kugler EM, Klooker T, Wouters M, Donovan J, Vignali S, Mazzuoli-Weber G, Grundy D, Boeckstaens G, and Schemann M. Neuronal activation by mucosal biopsy supernatants from irritable bowel syndrome patients is linked to visceral sensitivity. *Exp Physiol* 99: 1299-1311, 2014.
 17. Buhner S, Li Q, Berger T, Vignali S, Barbara G, De Giorgio R, Stanghellini V, and Schemann M. Submucous rather than myenteric neurons are activated by mucosal biopsy supernatants from irritable bowel syndrome patients. *Neurogastroenterol Motil* 24: 1134-e1572, 2012.
 18. Buhner S, Li Q, Vignali S, Barbara G, De Giorgio R, Stanghellini V, Cremon C, Zeller F, Langer R, Daniel H, Michel K, and Schemann M. Activation of human enteric neurons by supernatants of colonic biopsy specimens from patients with irritable bowel syndrome. *Gastroenterology* 137: 1425-1434, 2009.
 19. Camilleri M. Physiological underpinnings of irritable bowel syndrome: neurohormonal mechanisms. *J Physiol* 592: 2967-2980, 2014.
 20. Camilleri M, Andrews CN, Bharucha AE, Carlson PJ, Ferber I, Stephens D, Smyrk TC, Urrutia R, Aerssens J, Thielemans L, Gohlmann H, van den Wyngaert I, and Coulie B. Alterations in expression of p11 and SERT in mucosal biopsy specimens of patients with irritable bowel syndrome. *Gastroenterology* 132: 17-25, 2007.
 21. Camilleri M, Carlson P, Acosta A, and Busciglio I. Colonic mucosal gene expression and genotype in irritable bowel syndrome patients with normal or elevated fecal bile acid excretion. *Am J Physiol Gastrointest Liver Physiol* 309: G10-20, 2015.
 22. Camilleri M, Carlson P, Acosta A, Busciglio I, Nair AA, Gibbons SJ, Farrugia G, and Klee EW. RNA sequencing shows transcriptomic changes in rectosigmoid mucosa in patients with irritable bowel syndrome-diarrhea: a pilot case-control study. *Am J Physiol Gastrointest Liver Physiol* 306: G1089-1098, 2014.
 23. Camilleri M, Carlson P, Valentin N, Acosta A, O'Neill J, Eckert D, Dyer R, Na J, Klee EW, and Murray JA. Pilot study of small bowel mucosal gene expression in patients with

- irritable bowel syndrome with diarrhea. *Am J Physiol Gastrointest Liver Physiol* 311: G365-376, 2016.
24. Camilleri M, Oduyebo I, and Halawi H. Chemical and molecular factors in irritable bowel syndrome: current knowledge, challenges, and unanswered questions. *Am J Physiol Gastrointest Liver Physiol* 311: G777-G784, 2016.
 25. Canavan C, West J, and Card T. Review article: the economic impact of the irritable bowel syndrome. *Aliment Pharmacol Ther* 40: 1023-1034, 2014.
 26. Carrel D, Masson J, Al Awabdh S, Capra CB, Lenkei Z, Hamon M, Emerit MB, and Darmon M. Targeting of the 5-HT_{1A} serotonin receptor to neuronal dendrites is mediated by Yif1B. *J Neurosci* 28: 8063-8073, 2008.
 27. Cenac N, Andrews CN, Holzhausen M, Chapman K, Cottrell G, Andrade-Gordon P, Steinhoff M, Barbara G, Beck P, Bunnett NW, Sharkey KA, Ferraz JG, Shaffer E, and Vergnolle N. Role for protease activity in visceral pain in irritable bowel syndrome. *J Clin Invest* 117: 636-647, 2007.
 28. Cenac N, Bautzova T, Le Faouder P, Veldhuis NA, Poole DP, Rolland C, Bertrand J, Liedtke W, Dubourdeau M, Bertrand-Michel J, Zecchi L, Stanghellini V, Bunnett NW, Barbara G, and Vergnolle N. Quantification and Potential Functions of Endogenous Agonists of Transient Receptor Potential Channels in Patients With Irritable Bowel Syndrome. *Gastroenterology* 149: 433-444 e437, 2015.
 29. Chang K, Seabold GK, Wang CY, and Wentholt RJ. Reticulon 3 is an interacting partner of the SALM family of adhesion molecules. *J Neurosci Res* 88: 266-274, 2010.
 30. Chang L, Adeyemo M, Karagiannides I, Videlock EJ, Bowe C, Shih W, Presson AP, Yuan PQ, Cortina G, Gong H, Singh S, Licudine A, Mayer M, Tache Y, Pothoulakis C, and Mayer EA. Serum and colonic mucosal immune markers in irritable bowel syndrome. *Am J Gastroenterol* 107: 262-272, 2012.
 31. Chen JB, Zhang J, Hu HZ, Xue M, and Jin YJ. Polymorphisms of TGFB1, TLE4 and MUC22 are associated with childhood asthma in Chinese population. *Allergol Immunopathol (Madr)* 2017.
 32. Chen X, Li L, Khan MN, Shi L, Wang Z, Zheng F, Gong F, and Fang M. HMGB1 exacerbates experimental mouse colitis by enhancing innate lymphoid cells 3 inflammatory responses via promoted IL-23 production. *Innate Immun* 2016.
 33. Cherubini E, and North RA. Actions of gamma-aminobutyric acid on neurones of guinea-pig myenteric plexus. *Br J Pharmacol* 82: 93-100, 1984.
 34. Chin A, Svejda B, Gustafsson BI, Granlund AB, Sandvik AK, Timberlake A, Sumpio B, Pfragner R, Modlin IM, and Kidd M. The role of mechanical forces and adenosine in the regulation of intestinal enterochromaffin cell serotonin secretion. *Am J Physiol Gastrointest Liver Physiol* 302: G397-405, 2012.
 35. Chua HC, and Chebib M. GABAA Receptors and the Diversity in their Structure and Pharmacology. *Adv Pharmacol* 79: 1-34, 2017.

36. Coates MD, Mahoney CR, Linden DR, Sampson JE, Chen J, Blaszyk H, Crowell MD, Sharkey KA, Gershon MD, Mawe GM, and Moses PL. Molecular defects in mucosal serotonin content and decreased serotonin reuptake transporter in ulcerative colitis and irritable bowel syndrome. *Gastroenterology* 126: 1657-1664, 2004.
37. Collins FS, and Tabak LA. Policy: NIH plans to enhance reproducibility. *Nature* 505: 612-613, 2014.
38. Cremon C, Carini G, Wang B, Vasina V, Cogliandro RF, De Giorgio R, Stanghellini V, Grundy D, Tonini M, De Ponti F, Corinaldesi R, and Barbara G. Intestinal serotonin release, sensory neuron activation, and abdominal pain in irritable bowel syndrome. *Am J Gastroenterol* 106: 1290-1298, 2011.
39. Cremon C, Gargano L, Morselli-Labate AM, Santini D, Cogliandro RF, De Giorgio R, Stanghellini V, Corinaldesi R, and Barbara G. Mucosal immune activation in irritable bowel syndrome: gender-dependence and association with digestive symptoms. *Am J Gastroenterol* 104: 392-400, 2009.
40. Das N, Dewan V, Grace PM, Gunn RJ, Tamura R, Tzarum N, Watkins LR, Wilson IA, and Yin H. HMGB1 Activates Proinflammatory Signaling via TLR5 Leading to Allodynia. *Cell Rep* 17: 1128-1140, 2016.
41. Dunlop SP, Hebden J, Campbell E, Naesdal J, Olbe L, Perkins AC, and Spiller RC. Abnormal intestinal permeability in subgroups of diarrhea-predominant irritable bowel syndromes. *Am J Gastroenterol* 101: 1288-1294, 2006.
42. Ek WE, Reznichenko A, Ripke S, Niesler B, Zucchelli M, Rivera NV, Schmidt PT, Pedersen NL, Magnusson P, Talley NJ, Holliday EG, Houghton L, Gazouli M, Karamanolis G, Rappold G, Burwinkel B, Surowy H, Rafter J, Assadi G, Li L, Papadaki E, Gambaccini D, Marchi S, Colucci R, Blandizzi C, Barbaro R, Karling P, Walter S, Ohlsson B, Tornblom H, Bresso F, Andreasson A, Dlugosz A, Simren M, Agreus L, Lindberg G, Boeckxstaens G, Bellini M, Stanghellini V, Barbara G, Daly MJ, Camilleri M, Wouters MM, and D'Amato M. Exploring the genetics of irritable bowel syndrome: a GWA study in the general population and replication in multinational case-control cohorts. *Gut* 2014.
43. Fabregat A, Sidiropoulos K, Garapati P, Gillespie M, Hausmann K, Haw R, Jassal B, Jupe S, Korninger F, McKay S, Matthews L, May B, Milacic M, Rothfels K, Shamovsky V, Webber M, Weiser J, Williams M, Wu G, Stein L, Hermjakob H, and D'Eustachio P. The Reactome pathway Knowledgebase. *Nucleic Acids Res* 44: D481-487, 2016.
44. Falcon S, and Gentleman R. Using GOstats to test gene lists for GO term association. *Bioinformatics* 23: 257-258, 2007.
45. Faure C, Patey N, Gauthier C, Brooks EM, and Mawe GM. Serotonin signaling is altered in irritable bowel syndrome with diarrhea but not in functional dyspepsia in pediatric age patients. *Gastroenterology* 139: 249-258, 2010.
46. Gautier L, Cope L, Bolstad BM, and Irizarry RA. affy--analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* 20: 307-315, 2004.

47. Gentleman R. annotate: Annotation for microarrays. In: *R package version 1550* 2017.
48. Ghaleb AM, Bialkowska AB, Snider AJ, Gnatenko DV, Hannun YA, Yang VW, and Schmidt VA. IQ Motif-Containing GTPase-Activating Protein 2 (IQGAP2) Is a Novel Regulator of Colonic Inflammation in Mice. *PLoS One* 10: e0129314, 2015.
49. Gralnek IM, Hays RD, Kilbourne A, and Mayer EA. Impact of irritable bowel syndrome on health-related quality of life (HRQOL). *Gastroenterology* 116: A1038, 1999.
50. Guo X, and Qi X. VCP cooperates with UBXD1 to degrade mitochondrial outer membrane protein MCL1 in model of Huntington's disease. *Biochim Biophys Acta* 1863: 552-559, 2017.
51. Hickey FB, Corcoran JB, Docherty NG, Griffin B, Bhreathnach U, Furlong F, Martin F, Godson C, and Murphy M. IHG-1 promotes mitochondrial biogenesis by stabilizing PGC-1alpha. *J Am Soc Nephrol* 22: 1475-1485, 2011.
52. Hijikata M, Matsushita I, Tanaka G, Tsuchiya T, Ito H, Tokunaga K, Ohashi J, Homma S, Kobashi Y, Taguchi Y, Azuma A, Kudoh S, and Keicho N. Molecular cloning of two novel mucin-like genes in the disease-susceptibility locus for diffuse panbronchiolitis. *Hum Genet* 129: 117-128, 2011.
53. Hofer-Warbinek R, Schmid JA, Mayer H, Winsauer G, Orel L, Mueller B, Wiesner C, Binder BR, and de Martin R. A highly conserved proapoptotic gene, IKIP, located next to the APAF1 gene locus, is regulated by p53. *Cell Death Differ* 11: 1317-1325, 2004.
54. Houghton LA, Atkinson W, Lockhart C, Whorwell PJ, and Keevil B. Sigmoid-colonic motility in health and irritable bowel syndrome: a role for 5-hydroxytryptamine. *Neurogastroenterol Motil* 19: 724-731, 2007.
55. Huang da W, Sherman BT, and Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* 37: 1-13, 2009.
56. Huang da W, Sherman BT, and Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4: 44-57, 2009.
57. Hughes PA, Harrington AM, Castro J, Liebrechts T, Adam B, Grasby DJ, Isaacs NJ, Maldeniya L, Martin CM, Persson J, Andrews JM, Holtmann G, Blackshaw LA, and Brierley SM. Sensory neuro-immune interactions differ between irritable bowel syndrome subtypes. *Gut* 62: 1456-1465, 2013.
58. Ichikawa K, Kubota Y, Nakamura T, Weng JS, Tomida T, Saito H, and Takekawa M. MCRIPI, an ERK substrate, mediates ERK-induced gene silencing during epithelial-mesenchymal transition by regulating the co-repressor CtBP. *Mol Cell* 58: 35-46, 2015.
59. Ito T, Niwa J, Hishikawa N, Ishigaki S, Doyu M, and Sobue G. Dofin localizes to Lewy bodies and ubiquitylates synphilin-1. *J Biol Chem* 278: 29106-29114, 2003.
60. Kanehisa M, and Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28: 27-30, 2000.

61. Kauffmann A, Rayner TF, Parkinson H, Kapushesky M, Lukk M, Brazma A, and Huber W. Importing ArrayExpress datasets into R/Bioconductor. *Bioinformatics* 25: 2092-2094, 2009.
62. Kellenberger S, and Schild L. Epithelial sodium channel/degenerin family of ion channels: a variety of functions for a shared structure. *Physiol Rev* 82: 735-767, 2002.
63. Kerckhoffs AP, ter Linde JJ, Akkermans LM, and Samsom M. SERT and TPH-1 mRNA expression are reduced in irritable bowel syndrome patients regardless of visceral sensitivity state in large intestine. *Am J Physiol Gastrointest Liver Physiol* 302: G1053-1060, 2012.
64. Keszthelyi D, Troost FJ, Jonkers DM, Helyes Z, Hamer HM, Ludidi S, Vanhoutvin S, Venema K, Dekker J, Szolcsanyi J, and Masclee AA. Alterations in mucosal neuropeptides in patients with irritable bowel syndrome and ulcerative colitis in remission: a role in pain symptom generation? *Eur J Pain* 17: 1299-1306, 2013.
65. Khom S, Baburin I, Timin EN, Hohaus A, Sieghart W, and Hering S. Pharmacological properties of GABAA receptors containing gamma1 subunits. *Mol Pharmacol* 69: 640-649, 2006.
66. Kouzoukas DE, Ma F, Meyer-Siegler KL, Westlund KN, Hunt DE, and Vera PL. Protease-activated receptor 4 induces bladder pain through high mobility group box-1. *PLoS One* 11: e0152055, 2016.
67. Langfelder P, and Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9: 559, 2008.
68. Larhammar M, Patra K, Blunder M, Emilsson L, Peuckert C, Arvidsson E, Ronnlund D, Preobraschenski J, Birgner C, Limbach C, Widengren J, Blom H, Jahn R, Wallen-Mackenzie A, and Kullander K. SLC10A4 is a vesicular amine-associated transporter modulating dopamine homeostasis. *Biol Psychiatry* 77: 526-536, 2015.
69. Law CW, Chen Y, Shi W, and Smyth GK. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol* 15: R29, 2014.
70. Lee AK, and Potts PR. A Comprehensive Guide to the MAGE Family of Ubiquitin Ligases. *J Mol Biol* 429: 1114-1142, 2017.
71. Lee H, Park M, Shin N, Kim G, Kim YG, Shin JS, and Kim H. High mobility group box-1 is phosphorylated by protein kinase C zeta and secreted in colon cancer cells. *Biochem Biophys Res Commun* 424: 321-326, 2012.
72. Li S, Fei G, Fang X, Yang X, Sun X, Qian J, Wood JD, and Ke M. Changes in enteric neurons of small intestine in a rat model of irritable bowel syndrome with diarrhea. *J Neurogastroenterol Motil* 22: 310-320, 2016.
73. Li Y, Xiang YY, Lu WY, Liu C, and Li J. A novel role of intestine epithelial GABAergic signaling in regulating intestinal fluid secretion. *Am J Physiol Gastrointest Liver Physiol* 303: G453-460, 2012.

74. Li Z, Mou H, Wang T, Xue J, Deng B, Qian L, Zhou Y, Gong W, Wang JM, Wu G, Zhou CF, Fang J, and Le Y. A non-secretory form of FAM3B promotes invasion and metastasis of human colon cancer cells by upregulating Slug expression. *Cancer Lett* 328: 278-284, 2013.
75. Liang WJ, Zhang G, Luo HS, Liang LX, Huang D, and Zhang FC. Tryptase and Protease-Activated Receptor 2 Expression Levels in Irritable Bowel Syndrome. *Gut Liver* 10: 382-390, 2016.
76. Liao XJ, Mao WM, Wang Q, Yang GG, Wu WJ, and Shao SX. MicroRNA-24 inhibits serotonin reuptake transporter expression and aggravates irritable bowel syndrome. *Biochem Biophys Res Commun* 469: 288-293, 2016.
77. Lin SW, Chen JC, Hsu LC, Hsieh CL, and Yoshida A. Human gamma-aminobutyraldehyde dehydrogenase (ALDH9): cDNA sequence, genomic organization, polymorphism, chromosomal localization, and tissue expression. *Genomics* 34: 376-380, 1996.
78. Liu P, and Hwang JT. Quick calculation for sample size while controlling false discovery rate with application to microarray analysis. *Bioinformatics* 23: 739-746, 2007.
79. Liu Y, Lear T, Zhao Y, Zhao J, Zou C, Chen BB, and Mallampalli RK. F-box protein Fbxl18 mediates polyubiquitylation and proteasomal degradation of the pro-apoptotic SCF subunit Fbxl7. *Cell Death Dis* 6: e1630, 2015.
80. Longstreth GF, Thompson WG, Chey WD, Houghton LA, Mearin F, and Spiller RC. Functional bowel disorders. *Gastroenterology* 130: 1480-1491, 2006.
81. Luo Y, Feng C, Wu J, Wu Y, Liu D, Wu J, Dai F, and Zhang J. P2Y1, P2Y2, and TRPV1 Receptors Are Increased in Diarrhea-Predominant Irritable Bowel Syndrome and P2Y2 Correlates with Abdominal Pain. *Dig Dis Sci* 61: 2878-2886, 2016.
82. Mahurkar S, Polytarchou C, Iliopoulos D, Pothoulakis C, Mayer EA, and Chang L. Genome-wide DNA methylation profiling of peripheral blood mononuclear cells in irritable bowel syndrome. *Neurogastroenterol Motil* 28: 410-422, 2016.
83. Marger F, Gelot A, Alloui A, Matricon J, Ferrer JF, Barrere C, Pizzoccaro A, Muller E, Nargeot J, Snutch TP, Eschalier A, Bourinet E, and Ardid D. T-type calcium channels contribute to colonic hypersensitivity in a rat model of irritable bowel syndrome. *Proc Natl Acad Sci U S A* 108: 11268-11273, 2011.
84. Mearin F, Lacy BE, Chang L, Chey WD, Lembo AJ, Simren M, and Spiller R. Bowel Disorders. *Gastroenterology* 2016.
85. Meiniel A, Meiniel R, Goncalves-Mendes N, Creveaux I, Didier R, and Dastugue B. The thrombospondin type 1 repeat (TSR) and neuronal differentiation: roles of SCO-spondin oligopeptides on neuronal cell types and cell lines. *Int Rev Cytol* 230: 1-39, 2003.
86. Meiniel O, and Meiniel A. The complex multidomain organization of SCO-spondin protein is highly conserved in mammals. *Brain Research Reviews* 53: 321-327, 2007.

87. Melief EJ, Gibbs JT, Li X, Morgan RG, Keene CD, Montine TJ, Palmiter RD, and Darvas M. Characterization of cognitive impairments and neurotransmitter changes in a novel transgenic mouse lacking Slc10a4. *Neuroscience* 324: 399-406, 2016.
88. Missler M, and Sudhof TC. Neurexophilins form a conserved family of neuropeptide-like glycoproteins. *J Neurosci* 18: 3630-3638, 1998.
89. Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J, Puigserver P, Carlsson E, Ridderstrale M, Laurila E, Houstis N, Daly MJ, Patterson N, Mesirov JP, Golub TR, Tamayo P, Spiegelman B, Lander ES, Hirschhorn JN, Altshuler D, and Groop LC. PGC-1 α -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet* 34: 267-273, 2003.
90. Munakata J, Naliboff B, Harraf F, Kodner A, Lembo T, Chang L, Silverman DH, and Mayer EA. Repetitive sigmoid stimulation induces rectal hyperalgesia in patients with irritable bowel syndrome. *Gastroenterology* 112: 55-63, 1997.
91. Nair A, Simonetti M, Fabbretti E, and Nistri A. The Cdk5 kinase downregulates ATP-gated ionotropic P2X3 receptor function via serine phosphorylation. *Cell Mol Neurobiol* 30: 505-509, 2010.
92. Nalaskowski MM, Fliegert R, Ernst O, Brehm MA, Fanick W, Windhorst S, Lin H, Giehler S, Hein J, Lin YN, and Mayr GW. Human inositol 1,4,5-trisphosphate 3-kinase isoform B (IP3KB) is a nucleocytoplasmic shuttling protein specifically enriched at cortical actin filaments and at invaginations of the nuclear envelope. *J Biol Chem* 286: 4500-4510, 2011.
93. O'Malley D, Buckley MM, McKernan DP, Quigley EM, Cryan JF, and Dinan TG. Soluble mediators in plasma from irritable bowel syndrome patients excite rat submucosal neurons. *Brain Behav Immun* 44: 57-67, 2015.
94. Ohman L, Lindmark AC, Isaksson S, Posserud I, Strid H, Sjovall H, and Simren M. Increased TLR2 expression on blood monocytes in irritable bowel syndrome patients. *Eur J Gastroenterol Hepatol* 24: 398-405, 2012.
95. Ostertag D, Buhner S, Michel K, Pehl C, Kurjak M, Gotzberger M, Schulte-Frohlinde E, Frieling T, Enck P, Phillip J, and Schemann M. Reduced responses of submucous neurons from irritable bowel syndrome patients to a cocktail containing histamine, serotonin, tnfalpa, and tryptase (IBS-cocktail). *Front Neurosci* 9: 465, 2015.
96. Ouederni M, Vincent QB, Frange P, Touzot F, Scerra S, Bejaoui M, Bousfiha A, Levy Y, Lisowska-Grospierre B, Canioni D, Bruneau J, Debre M, Blanche S, Abel L, Casanova JL, Fischer A, and Picard C. Major histocompatibility complex class II expression deficiency caused by a RFXANK founder mutation: a survey of 35 patients. *Blood* 118: 5108-5118, 2011.
97. Peters SA, Edogawa S, Sundt WJ, Dyer RB, Dalenberg DA, Mazzone A, Singh RJ, Moses N, Smyrk TC, Weber C, Linden DR, MacNaughton WK, Turner JR, Camilleri M, Katzka DA, Farrugia G, and Grover M. Constipation-predominant irritable bowel syndrome females have normal colonic barrier and secretory function. *Am J Gastroenterol* 112: 913-923, 2017.

98. Qin B, Dong L, Guo X, Jiang J, He Y, Wang X, Li L, and Zhao J. Expression of G protein-coupled estrogen receptor in irritable bowel syndrome and its clinical significance. *Int J Clin Exp Pathol* 7: 2238-2246, 2014.
99. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2013.
100. Rolland-Fourcade C, Denadai-Souza A, Cirillo C, Lopez C, Jaramillo JO, Desormeaux C, Cenac N, Motta JP, Larauche M, Tache Y, Berghe PV, Neunlist M, Coron E, Kirzin S, Portier G, Bonnet D, Alric L, Vanner S, Deraison C, and Vergnolle N. Epithelial expression and function of trypsin-3 in irritable bowel syndrome. *Gut* 66: 1767-1778, 2017.
101. Rong Z, Wang A, Li Z, Ren Y, Cheng L, Li Y, Wang Y, Ren F, Zhang X, Hu J, and Chang Z. IL-17RD (Sef or IL-17RLM) interacts with IL-17 receptor and mediates IL-17 signaling. *Cell Res* 19: 208-215, 2009.
102. Scanzi J, Accarie A, Muller E, Pereira B, Aissouni Y, Goutte M, Joubert-Zakeyh J, Picard E, Boudieu L, Mallet C, Gelot A, Ardid D, Carvalho FA, and Dapoigny M. Colonic overexpression of the T-type calcium channel Cav 3.2 in a mouse model of visceral hypersensitivity and in irritable bowel syndrome patients. *Neurogastroenterol Motil* 28: 1632-1640, 2016.
103. Schmidt S, Moncada M, Burger S, and Geyer J. Expression, sorting and transport studies for the orphan carrier SLC10A4 in neuronal and non-neuronal cell lines and in *Xenopus laevis* oocytes. *BMC Neurosci* 16: 35, 2015.
104. Scoumanne A, Molina-Ortiz P, Monteyne D, Perez-Morga D, Erneux C, and Schurmans S. Specific expression and function of inositol 1,4,5-trisphosphate 3-kinase C (ITPKC) in wild type and knock-out mice. *Adv Biol Regul* 62: 1-10, 2016.
105. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R, and Dunbar GC. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry* 59 Suppl 20: 22-33;quiz 34-57, 1998.
106. Shiotani A, Kusunoki H, Kimura Y, Ishii M, Imamura H, Tarumi K, Manabe N, Kamada T, Hata J, and Haruma K. S100A expression and interleukin-10 polymorphisms are associated with ulcerative colitis and diarrhea predominant irritable bowel syndrome. *Dig Dis Sci* 58: 2314-2323, 2013.
107. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 3: Article3, 2004.
108. Song X, Zhu S, Shi P, Liu Y, Shi Y, Levin SD, and Qian Y. IL-17RE is the functional receptor for IL-17C and mediates mucosal immunity to infection with intestinal pathogens. *Nat Immunol* 12: 1151-1158, 2011.
109. Splinter PL, Lazaridis KN, Dawson PA, and LaRusso NF. Cloning and expression of SLC10A4, a putative organic anion transport protein. *World J Gastroenterol* 12: 6797-6805, 2006.

110. Stygelbout V, Leroy K, Pouillon V, Ando K, D'Amico E, Jia Y, Luo HR, Duyckaerts C, Erneux C, Schurmans S, and Brion JP. Inositol trisphosphate 3-kinase B is increased in human Alzheimer brain and exacerbates mouse Alzheimer pathology. *Brain* 137: 537-552, 2014.
111. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, and Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 102: 15545-15550, 2005.
112. Swan C, Duroudier NP, Campbell E, Zaitoun A, Hastings M, Dukes GE, Cox J, Kelly FM, Wilde J, Lennon MG, Neal KR, Whorwell PJ, Hall IP, and Spiller RC. Identifying and testing candidate genetic polymorphisms in the irritable bowel syndrome (IBS): association with TNFSF15 and TNF α . *Gut* 62: 985-994, 2013.
113. Szlufcik K, Missiaen L, Parys JB, Callewaert G, and De Smedt H. Uncoupled IP3 receptor can function as a Ca²⁺-leak channel: cell biological and pathological consequences. *Biol Cell* 98: 1-14, 2006.
114. Tay PN, Tan P, Lan Y, Leung CH, Laban M, Tan TC, Ni H, Manikandan J, Rashid SB, Yan B, Yap CT, Lim LH, Lim YC, and Hooi SC. Palladin, an actin-associated protein, is required for adherens junction formation and intercellular adhesion in HCT116 colorectal cancer cells. *Int J Oncol* 37: 909-926, 2010.
115. Thompson WG, Longstreth GF, Drossman DA, Heaton KW, Irvine EJ, and Muller-Lissner SA. Functional bowel disorders and functional abdominal pain. *Gut* 45 Suppl 2: II43-47, 1999.
116. Traini C, Evangelista S, Girod V, Faussone-Pellegrini MS, and Vannucchi MG. Changes of excitatory and inhibitory neurotransmitters in the colon of rats underwent to the wrap partial restraint stress. *Neurogastroenterol Motil* 28: 1172-1185, 2016.
117. Ugrinova I, Pashev IG, and Pasheva EA. Cyclin-dependent kinase 5 phosphorylates mammalian HMGB1 protein only if acetylated. *J Biochem* 149: 563-568, 2011.
118. Valdez-Morales EE, Overington J, Guerrero-Alba R, Ochoa-Cortes F, Ibeakanma CO, Spreadbury I, Bunnett NW, Beyak M, and Vanner SJ. Sensitization of peripheral sensory nerves by mediators from colonic biopsies of diarrhea-predominant irritable bowel syndrome patients: a role for PAR2. *Am J Gastroenterol* 108: 1634-1643, 2013.
119. Vazquez-Roque MI, Camilleri M, Smyrk T, Murray JA, O'Neill J, Carlson P, Lamsam J, Eckert D, Janzow D, Burton D, Ryks M, Rhoten D, and Zinsmeister AR. Association of HLA-DQ gene with bowel transit, barrier function, and inflammation in irritable bowel syndrome with diarrhea. *Am J Physiol Gastrointest Liver Physiol* 303: G1262-1269, 2012.
120. Waggott D, Chu K, Yin S, Wouters BG, Liu FF, and Boutros PC. NanoStringNorm: an extensible R package for the pre-processing of NanoString mRNA and miRNA data. *Bioinformatics* 28: 1546-1548, 2012.

121. Wang JP, and Hou XH. Expression of aquaporin 8 in colonic epithelium with diarrhoea-predominant irritable bowel syndrome. *Chin Med J (Engl)* 120: 313-316, 2007.
122. Wang YM, Chang Y, Chang YY, Cheng J, Li J, Wang T, Zhang QY, Liang DC, Sun B, and Wang BM. Serotonin transporter gene promoter region polymorphisms and serotonin transporter expression in the colonic mucosa of irritable bowel syndrome patients. *Neurogastroenterol Motil* 24: 560-565, e254-565, 2012.
123. Wedel T, Spiegler J, Soellner S, Roblick UJ, Schiedeck TH, Bruch HP, and Krammer HJ. Enteric nerves and interstitial cells of Cajal are altered in patients with slow-transit constipation and megacolon. *Gastroenterology* 123: 1459-1467, 2002.
124. Wouters MM, Balemans D, Van Wanrooy S, Dooley J, Cibert-Goton V, Alpizar YA, Valdez-Morales EE, Nasser Y, Van Veldhoven PP, Vanbrabant W, Van der Merwe S, Mols R, Ghesquiere B, Cirillo C, Kortekaas I, Carmeliet P, Peetermans WE, Vermeire S, Rutgeerts P, Augustijns P, Hellings PW, Belmans A, Vanner S, Bulmer DC, Talavera K, Vanden Berghe P, Liston A, and Boeckxstaens GE. Histamine receptor H1-mediated sensitization of TRPV1 mediates visceral hypersensitivity and symptoms in patients with irritable bowel syndrome. *Gastroenterology* 150: 875-887 e879, 2016.
125. Wouters MM, Lambrechts D, Knapp M, Cleynen I, Whorwell P, Agréus L, Dlugosz A, Schmidt PT, Halfvarson J, Simrén M, Ohlsson B, Karling P, Van Wanrooy S, Mondelaers S, Vermeire S, Lindberg G, Spiller R, Dukes G, D'Amato M, and Boeckxstaens G. Genetic variants in CDC42 and NXPH1 as susceptibility factors for constipation and diarrhoea predominant irritable bowel syndrome. *Gut* 63: 1103-1111, 2014.
126. Yang YR, He Y, Zhang Y, Li Y, Li Y, Han Y, Zhu H, and Wang Y. Activation of cyclin-dependent kinase 5 (Cdk5) in primary sensory and dorsal horn neurons by peripheral inflammation contributes to heat hyperalgesia. *Pain* 127: 109-120, 2007.
127. Yunker AM. Modulation and pharmacology of low voltage-activated ("T-Type") calcium channels. *J Bioenerg Biomembr* 35: 577-598, 2003.
128. Zelano J, Mikulovic S, Patra K, Kuhnemund M, Larhammar M, Emilsson L, Leao RN, and Kullander K. The synaptic protein encoded by the gene Slc10A4 suppresses epileptiform activity and regulates sensitivity to cholinergic chemoconvulsants. *Exp Neurol* 239: 73-81, 2013.
129. Zhang R, Zou N, Li J, Lv H, Wei J, Fang X-C, and Qian J-M. Elevated expression of c-fos in central nervous system correlates with visceral hypersensitivity in irritable bowel syndrome (IBS): a new target for IBS treatment. *Int J Colorectal Dis* 26: 1035-1044, 2011.
130. Zhao JH, Dong L, Shi HT, Wang ZY, Shi HY, and Ding H. The expression of protease-activated receptor 2 and 4 in the colon of irritable bowel syndrome patients. *Dig Dis Sci* 57: 58-64, 2012.
131. Zhou Q, Yang L, Larson S, Basra S, Merwat S, Tan A, Croce C, and Verne GN. Decreased miR-199 augments visceral pain in patients with IBS through translational upregulation of TRPV1. *Gut* 65: 797-805, 2016.

132. Zhou Q, Zhang B, and Verne GN. Intestinal membrane permeability and hypersensitivity in the irritable bowel syndrome. *Pain* 146: 41-46, 2009.
133. Zigmond AS, and Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand* 67: 361-370, 1983

Chapter 3:

Differential Expression of Long non-Coding RNAs in Irritable Bowel Syndrome

ABSTRACT

Chapter 2 of this dissertation reported analysis and results of the mRNAs (protein-coding RNA transcripts) present on the microarray platform used for gene expression profiling in irritable bowel syndrome (IBS) vs. healthy controls (HCs). This microarray platform also included over 30,000 long non-coding RNAs (lncRNAs). While lncRNAs do not encode protein, they can regulate gene expression and thus play an epigenetic role in phenotype expression. The goal of the experiments described in this chapter was to identify differentially expressed lncRNAs in IBS, and to select a lncRNA for further study in intestinal epithelial cell culture models. Based on microarray data, quantitative real-time PCR and *in-situ* hybridization, the lncRNA afadin divergent transcript (*AFDN-DT*) was selected for further study, and the subsequent cell culture experiments are the subject of Chapter 4.

INTRODUCTION

The majority of genes do not encode proteins. Long non-coding RNAs (lncRNAs) are now known to be important regulators of gene expression and biological function,(1) and have been identified as important regulators of multiple diseases. They have diverse molecular functions. The best-known example is as allelic repressors. The lncRNA Xist, for example, regulates X-inactivation by recruitment of silencing factors.(2) Examples of other lncRNAs with known functions are *DACOR1* which up-regulates tumor suppressor genes and is down-regulated in colon cancer,(3) and *Lethe* and *PACER*, which bind subunits of nuclear factor kappa B (NF- κ B) and prevent expression of pro-inflammatory genes.(4) Molecular mechanisms of lncRNAs are diverse and include recruitment of transcriptional machinery, transcriptional silencing through regulation of chromatin structure, stabilization of mRNA, and posttranscriptional gene regulation (protein synthesis, RNA transport).(5) Advances in the field of lncRNAs have led to novel therapeutics based on targeting lncRNAs with antisense oligonucleotides, and several compounds have advanced to clinical trials.(6)

lncRNAs have been associated with multiple gastrointestinal (GI) diseases, such as colon cancer, inflammatory bowel disease, and celiac disease, but have not been extensively studied in irritable bowel syndrome (IBS).(7-9) The hypothesis that lncRNAs are dysregulated in IBS is based on several factors. Firstly, epigenetics is important in the response to stimuli and environment,(10) and onset or symptom exacerbation in IBS is stimulated by a variety of triggers, such as stress, history of early adverse life events, antibiotic use, infection and diet.(11,12) In addition, other epigenetic mechanisms such as methylation/histone modification and microRNAs have been associated with IBS.(13,14)

Sigmoid mucosal expression of lncRNAs was determined by microarray in twenty Rome III+ IBS patients (ten IBS with constipation (IBS-C), ten IBS with diarrhea (IBS-D), five men/women each), and ten age- and sex-matched healthy controls (HCs). This chapter reports the results of

lncRNA profiling as well as the process and experiments used to select a differentially expressed lncRNA for further study in cellular models.

Two approaches were used in the selection of a lncRNA for further study. The first was a data-driven approach. In Chapter 2, a weighted gene co-expression network analysis (WGCNA (15)) was performed to identify clusters (modules) of highly correlated mRNAs, and determine module association with IBS.(16) Since lncRNAs can regulate expression of mRNAs, a lncRNA correlated with the module eigengene (first principal component) of an IBS-associated mRNA module may regulate and thus cause differential expression of these mRNAs, and would therefore be an promising target for further study. Correlation of a pre-filtered list of candidate lncRNAs (based on magnitude of fold change and other features described below) with IBS-associated module eigengenes identified the lncRNA ghrelin opposite strand (*GHRLOS*) as a promising target; however, *in-situ* hybridization localized *GHRLOS* to the lamina propria, and thus it was ultimately not suitable for studies in an epithelial cell culture model.

The second approach to selection of a lncRNA for further study was hypothesis-driven. IBS has been associated with increased intestinal permeability,(17-21) a phenotype linked to abdominal pain.(18,21-23) The lncRNA afadin divergent transcript (*AFDN-DT*) was selected due to differential expression in IBS-C vs. HCs and proximity to afadin (*AFDN*), also differentially expressed in IBS-C vs. HCs, which encodes an adherens junction protein. Since real time quantitative PCR supported microarray findings and *in-situ* hybridization revealed expression of *AFDN-DT* in epithelial cells, *AFDN-DT* was selected for further investigation in cellular models (described in Chapter 4).

METHODS

Participants, protocol and microarray

This chapter focuses on the lncRNAs in the ArrayStar (Rockville, MD) Human LncRNA Microarray V3.0 (includes 30,586 lncRNAs and 26,109 mRNAs). The microarray sample and methods are described in Chapter 2 (page 43).

WGCNA

WGCNA(15) was performed in R (24) on targets with mean raw signal intensity > 200 using a soft-thresholding power 14 and a minimum module size of 30. Highly-correlated modules ($\rho > 0.75$) were merged.

Real-time Quantitative Polymerase Chain Reaction

Total RNA (100 ng) was extracted from the 30 samples analyzed by microarray and eighteen additional samples (identical recruitment and protocol) with TRIzol Reagent (Thermo Fisher Scientific) or with the miRNeasy mini Kit plus on-column DNase treatment (Qiagen, Germantown, MD). RNA was quantified with the NanoDrop spectrophotometer (Thermo Fisher Scientific). Complementary DNA (cDNA) was synthesized from 250-500 ng RNA in a 10- μ L reaction using the iScript™ gDNA Clear cDNA Synthesis Kit (Bio-Rad Laboratories, Hercules, CA) and was diluted 1:2 with nuclease free water. Real-time quantitative PCR (RT-PCR) was performed in 384-well plates using the CFX384 Touch Real-Time PCR Detection System (Bio-Rad). Each 5- μ L reaction contained 1 μ L diluted cDNA (concentration normalized across samples), 2.5 μ L SYBR® Green Supermix (Bio-Rad), and 0.5 μ L each of 10 μ M forward and reverse primers (Integrated DNA Technologies, Coralville, IA). Primer sequences are shown in Table 1. Resulting data were calculated by using the $2^{-[\Delta\Delta Ct]}$ method.

Table 1: Primer sequences

<i>PPIA</i>	F - AGACAAGGTCCCAAAGAC R - ACCACCCTGACACATAAA
<i>18S</i>	F - AGAAACGGCTACCACATCCA R - CACCAGACTTGCCCTCCA
<i>GAPDH</i>	F - ATGTTTCGTCATGGGTGTGAA R - GGTGCTAAGCAGTTGGTGGT
<i>B-Actin</i>	F - CCCAGCACAATGAAGATCAA R - ACATCTGCTGGAAGGTGGAC
<i>AFDN-DT</i>	F - TGAGAGATGCCACGTGATCG R - CACCGGGATGCTCACATAGT
<i>AFDN</i>	F - GAGCATGATGGAGGGTGTCA R - TCCGACTAAGGTCTCGGTCT
<i>ADD1</i>	F - AACCCCTTGCTGCAGTCTCTC R - TAGGGACAGCAGAGGTGGAA
<i>NECTIN</i>	F - TACAACCGGCAGCAGAAGAG R - CTCAGGCACAAGGGAGCTTT
<i>PVR</i>	F - AATGTTCCCGTGAGGTCCTT R - AGCTGAATAGGAGACATGCCC
<i>GHRLOS</i>	F - CTAACACCTGGCGCTGTG R - ATTCAGCCTCCCAGAGCTTT

F, forward; *R*, reverse

***In-Situ* Hybridization**

In-situ hybridization was performed on formalin-fixed paraffin-embedded colonic biopsies using the RNA Scope system (Advanced Cell Diagnostics, Newark, CA) with custom probes designed to target *GHRLOS* and *AFDN-DT*.

Statistical Analysis

All statistics were performed using the R software version 3.3.(24)

Microarray differential expression

Differentially expressed transcripts (DETs) were determined using the Linear Models for Microarray Analysis (limma)(25) package in R version 3.3.1(24) following log2 transformation. An adjusted p value threshold (Benjamini-Hochberg) of 0.05 determined statistical significance. Association of lncRNAs with IBS-associated gene modules was determined Pearson correlation.

RESULTS

Clinical characteristics of the microarray sample are shown in Chapter 2, Table 1 (page 48). The numbers of differentially expressed transcripts (DETs, Total (up-regulated, down-regulated) are shown in Table 2. Similar to the mRNA analysis (Chapter 2), most differences were in IBS-C vs. HC.

Table 2: Differentially expressed lncRNAs in IBS vs. HCs

Group vs. HC	All		Fold Change > 1.5		Fold Change > 2	
	FDR	P	FDR	P	FDR	P
IBS	4 (3,1)	1236 (482,754)	4 (3,1)	252 (103,149)	2 (1,1)	70 (14,56)
IBS-D	0	1022 (718,304)	0	391 (312,79)	0	113 (86,27)
IBS-C	896 (177,719)	2460 (681,1779)	520 (87,433)	747 (144,603)	96 (6,90)	128 (17,111)

Total (up-regulated, down-regulated) lncRNAs for different fold change thresholds. *IBS*, irritable bowel syndrome; *IBS-C*, IBS with constipation; *IBS-D*, IBS with diarrhea

Selection of candidate lncRNAs using a data-driven approach

Stage I selection process

Since very little is known about the function of most lncRNAs, an unbiased approach was used initially to select candidate DETs for further analysis. First, 224 DETs with a fold change > 2 and $p < 0.05$ in any comparison (IBS/IBS-D/IBS-C vs. HC) were identified. From these, we retained 91 DETs that had been assigned an Entrez gene identification number by the National Center for Biotechnology Information (NCBI), as these were more likely to be valid transcripts. For all 91, the following additional evaluation steps were performed: 1) visualize dot plots showing expression by group to determine whether differences were consistent in the groups (i.e. high expression of an up-regulated DET in most IBS samples rather than very high expression in a subset); 2) determine expression profile in colon and colon cancer cell lines through search of “Expression Atlas,” a resource of the European Bioinformatics Institute (EMBL-EBI) which shows expression across tissues and cell lines (26); and 3) literature search. Following this

process, the list of 91 was narrowed to seven candidates which are shown in Table 3. The pseudogene guanylate cyclase 1 soluble subunit beta 2 (*GUCY1B2*) was included in this list, because guanylate cyclase C agonists are effective therapies in IBS-C and chronic idiopathic constipation.(27) There were limited reports of functional pseudogenes at the time, which is why we did not investigate *GUCY1B2* further; however, it should be noted that pseudogenes are now considered to be an important class of lncRNAs with regulatory functions.(28)

Table 3: Fold change (p value) of candidate lncRNAs

Group vs. HC		IBS	IBS-D	IBS-C
GUCY1B2		1.7 (0.046)	2.4 (0.0043)	
MTMR9LP		1.7 (0.044)	2.0 (0.0014)	
SACSAS1		2.0 (0.0058)	2.6 (0.0015)	
LINC01482		1.7 (0.024)	2.5 (0.0014)	
GHRLOS	GHRLOS A	0.5 (0.0014)	0.61 (0.04)	0.41 (0.0004)
	GHRLOS B	0.54 (0.0039)		0.43 (0.001)
BCAR4		-1.5 (0.034)		-2.1 (0.0014)
HYMAI		-1.5 (0.003)		-2.1 (<0.0005)

P values are unadjusted. IBS-C comparisons are significant at FDR<0.05. *IBS*, irritable bowel syndrome; *IBS-C*, IBS with constipation; *IBS-D*, IBS with diarrhea

Stage II selection

Next, we selected a candidate lncRNA based on the potential role of lncRNAs in regulation of mRNA expression. As part of the analysis of microarray mRNA results (Chapter 2), WGCNA was performed to identify groups (“modules”) of highly correlated mRNAs and determine the association of modules with IBS status or symptoms. If a lncRNA regulated expression of mRNAs in these modules, its expression would likely correlate with that module’s eigengene (first principle component), which may predict an important regulatory role in IBS. The WGCNA analysis described in Chapter 2 identified 45 modules and seven of these were associated with IBS status (Chapter 2, Figure 2, page 53). Six of these IBS-associated module eigengenes were significantly correlated with one or more of six out of the seven candidate lncRNAs. A heatmap of these correlations is shown in Figure 1.

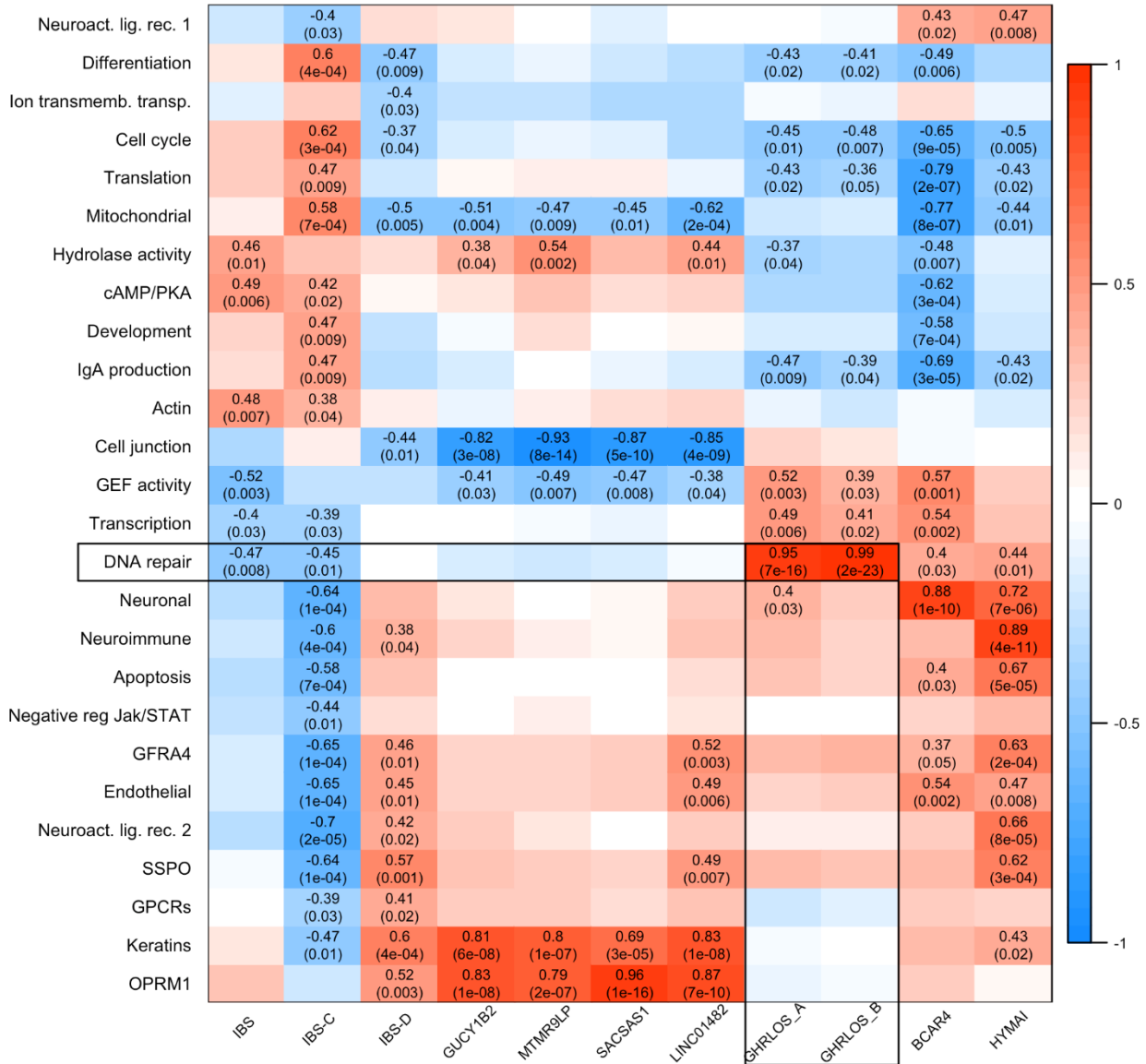


Figure 1: Correlation of candidate lncRNAs with IBS-associated mRNA modules

GHRLOS and the “DNA repair” module

Two probes of ghrelin opposite strand (*GHRLOS*) were down-regulated in IBS (Figure 2, Table 3) and had a strong positive correlation with the “DNA repair” module, which was also down-regulated in IBS (Figure 1). Gene ontology terms overrepresented in the “DNA repair” module are shown in Table 4. Both microarray probes matched the same 5 transcript variants (Figure 3).

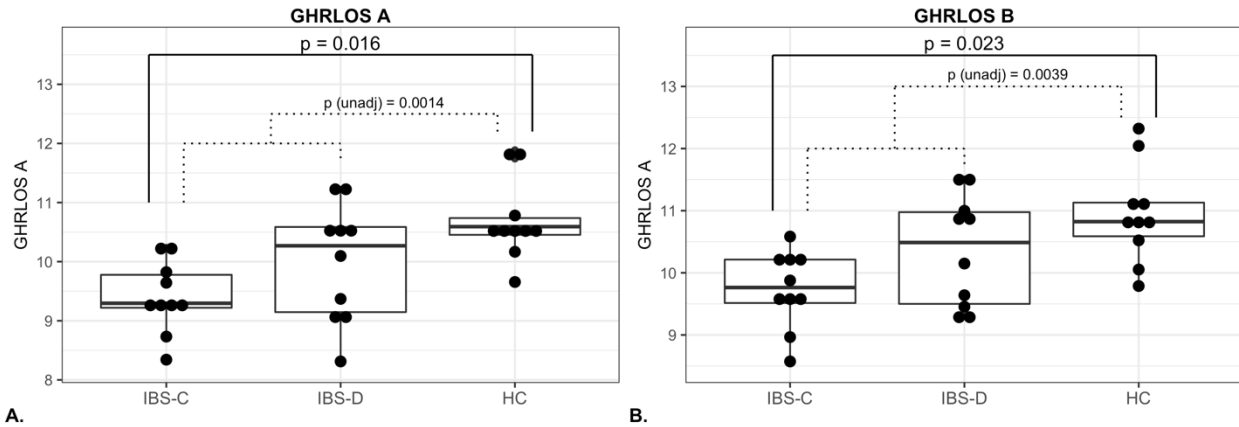


Figure 2: *GHRLOS* is down-regulated in IBS

Table 4: Top ten gene ontology terms associated with the “DNA repair” module

Term	p value
Double-strand break repair via nonhomologous end joining	0.001
MyD88-dependent toll-like receptor signaling pathway	0.005
Interstrand cross-link repair	0.008
Internal peptidyl-lysine acetylation	0.008
Positive regulation of DNA repair	0.009
Histone modification	0.009
Ossification	0.009
Muscle system process	0.010
Modification-dependent protein catabolic process	0.010
Regulation of canonical Wnt signaling pathway	0.010

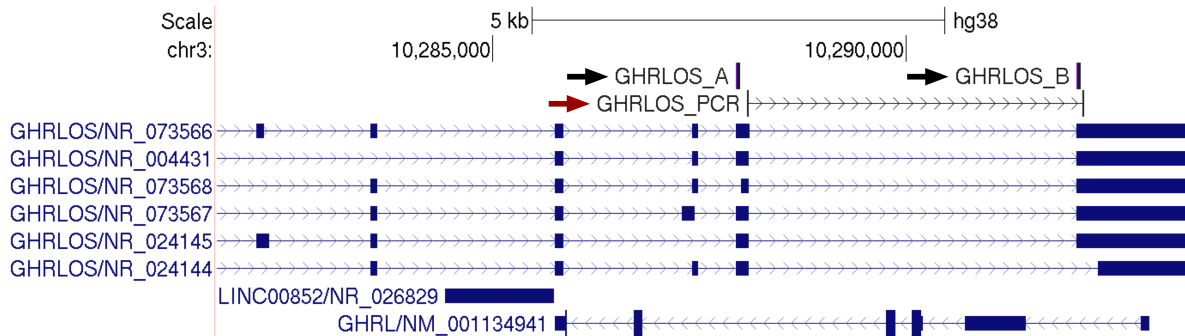


Figure 3: *GHRLOS* microarray probe and PCR primer alignments

BLAT(29) alignment of microarray probes (Black arrows, *GHRLOS A*, *GHRLOS B*) and forward and reverse PCR primers (red arrow). Image created with the UCSC genome browser, <http://genome.ucsc.edu> (30) using the Human Dec. 2013 (GRCh38/hg38) assembly.

Table 5: “DNA repair” genes correlated with *GHRLOS* and IBS-C

Symbol, Name	<i>GHRLOS</i> Correlation	Fold Change	
		IBS-C vs. HC	IBS vs. HC*
<i>APCDD1</i> , APC down-regulated 1	0.72 (0.002)	0.69 (0.007)	0.81 (0.023)
<i>KCNMB3</i> , potassium calcium-activated channel subfamily M regulatory beta subunit 3	0.94 (<0.005)	0.59 (0.010)	0.67 (0.001)
<i>TMCO6</i> , Transmembrane and coiled-coil domains 6	-0.68 (0.010)	1.52 (0.010)	1.33 (0.005)
<i>MLN</i> , motilin	-0.64 (0.026)	1.49 (0.012)	1.38 (0.001)
<i>ROBO2</i> , roundabout guidance receptor 2	0.71 (0.003)	0.66 (0.014)	0.74 (0.003)
<i>FAM153A</i> , family with sequence similarity 153 member A	0.69 (0.006)	0.78 (0.016)	0.86 (0.022)
<i>ZNF45</i> , zinc finger protein 45	0.97 (<0.005)	0.34 (0.018)	0.49 (0.011)
<i>TEK</i> , TEK receptor tyrosine kinase	0.96 (<0.005)	0.30 (0.021)	0.43 (0.008)
<i>PDE7A</i> , phosphodiesterase 7A	0.98 (<0.005)	0.36 (0.021)	0.50 (0.010)
<i>CALCA</i> , calcitonin related polypeptide alpha	0.96 (<0.005)	0.27 (0.023)	0.40 (0.008)

*unadjusted p value. P values for *GHRLOS* correlation and fold change in IBS-C vs. HCs are adjusted for multiple comparisons (Benjamini-Hochberg)

Table 5 shows the top ten genes in the “DNA repair” module based on differential expression in IBS-C vs. HC as well as the correlation of those genes with *GHRLOS*. Interestingly, these genes included motilin (*MLN*), which was up-regulated in IBS-C (fold change = 1.49, $p = 0.012$) and negatively correlated with *GHRLOS* (Figure 4, $r = -0.064$, $p = 0.026$). Dysregulation of motilin in IBS is supported by other studies,(31-33) and motilin receptor agonists, such as erythromycin, accelerate GI transit. In addition to GO terms shown in Table 4, “positive regulation of ERK1 and ERK2 cascade” (GO:0070374, $p=0.003$) and “positive regulation of NF- κ B transcription factor activity” (GO:0051092, $p=0.0016$) were also enriched in the “DNA repair” module. Extracellular signal-regulated kinases (ERK1/ERK2) are downstream effectors of G protein-coupled receptors (GPCRs), such as the motilin receptor, and binding of the closely related hormone ghrelin to its receptor activates NF- κ B.(34)

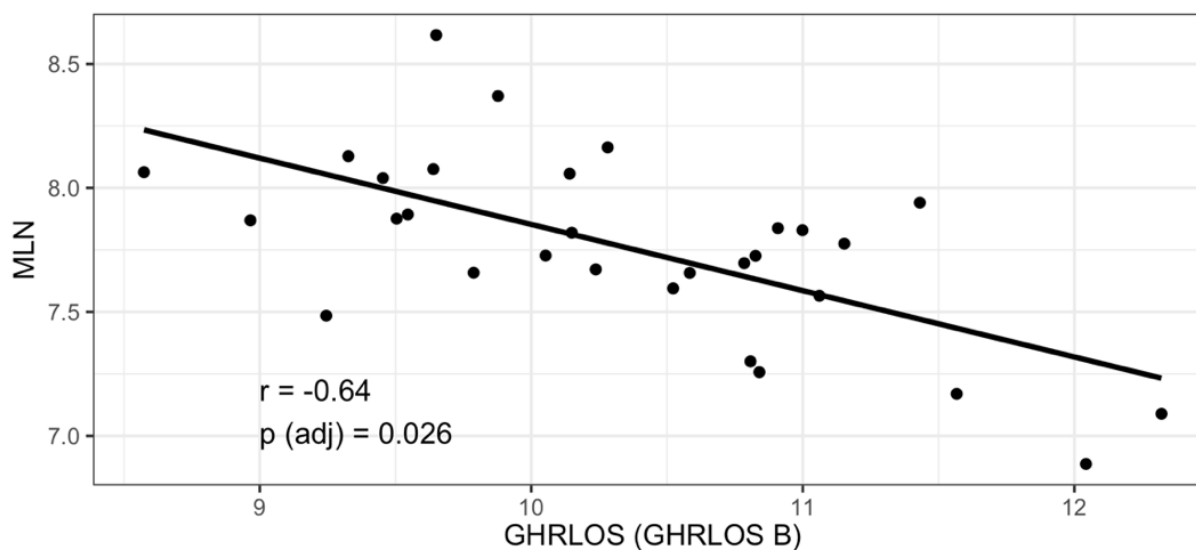


Figure 4: *GHRLOS* and *MLN* are negatively correlated

PCR validation of *GHRLOS*

Microarray findings were validated with RT-PCR. Primers (Table 1, Figure 3) were designed to amplify the region of *GHRLOS* targeted by both microarray probes (*GHRLOS A*, *GHRLOS B*).

Characteristics of the PCR cohort are shown in **Error! Not a valid bookmark self-reference..**

Results of the PCR are shown in Figure 5. There was a trend for decreased expression in IBS-C vs. HC ($p = 0.09$).

Table 6: *GHRLOS* PCR sample demographics

		IBS	HC
N (% women)		38 (53%)	16 (44%)
Cohort	Microarray	15 (9 IBS-C, 6 IBS-D)	10
	Validation	23 (10 IBS-C, 9 IBS-D, 2 IBS-M, 1 IBS-U)	6
Age		33.9 (12.5)	33.1 (9.2)
BMI		25 (5)	25.5 (3.4)
Hispanic		3 (7.9%)	4 (25%)
Asian		8 (21.1%)	3 (18.8%)
Black		4 (10.5%)	1 (6.2%)
White		20 (52.6%)	8 (50%)
Other		6 (15.8%)	3 (18.8%)

BMI, body mass index; *IBS-M*, IBS with mixed bowel habit; *IBS-U*, IBS unsubtyped.

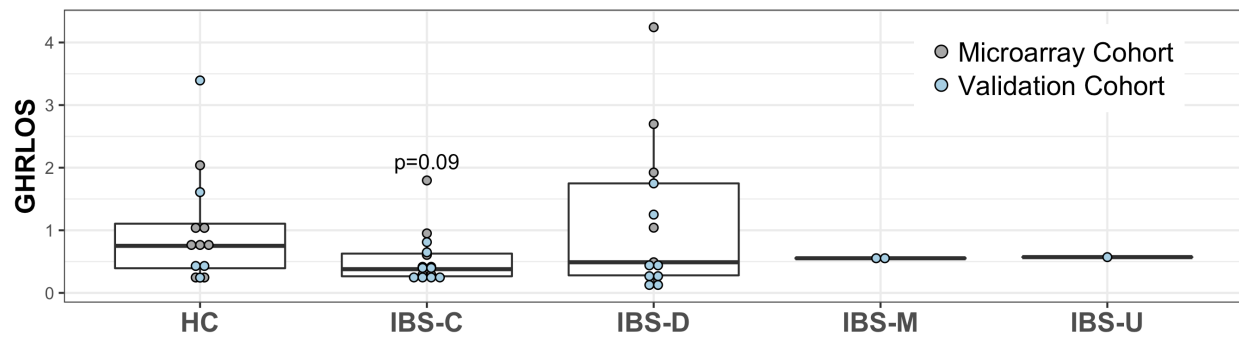


Figure 5: *GHRLOS* PCR
RT-PCR of *GHRLOS* normalized to 18S and PPIA

In-situ hybridization of *GHRLOS*

In order to verify expression of *GHRLOS* in colonic mucosal biopsies, *in-situ* hybridization was performed on formalin-fixed paraffin embedded colonic biopsies using a custom probe designed to target *GHRLOS*. *GHRLOS* was predominantly expressed in the lamina propria (Figure 6).

Because the goal of this project was to test a lncRNA in colonic epithelial cell culture models, no further investigation of *GHRLOS* was performed.

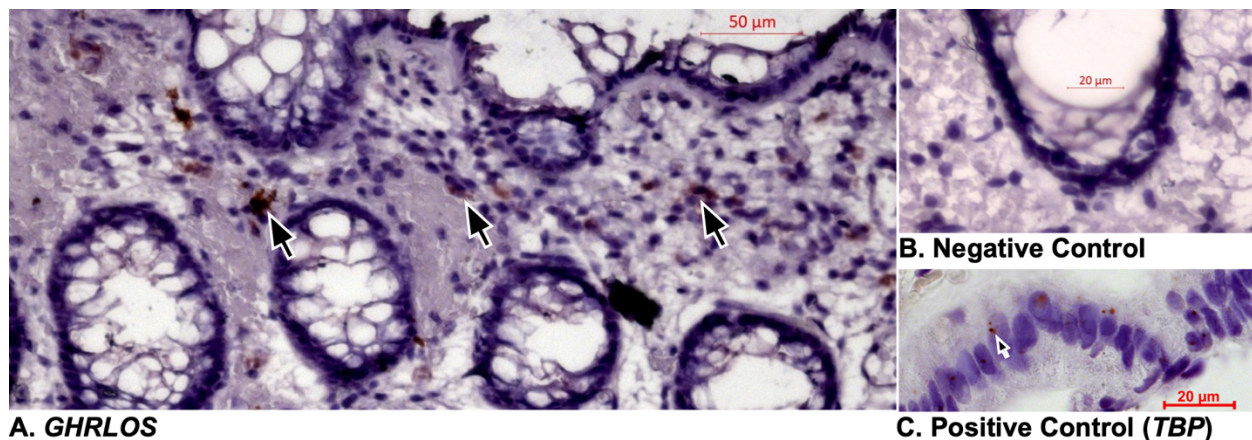


Figure 6: *GHRLOS* gene expression by *in-situ* hybridization

Selection of candidate lncRNAs using a hypothesis-based approach

Identification of *GHRLOS* as a candidate lncRNA for further study used an unbiased approach. Subsequently, a hypothesis-based approach was used, based on evidence indicating that impaired intestinal permeability is a consistent finding in IBS.(35) We therefore hypothesized that differential expression of lncRNAs may underlie alterations in intestinal permeability in IBS.

Differential expression of adherens junction genes in IBS-C vs. HCs

In Chapter 2 (16), differentially expressed mRNAs from this study were compared with publicly available microarray data from two additional cohorts (Mayo Clinic (36) and the University of Nottingham (37)). One of the mRNAs that was differentially expressed in our study as well as in the sigmoid biopsies from the Mayo Clinic was nectin cell adhesion molecule 1 (*NECTIN1*), which encodes an adherens junction protein. Formation of adherens junctions is an early step in the cellular adhesion process, and it is necessary for proper formation of tight junctions.(38) In addition to *NECTIN1*, genes encoding the adherens junction proteins poliovirus receptor (*PVR*), and afadin (*AFDN*) as well as the actin-capping protein adducin 1 (*ADD1*), were also differentially expressed in IBS-C vs. HCs (FDR<0.05, Table 7). The role of the adherens junction and these molecules specifically in the formation of cellular junctions is outlined in Figure 7.

Microarray expression of *NECTIN1*, *ADD1*, *PVR* and *AFDN* is shown in Figure 8A-D.

Expression measured by RT-PCR in a larger cohort (Table 8) is shown in Figure 8E-H.

Microarray findings were not replicated overall, though the direction of fold change in IBS vs. HCs was maintained for *PVR* (down) and *AFDN* (up). Despite these results, based on findings in the literature, selection of a target involved in cell adhesion remained a primary focus.

Table 7: Differentially expressed adherens junction genes

	IBS-C vs. HCs			IBS-D vs. HCs			IBS vs. HCs		
Symbol	FC	P Value	Adj. P Value	FC	P Value	Adj. P Value	FC	P Value	Adj. P Value
<i>NECTIN1</i>	0.63	0.0001	0.01	1.1	0.36	0.8	0.83	0.056	0.483
<i>ADD1</i>	0.71	0.0001	0.009	1.11	0.18	0.68	0.89	0.087	0.54
<i>PVR</i>	0.8	0.0028	0.04	1.03	0.63	0.9	0.91	0.122	0.581
<i>AFDN</i>	1.51	0.0002	0.011	1.09	0.37	0.8	1.28	0.005	0.282

FC, fold change; Adj., adjusted

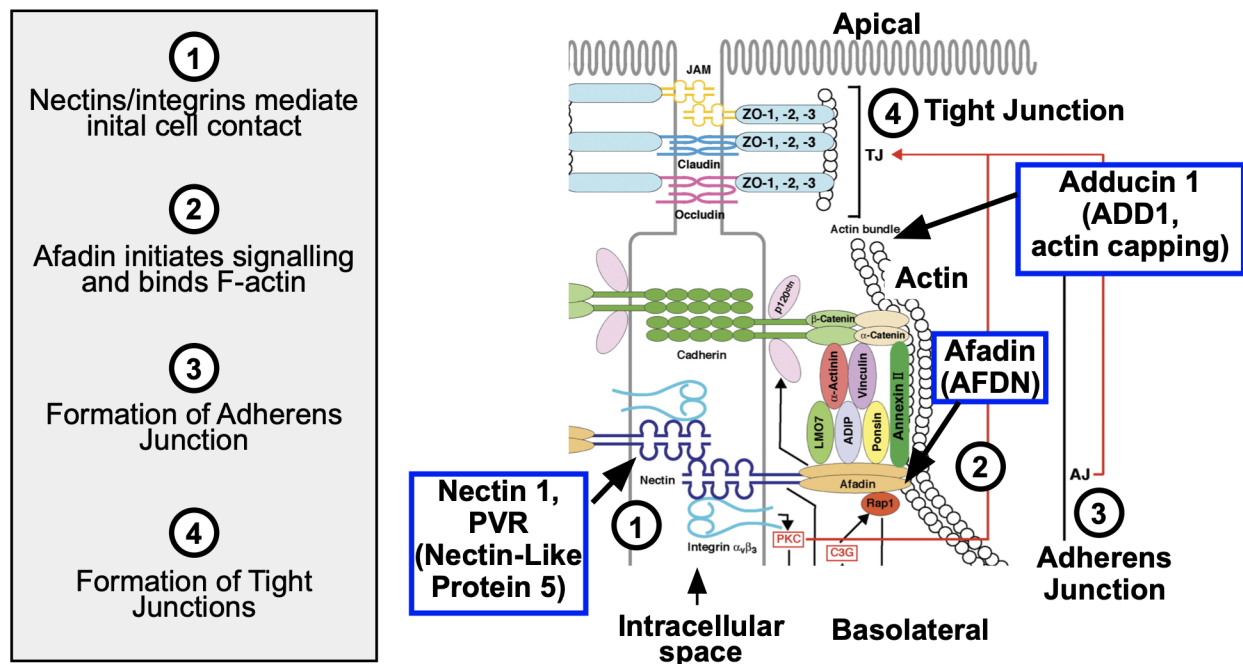


Figure 7: Cellular junction formation

Modified from Current Opinion in Cell Biology, Vol. 19, Issue 5, Sakisaka et al, (39), *The roles of nectins in cell adhesions: cooperation with other cell adhesion molecules and growth factor receptors*, pages 593-602; Copyright 2007, with permission from Elsevier.

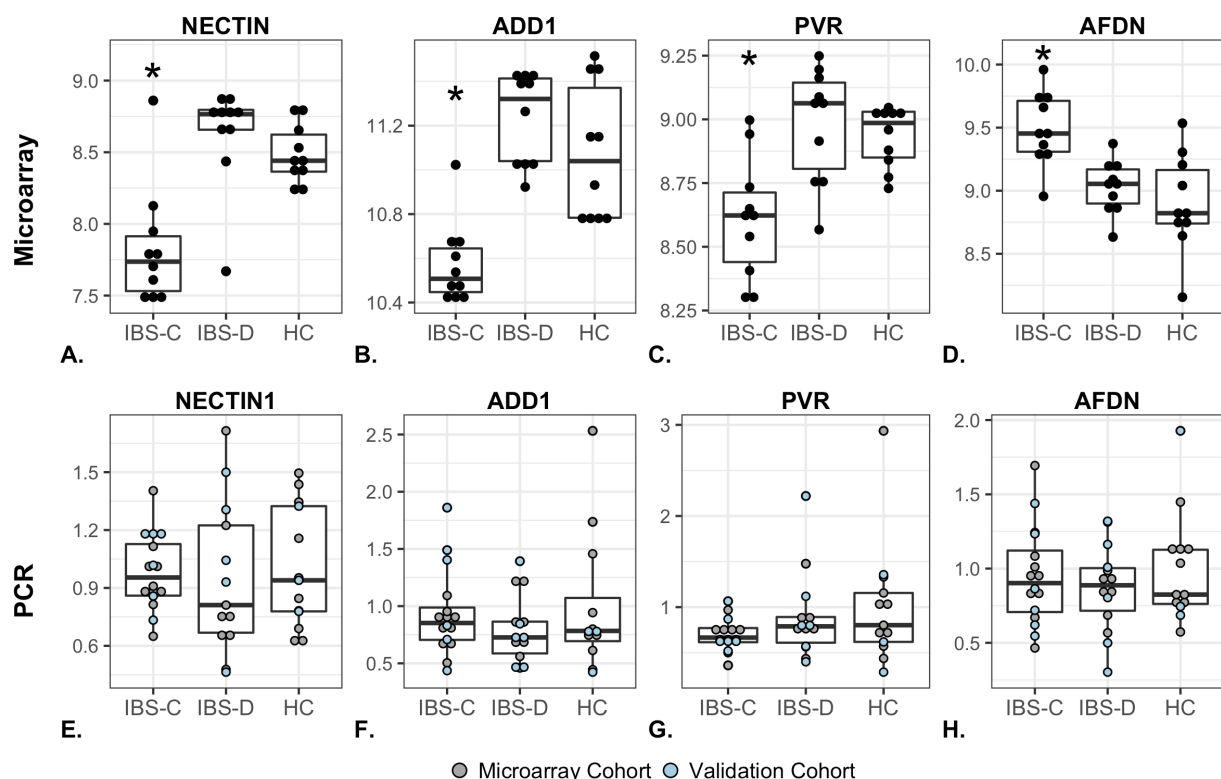


Figure 8: Expression of adherens junction genes by microarray and RT-PCR
PCR expression is normalized to *ACTB* and *GAPDH*

Table 8: Adherens junction genes PCR sample demographics

		IBS	HC
N (% women)		30 (50%)	13 (46%)
Cohort	Microarray	18 (10 IBS-C, 8 IBS-D)	10
	Validation	12 (6 IBS-C, 10 IBS-D)	3
Age		34.9 (12.2)	32.2 (7.8)
BMI		25 (5.5)	25.5 (2.9)
Hispanic		4 (13.3%)	2 (15.4%)
Asian		6 (20.0%)	3 (23.1%)
Black		4 (13.3%)	1 (7.7%)
White		15 (50%)	6 (46.2%)
Other		5 (16.7%)	3 (23.1%)

BMI, body mass index

Afadin divergent transcript (*AFDN-DT*) is up-regulated in IBS-C vs. HCs

A search of lncRNAs in close proximity to the cell adhesion genes in

Table 7 identified *AFDN* divergent transcript (*AFDN-DT*). *AFDN-DT* (NCBI gene ID 653483) has two exons and encodes a 2,311-base-pair lncRNA (NR 027906). *AFDN-DT* is in a head-to-head relationship with *AFDN*, which means that there is overlap of the 5' regions (Figure 9). In the microarray data, *AFDN-DT* was up-regulated in IBS-C vs. HCs with an FDR of 0.061 and in IBS overall vs. HCs with an unadjusted p value of 0.013 (Figure 10A). It was also positively correlated with *AFDN* (Figure 10B).

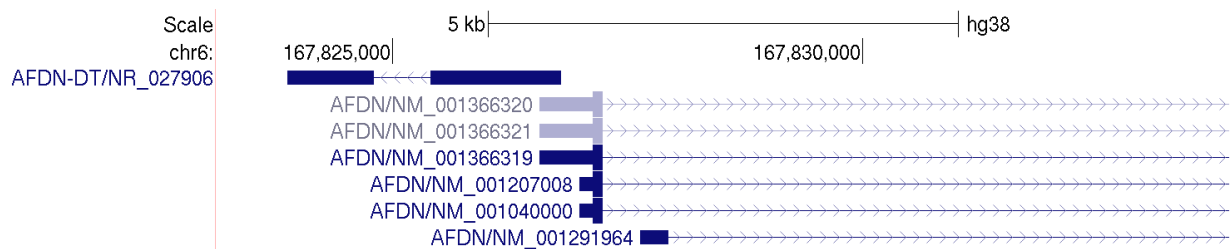


Figure 9: Afadin divergent transcript (*AFDN-DT*)

Image created with the UCSC genome browser, <http://genome.ucsc.edu> (30) using the Human Dec. 2013 (GRCh38/hg38) assembly.

Validation of *AFDN-DT* by RT-PCR

RT-PCR was performed in two larger overlapping cohorts (n=55, n=63). Sample demographics are shown in Table 9. PCR results are shown in Figure 11. Mean expression of *AFDN-DT* was greater in IBS-C vs. HCs but the difference was not statistically significant. There was a trend for increased expression in IBS overall vs. HCs ($p = 0.08$, $p = 0.07$ for PCR 1 and 2). There was a significant ($p < 0.005$) positive correlation between *AFDN-DT* and *AFDN* (Figure 12).

Table 9: *AFDN-DT* PCR cohort demographics

		PCR 1 (n=56)		PCR 2 (n=65)	
		IBS	HC	IBS	HC
N (% women)		39 (54%)	17 (47%)	45 (62%)	20 (45%)
Cohort	Microarray	15 (9 IBS-C, 6 IBS-D)	10	18 (10 IBS-C, 8 IBS-D)	10
	Validation	23 (11 IBS-C, 9 IBS-D, 2 IBS-M, 1 IBS-U)	7	25 (12 IBS-C, 10 IBS-D, 2 IBS-M, 1 IBS-U)	10
Age		34.2 (12.5)	32.4 (9.3)	34.6 (12.3)	33.8 (10.4)
BMI		25 (4.9)	25.7 (3.3)	25.1 (5.4)	25.7 (3.6)
Hispanic		3 (7.7%)	4 (23.5%)	5 (11.1%)	4 (20%)
Asian		8 (20.5%)	4 (23.5%)	10 (22.2%)	4 (20%)
Black		5 (12.8%)	1 (5.9%)	8 (17.8%)	2 (10%)
White		20 (51.3%)	8 (47.1%)	22 (48.9%)	10 (50%)
Other		6 (15.4%)	3 (17.6%)	5 (11.1%)	3 (15%)

BMI, body mass index; IBS-M, IBS with mixed bowel habit; IBS-U, IBS unsubtyped

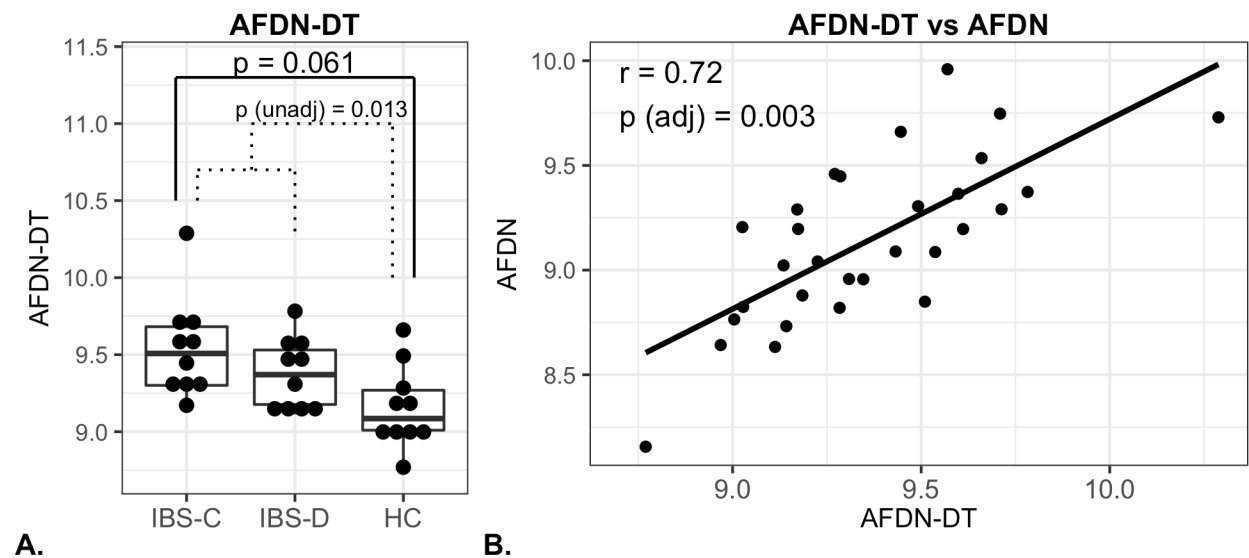


Figure 10: *AFDN-DT* microarray differential expression and correlation with *AFDN*

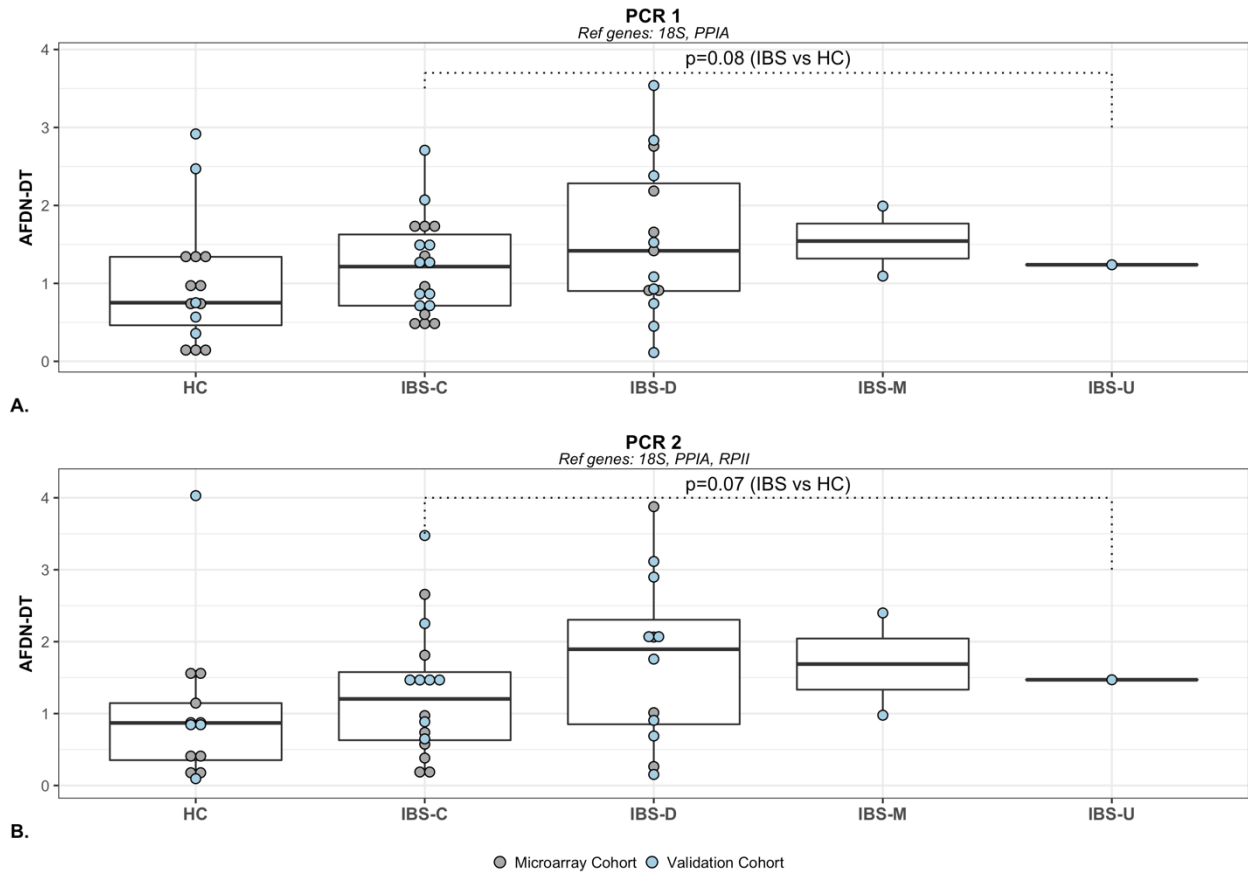


Figure 11: PCR validation of *AFDN-DT*

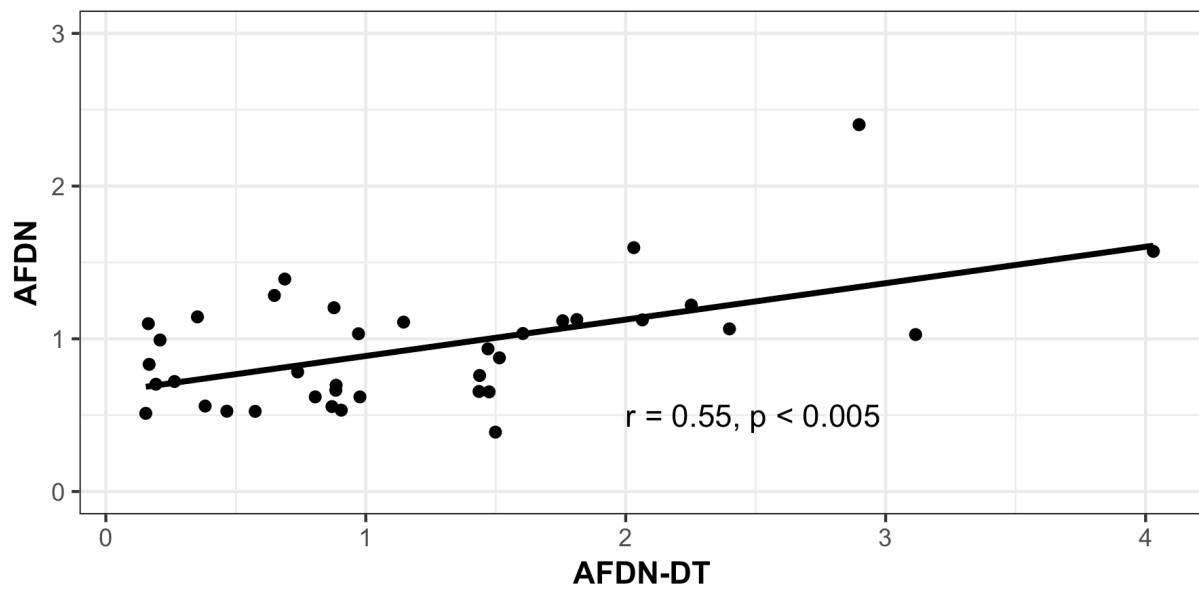


Figure 12: PCR correlation of *AFDN-DT* and *AFDN*
Results are from PCR 2 with normalization to *18S*, *PPIA* and *RPII*.

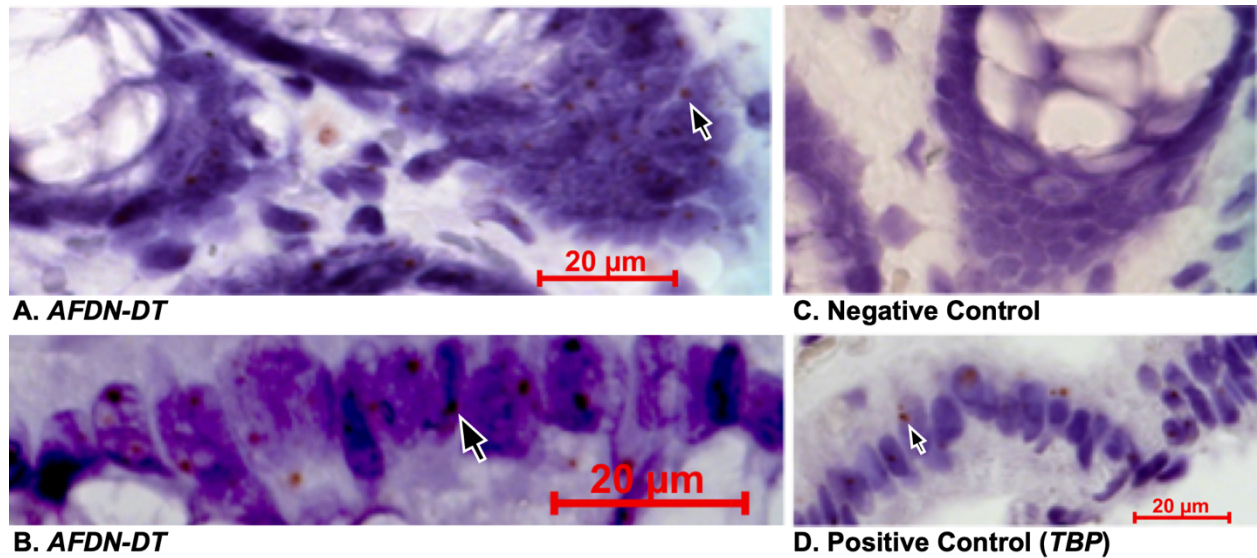


Figure 13: *AFDN-DT* expression by *in situ* hybridization

In-situ hybridization of *AFDN-DT*

In order to verify the expression of *AFDN-DT* in colonic mucosal biopsies and to determine whether expression was present in epithelial cells, *in-situ* hybridization was performed on formalin-fixed paraffin embedded colonic biopsies using a custom probe designed to target *AFDN-DT*. *AFDN-DT* was expressed in epithelial cells (Figure 13A, arrows indicate positive signal).

AFDN-DT selected for further study in cellular models

In-situ hybridization confirmed expression of *AFDN-DT* in colonic epithelial cells. In Chapter 4, cell culture models are used to test the hypothesis that *AFDN-DT* overexpression affects the permeability of colonic epithelial monolayers. This specific hypothesis addresses the overall hypothesis that dysregulation of *AFDN-DT* may impact IBS symptoms and pathophysiology by affecting intestinal permeability.

DISCUSSION

Dysregulation of lncRNAs in IBS

Similar to the mRNA findings discussed in Chapter 2, differential expression of lncRNAs was primarily seen in IBS-C vs. HCs. As discussed in Chapter 2, the presence of positive findings in IBS-C vs. HCs may be due to careful age- and sex-matching of the samples and exclusion of those with medication use. Absence of significant differences in IBS overall and IBS-D vs. HCs should be interpreted with caution as the microarray study was only powered to make negative conclusions for genes with fold change > 2. Absence of smaller, but potentially biologically meaningful differences cannot be definitively concluded.

GHRLOS

GHRLOS was identified as a promising target for further study through a data-driven analytic approach to differentially expressed lncRNAs in IBS. Several transcripts of *GHRLOS*, the natural antisense transcript (NAT) of the ghrelin gene (*GHRL*), have been identified.(40) The first exon of the majority of transcripts overlaps 1.4 kb of the *GHRL* promoter and 106 nucleotides of *GHRL* exon 4. Evidence that *GHRLOS* is a functional lncRNA includes polyadenylation, 5' capping, extensive splicing and tissue-specific expression.(41) Expression is low in the stomach and highest in tissues known to have high expression of lncRNAs, such as the brain and thymus.(41) An association between decreased expression of *GHRLOS* and poor prognosis in colon cancer has been reported.(42)

GHRLOS is down-regulated in IBS

In this study, *GHRLOS* was down-regulated in sigmoid mucosal biopsies of twenty patients with IBS (ten with constipation, ten with diarrhea) in comparison to ten HCs (Figure 2). PCR validation was supportive ($p = 0.09$, Figure 5). A weighted gene correlation network analysis (WGCNA)(15) identified a strong positive association of *GHRLOS* with genes related to activation of NF- κ B and ERK1/ERK2, which are both downstream targets of ghrelin signaling,

which parallels motilin signaling.(34,43) Expression of *GHRLOS* was also negatively associated with motilin (*MLN*), which was up-regulated in IBS (fold change 1.37, $p = 0.0006$). Dysregulation of motilin in IBS is supported in other studies,(31,33,44) and motilin receptor agonists, such as erythromycin accelerate GI transit.

Ghrelin and motilin in gastrointestinal physiology

Ghrelin and motilin are closely related hormones released from the GI tract. They are released by similar stimuli, such as fasting, but have different physiologic effects. Motilin is secreted in association with phase III of the migrating motor complex (MMC).(45) Motilin's prokinetic effect is hypothesized to occur via facilitation of myenteric cholinergic activity at low concentrations and direct action on smooth muscle at high concentrations.(46) Motilin agonists are more potent prokinetics than 5HT4 agonists;(47) however, few of the agents studied in IBS clinical trials have succeeded in meeting clinical endpoints.(45,48) Additionally, despite observations that motilin increases contractility of human stomach or colon strips,(49,50) the non-peptide motilin receptor agonist erythromycin did not increase colonic motility or transit in the majority of studies in healthy individuals or those with chronic constipation.(51-53) Broad et al. proposed that this discrepancy is due to the longer duration of action with non-peptides, which prevents regulation of MMCs.(49)

Co-regulation of *GHRLOS* and motilin may involve NF- κ B activation

The receptors for ghrelin and motilin are the growth hormone secretagogue receptor (GHSR) and the motilin receptor (MLNR) and they are both GPCRs. Ghrelin receptors are widely expressed outside the GI tract, and signaling can occur in autocrine or paracrine fashions, such as in adipocytes,(54) or activation of vagal afferents.(55) NF- κ B is a known downstream target of GPCRs and ghrelin signaling,(56,57) and NF- κ B activation was a GO term associated with the "DNA repair" module which contains motilin and was positively correlated with *GHRLOS* expression. Though data from the cell lines studied in the ENCODE project (58) may have

limited relevance to regulation of colonic gene expression, the presence of chromatin immunoprecipitation sequencing (ChIP-Seq) peaks in the *GHRLOS* promoter region for MAX and JUND, two transcription factors activation by ERK1/ERK2 signaling (downstream of GPCRs), provides support for the hypothesis that *GHRLOS* is regulated by GPCR-signaling. In addition, there is a NF- κ B binding site in intron 2 of *GHRLOS*, and intronic NF- κ B binding has been shown to induce gene expression.(59) Since signaling down-stream of GPCRs may increase *GHRLOS* expression, it can be hypothesized that *GHRLOS* could function as a feedback mechanism to regulate ligand expression and thereby influence autocrine signaling. This hypothetical mechanism is outlined in (Figure 14). While there was no significant correlation of *GHRLOS* and ghrelin (*GHRL*) RNA, correlations due to this type of mechanism might only be seen in single cell sequencing or cell culture experiments.

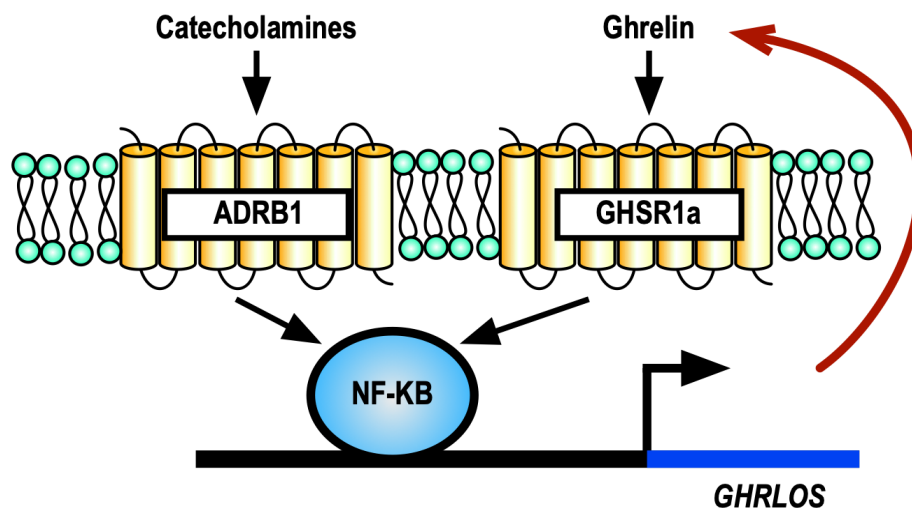


Figure 14: Hypothesized role of *GHRLOS*

If *GHRLOS* is regulated by signaling down-stream of GPCRs (NF- κ B), perhaps it has a role in feedback regulation of ghrelin autocrine signaling

AFDN-DT

AFDN-DT was selected for further investigation due to a putative role in epithelial permeability. It was up-regulated in IBS-C vs. HCs (Figure 10) and its expression in epithelial cells was confirmed by *in-situ* hybridization (Figure 13). Down-regulation of *AFDN-DT* (also known as *MLLT4* antisense 1) is associated with poor prognosis in gastric cancer.(60)

Mechanisms of antisense lncRNA regulation of mRNA

Antisense transcripts play functional regulatory roles through a variety of mechanisms, which are outlined in Figure 15.(61,62) Some antisense transcripts result in epigenetic silencing of neighboring genes, such as *KCNQ1* overlapping transcript 1 and the Air non-coding RNA, which both function through interactions with chromatin and histone methyltransferases.(63,64) Antisense transcripts can also act as positive regulators of neighboring genes. For example, *TALAM1*, positively regulates *MALAT1* through stabilization by 3' end processing.(65)

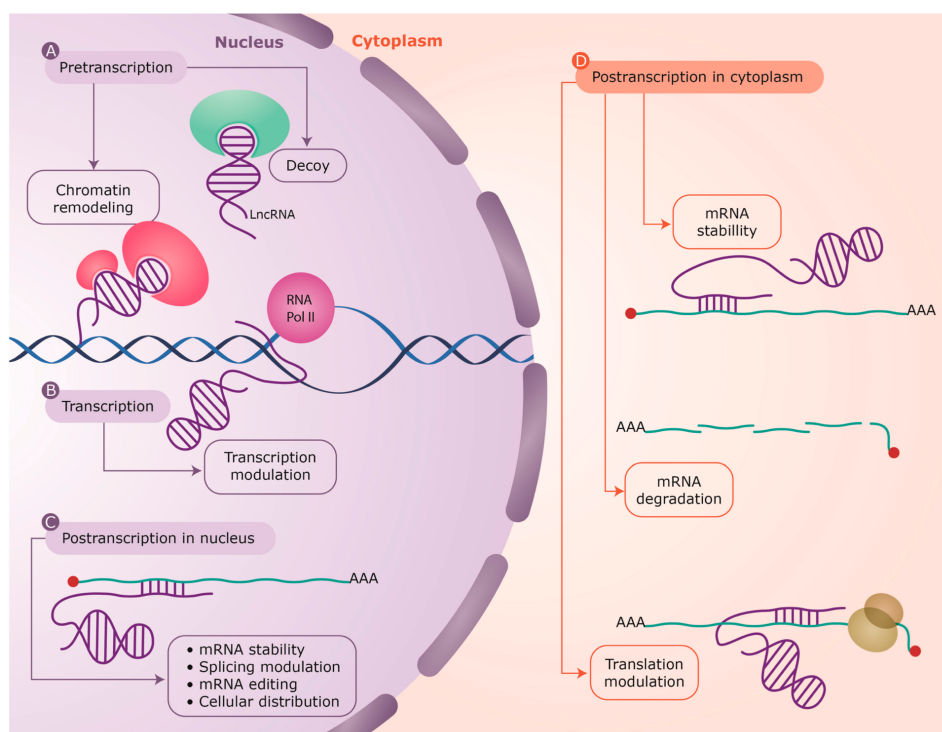


Figure 15: Mechanisms of antisense transcripts

Reprinted from Villegas VE and Zaphiropoulos PG, Neighboring gene regulation by antisense long non-coding RNAs. Int J Mol Sci, 2015. 16(2): 3251-66. Licensed under [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/).

Impaired intestinal permeability in IBS

Increased intestinal permeability is frequently seen in IBS,(17,20-23,66-72) and has been shown to correlate with abdominal pain.(21-23,71) The mechanism underlying this observation remains unknown; however, several groups have identified alterations in expression of tight junction proteins such as ZO-1 and occludin.(21,23,73) Decreased expression of the adherens junction protein E-cadherin has also been observed in IBS.(74)

Notably, our findings were greatest in IBS-C vs. HCs. While many of the studies on permeability in IBS have been conducted in IBS-D patients, there are several studies which have shown increased permeability in IBS-C (21,23,66,75) and others have shown increased permeability in IBS vs. HCs overall (including IBS-C).(19,67) Increased permeability has also been associated with chronic constipation.(76)

Adherens junctions in epithelial barrier formation

The epithelial barrier is maintained by junctional complexes which include tight junctions and adherens junctions (Figure 7).(39) The gene encoding the adaptor protein AFDN, the sense counterpart to *AFDN-DT*, plays a crucial role in formation and maintenance of adherens junctions through interactions with nectins and cadherins.(77) Adherens junctions initiate signaling resulting in organization of the actin cytoskeleton and recruitment of tight junction proteins. Interestingly, the phenotype of mice lacking intestinal expression of afadin is reminiscent of IBS with a morphologically normal gut, but with increased permeability.(78)

A potential role for *AFDN-DT* and *AFDN* in intestinal permeability

Based on the differential expression of both *AFDN-DT* and *AFDN* in IBS as well as the positive correlation between *AFDN-DT* and *AFDN*, we hypothesize that *AFDN-DT* is a positive regulator of *AFDN*, and that *AFDN-DT* expression impacts epithelial permeability. While the direction of change (up-regulation) is opposite of what might be expected based on the animal model of intestinal afadin knockout (increased permeability with afadin knockout), mRNA and protein

expression may not be directly correlated. It is also possible that the change in expression is a response to impaired permeability, and increased expression of *AFDN-DT* and *AFDN* may be related to restoration of barrier function. In the next Chapter of this dissertation, the hypothesis that *AFDN-DT* can affect epithelial permeability is tested in a cell culture model using colonic epithelial cells with inducible overexpression of *AFDN-DT*. Barrier impairment following overexpression of *AFDN-DT* would support a role for *AFDN-DT* in IBS pathogenesis. The possibility that *AFDN-DT* contributes to the maintenance of barrier function will also be tested as though overexpression during treatment of monolayers with pro-inflammatory cytokines known to increase permeability.

REFERENCES:

1. Fu XD. Non-coding RNA: a new frontier in regulatory biology. *Natl Sci Rev.* 2014;1(2):190-204.
2. Brown CJ, Hendrich BD, Rupert JL, Lafreniere RG, Xing Y, Lawrence J, Willard HF. The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. *Cell.* 1992;71(3):527-542.
3. Merry CR, Forrest ME, Sabers JN, Beard L, Gao XH, Hatzoglou M, Jackson MW, Wang Z, Markowitz SD, Khalil AM. DNMT1-associated long non-coding RNAs regulate global gene expression and DNA methylation in colon cancer. *Hum Mol Genet.* 2015;24(21):6240-6253.
4. Elling R, Chan J, Fitzgerald KA. Emerging role of long noncoding RNAs as regulators of innate immune cell development and inflammatory gene expression. *Eur J Immunol.* 2016;46(3):504-512.
5. Wang KC, Chang HY. Molecular mechanisms of long noncoding RNAs. *Mol Cell.* 2011;43(6):904-914.
6. Parasramka MA, Maji S, Matsuda A, Yan IK, Patel T. Long non-coding RNAs as novel targets for therapy in hepatocellular carcinoma. *Pharmacol Ther.* 2016;161:67-78.
7. Castellanos-Rubio A, Fernandez-Jimenez N, Kratchmarov R, Luo X, Bhagat G, Green PH, Schneider R, Kiledjian M, Bilbao JR, Ghosh S. A long noncoding RNA associated with susceptibility to celiac disease. *Science.* 2016;352(6281):91-95.
8. Ma Y, Yang Y, Wang F, Moyer MP, Wei Q, Zhang P, Yang Z, Liu W, Zhang H, Chen N, Wang H, Wang H, Qin H. Long non-coding RNA CCAL regulates colorectal cancer progression by activating Wnt/beta-catenin signalling pathway via suppression of activator protein 2alpha. *Gut.* 2015, 10.1136/gutjnl-2014-308392.
9. Padua D, Mahurkar-Joshi S, Law IK, Polytarchou C, Vu JP, Pisegna JR, Shih D, Iliopoulos D, Pothoulakis C. A long noncoding RNA signature for ulcerative colitis identifies IFNG-AS1 as an enhancer of inflammation. *Am J Physiol Gastrointest Liver Physiol.* 2016;311(3):G446-457.
10. Mohanty V, Gokmen-Polar Y, Badve S, Janga SC. Role of lncRNAs in health and disease-size and shape matter. *Brief Funct Genomics.* 2015;14(2):115-129.
11. Chey WD, Kurlander J, Eswaran S. Irritable bowel syndrome: a clinical review. *JAMA.* 2015;313(9):949-958.
12. Park SH, Videlock EJ, Shih W, Presson AP, Mayer EA, Chang L. Adverse childhood experiences are associated with irritable bowel syndrome and gastrointestinal symptom severity. *Neurogastroenterol Motil.* 2016;28(8):1252-1260.
13. Mahurkar S, Polytarchou C, Iliopoulos D, Pothoulakis C, Mayer EA, Chang L. Genome-wide DNA methylation profiling of peripheral blood mononuclear cells in irritable bowel syndrome. *Neurogastroenterol Motil.* 2016;28(3):410-422.

14. Mahurkar-Joshi S, Videlock EJ, Iliopoulos D, Pothoulakis C, Mayer EA, Chang L. Epigenetic Changes in Blood Cells and Colonic Mucosa are Associated with Irritable Bowel Syndrome (IBS). *Gastroenterology*: Elsevier; 2018. p. S-214.
15. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008;9:559.
16. Videlock EJ, Mahurkar-Joshi S, Hoffman JM, Iliopoulos D, Pothoulakis C, Mayer EA, Chang L. Sigmoid colon mucosal gene expression supports alterations of neuronal signaling in irritable bowel syndrome with constipation. *Am J Physiol Gastrointest Liver Physiol*. 2018;315(1):G140-G157.
17. Dunlop SP, Hebden J, Campbell E, Naesdal J, Olbe L, Perkins AC, Spiller RC. Abnormal intestinal permeability in subgroups of diarrhea-predominant irritable bowel syndromes. *Am J Gastroenterol*. 2006;101(6):1288-1294.
18. Zhou Q, Zhang B, Verne GN. Intestinal membrane permeability and hypersensitivity in the irritable bowel syndrome. *Pain*. 2009;146(1-2):41-46.
19. Marshall JK, Thabane M, Garg AX, Clark W, Meddings J, Collins SM, Investigators WEL. Intestinal permeability in patients with irritable bowel syndrome after a waterborne outbreak of acute gastroenteritis in Walkerton, Ontario. *Aliment Pharmacol Ther*. 2004;20(11-12):1317-1322.
20. Gecse K, Roka R, Sera T, Rosztoczy A, Annahazi A, Izbeki F, Nagy F, Molnar T, Szepes Z, Pavics L, Bueno L, Wittmann T. Leaky gut in patients with diarrhea-predominant irritable bowel syndrome and inactive ulcerative colitis. *Digestion*. 2012;85(1):40-46.
21. Piche T, Barbara G, Aubert P, Bruley des Varannes S, Dainese R, Nano JL, Cremon C, Stanghellini V, De Giorgio R, Galmiche JP, Neunlist M. Impaired intestinal barrier integrity in the colon of patients with irritable bowel syndrome: involvement of soluble mediators. *Gut*. 2009;58(2):196-201.
22. Shulman RJ, Jarrett ME, Cain KC, Broussard EK, Heitkemper MM. Associations among gut permeability, inflammatory markers, and symptoms in patients with irritable bowel syndrome. *J Gastroenterol*. 2014;49(11):1467-1476.
23. Bertiaux-Vandaele N, Youmba SB, Belmonte L, Lecleire S, Antonietti M, Gourcerol G, Leroi AM, Dechelotte P, Menard JF, Ducrotte P, Coeffier M. The expression and the cellular distribution of the tight junction proteins are altered in irritable bowel syndrome patients with differences according to the disease subtype. *Am J Gastroenterol*. 2011;106(12):2165-2173.
24. R Core Team. R: A language and environment for statistical computing. 2013. Vienna, Austria: R Foundation for Statistical Computing. Available from: <http://www.R-project.org/>.
25. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol*. 2004;3:Article3.

26. Petryszak R, Keays M, Tang YA, Fonseca NA, Barrera E, Burdett T, Fullgrabe A, Fuentes AM, Jupp S, Koskinen S, Mannion O, Huerta L, Megy K, Snow C, Williams E, Barzine M, Hastings E, Weisser H, Wright J, Jaiswal P, Huber W, Choudhary J, Parkinson HE, Brazma A. Expression Atlas update--an integrated database of gene and protein expression in humans, animals and plants. *Nucleic Acids Res.* 2016;44(D1):D746-752.
27. Chey WD, Lembo AJ, Lavins BJ, Shiff SJ, Kurtz CB, Currie MG, MacDougall JE, Jia XD, Shao JZ, Fitch DA, Baird MJ, Schneier HA, Johnston JM. Linacotide for irritable bowel syndrome with constipation: a 26-week, randomized, double-blind, placebo-controlled trial to evaluate efficacy and safety. *Am J Gastroenterol.* 2012;107(11):1702-1712.
28. Stewart GL, Enfield KSS, Sage AP, Martinez VD, Minatel BC, Pewarchuk ME, Marshall EA, Lam WL. Aberrant Expression of Pseudogene-Derived lncRNAs as an Alternative Mechanism of Cancer Gene Regulation in Lung Adenocarcinoma. *Front Genet.* 2019;10:138.
29. Kent WJ. BLAT--the BLAST-like alignment tool. *Genome Res.* 2002;12(4):656-664.
30. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. The human genome browser at UCSC. *Genome Res.* 2002;12(6):996-1006.
31. Simren M, Abrahamsson H, Bjornsson ES. An exaggerated sensory component of the gastrocolonic response in patients with irritable bowel syndrome. *Gut.* 2001;48(1):20-27.
32. Sjolund K, Ekman R, Lindgren S, Rehfeld JF. Disturbed motilin and cholecystokinin release in the irritable bowel syndrome. *Scand J Gastroenterol.* 1996;31(11):1110-1114.
33. Van Der Veek PP, Biemond I, Masclee AA. Proximal and distal gut hormone secretion in irritable bowel syndrome. *Scand J Gastroenterol.* 2006;41(2):170-177.
34. Zhao D, Zhan Y, Zeng H, Moyer MP, Mantzoros CS, Pothoulakis C. Ghrelin stimulates interleukin-8 gene expression through protein kinase C-mediated NF-kappaB pathway in human colonic epithelial cells. *J Cell Biochem.* 2006;97(6):1317-1327.
35. Camilleri M, Gorman H. Intestinal permeability and irritable bowel syndrome. *Neurogastroenterol Motil.* 2007;19(7):545-552.
36. Aerssens J, Camilleri M, Talloen W, Thielemans L, Gohlmann HW, Van Den Wyngaert I, Thielemans T, De Hoogt R, Andrews CN, Bharucha AE, Carlson PJ, Busciglio I, Burton DD, Smyrk T, Urrutia R, Coulie B. Alterations in mucosal immunity identified in the colon of patients with irritable bowel syndrome. *Clin Gastroenterol Hepatol.* 2008;6(2):194-205.
37. Swan C, Duroudier NP, Campbell E, Zaitoun A, Hastings M, Dukes GE, Cox J, Kelly FM, Wilde J, Lennon MG, Neal KR, Whorwell PJ, Hall IP, Spiller RC. Identifying and testing candidate genetic polymorphisms in the irritable bowel syndrome (IBS): association with TNFSF15 and TNFalpha. *Gut.* 2013;62(7):985-994.
38. Yamada T, Kuramitsu K, Rikitsu E, Kurita S, Ikeda W, Takai Y. Nectin and junctional adhesion molecule are critical cell adhesion molecules for the apico-basal alignment of adherens and tight junctions in epithelial cells. *Genes Cells.* 2013;18(11):985-998.

39. Sakisaka T, Ikeda W, Ogita H, Fujita N, Takai Y. The roles of nectins in cell adhesions: cooperation with other cell adhesion molecules and growth factor receptors. *Curr Opin Cell Biol.* 2007;19(5):593-602.
40. Seim I, Collet C, Herington AC, Chopin LK. Revised genomic structure of the human ghrelin gene and identification of novel exons, alternative splice variants and natural antisense transcripts. *BMC Genomics.* 2007;8:298.
41. Seim I, Carter SL, Herington AC, Chopin LK. Complex organisation and structure of the ghrelin antisense strand gene GHRLOS, a candidate non-coding RNA gene. *BMC Mol Biol.* 2008;9:95.
42. Wu S, Liu J, Wang X, Li M, Chen Z, Tang Y. Aberrant Expression of the Long Non-coding RNA GHRLOS and Its Prognostic Significance in Patients with Colorectal Cancer. *J Cancer.* 2017;8(19):4040-4047.
43. Sivertsen B, Lang M, Frimurer TM, Holliday ND, Bach A, Els S, Engelstoft MS, Petersen PS, Madsen AN, Schwartz TW, Beck-Sickinger AG, Holst B. Unique interaction pattern for a functionally biased ghrelin receptor agonist. *J Biol Chem.* 2011;286(23):20845-20860.
44. Sjolund K, Ekman R, Wierup N. Covariation of plasma ghrelin and motilin in irritable bowel syndrome. *Peptides.* 2010;31(6):1109-1112.
45. Sanger GJ, Furness JB. Ghrelin and motilin receptors as drug targets for gastrointestinal disorders. *Nat Rev Gastroenterol Hepatol.* 2016;13(1):38-48.
46. Coulie B, Tack J, Peeters T, Janssens J. Involvement of two different pathways in the motor effects of erythromycin on the gastric antrum in humans. *Gut.* 1998;43(3):395-400.
47. Jarvie EM, North Laidler VJ, Corcoran S, Bassil A, Sanger GJ. Differences between the abilities of tegaserod and motilin receptor agonists to stimulate gastric motility in vitro. *Br J Pharmacol.* 2007;150(4):455-462.
48. Sanger GJ. Ghrelin and motilin receptor agonists: time to introduce bias into drug design. *Neurogastroenterol Motil.* 2014;26(2):149-155.
49. Broad J, Takahashi N, Tajimi M, Sudo M, Goralczyk A, Parampalli U, Mannur K, Yamamoto T, Sanger GJ. RQ-00201894: A motilin receptor agonist causing long-lasting facilitation of human gastric cholinergically-mediated contractions. *J Pharmacol Sci.* 2016;130(2):60-65.
50. Van Assche G, Depoortere I, Thijs T, Missiaen L, Penninckx F, Takanashi H, Geboes K, Janssens J, Peeters TL. Contractile effects and intracellular Ca²⁺ signalling induced by motilin and erythromycin in the circular smooth muscle of human colon. *Neurogastroenterol Motil.* 2001;13(1):27-35.
51. Jameson JS, Rogers J, Misiewicz JJ, Raimundo AH, Henry MM. Oral or intravenous erythromycin has no effect on human distal colonic motility. *Aliment Pharmacol Ther.* 1992;6(5):589-595.

52. Bassotti G, Chiarioni G, Vantini I, Morelli A, Whitehead WE. Effect of different doses of erythromycin on colonic motility in patients with slow transit constipation. *Gastroenterology*. 1998;36(3):209-213.
53. Venkatasubramani N, Rudolph CD, Sood MR. Erythromycin lacks colon prokinetic effect in children with functional gastrointestinal disorders: a retrospective study. *BMC Gastroenterol*. 2008;8:38.
54. Gurriaran-Rodriguez U, Al-Massadi O, Crujeiras AB, Mosteiro CS, Amil-Diz M, Beiroa D, Nogueiras R, Seoane LM, Gallego R, Pazos Y, Casanueva FF, Camina JP. Preproghrelin expression is a key target for insulin action on adipogenesis. *J Endocrinol*. 2011;210(2):R1-7.
55. Avau B, De Smet B, Thijs T, Geuzens A, Tack J, Vanden Berghe P, Depoortere I. Ghrelin is involved in the paracrine communication between neurons and glial cells. *Neurogastroenterol Motil*. 2013;25(9):e599-608.
56. Fraser CC. G protein-coupled receptor connectivity to NF-kappaB in inflammation and cancer. *Int Rev Immunol*. 2008;27(5):320-350.
57. Ye RD. Regulation of nuclear factor kappaB activation by G-protein-coupled receptors. *J Leukoc Biol*. 2001;70(6):839-848.
58. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. *Nature*. 2012;489(7414):57-74.
59. Xu Y, Kiningham KK, Devalaraja MN, Yeh CC, Majima H, Kasarskis EJ, St Clair DK. An intronic NF-kappaB element is essential for induction of the human manganese superoxide dismutase gene by tumor necrosis factor-alpha and interleukin-1beta. *DNA Cell Biol*. 1999;18(9):709-722.
60. Lai Y, Xu P, Liu J, Li Q, Ren D, Zhang J, Wang J. Decreased expression of the long non-coding RNA MLLT4 antisense RNA 1 is a potential biomarker and an indicator of a poor prognosis for gastric cancer. *Oncol Lett*. 2017;14(3):2629-2634.
61. Villegas VE, Zaphiropoulos PG. Neighboring gene regulation by antisense long non-coding RNAs. *Int J Mol Sci*. 2015;16(2):3251-3266.
62. Faghihi MA, Wahlestedt C. Regulatory roles of natural antisense transcripts. *Nat Rev Mol Cell Biol*. 2009;10(9):637-643.
63. Pandey RR, Mondal T, Mohammad F, Enroth S, Redrup L, Komorowski J, Nagano T, Mancini-Dinardo D, Kanduri C. Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Mol Cell*. 2008;32(2):232-246.
64. Nagano T, Mitchell JA, Sanz LA, Pauler FM, Ferguson-Smith AC, Feil R, Fraser P. The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. *Science*. 2008;322(5908):1717-1720.

65. Zong X, Nakagawa S, Freier SM, Fei J, Ha T, Prasanth SG, Prasanth KV. Natural antisense RNA promotes 3' end processing and maturation of MALAT1 lncRNA. *Nucleic Acids Res.* 2016;44(6):2898-2908.
66. Annahazi A, Ferrier L, Bezirard V, Leveque M, Eutamene H, Ait-Belgnaoui A, Coeffier M, Ducrotte P, Roka R, Incze O, Gecse K, Rosztoczy A, Molnar T, Ringel-Kulka T, Ringel Y, Piche T, Theodorou V, Wittmann T, Bueno L. Luminal cysteine-proteases degrade colonic tight junction structure and are responsible for abdominal pain in constipation-predominant IBS. *Am J Gastroenterol.* 2013;108(8):1322-1331.
67. Coeffier M, Gloro R, Boukhattala N, Aziz M, Lecleire S, Vandaele N, Antonietti M, Savoye G, Bole-Feysot C, Dechelotte P, Reimund JM, Ducrotte P. Increased proteasome-mediated degradation of occludin in irritable bowel syndrome. *Am J Gastroenterol.* 2010;105(5):1181-1188.
68. Gecse K, Roka R, Ferrier L, Leveque M, Eutamene H, Cartier C, Ait-Belgnaoui A, Rosztoczy A, Izbeki F, Fioramonti J, Wittmann T, Bueno L. Increased faecal serine protease activity in diarrhoeic IBS patients: a colonic luminal factor impairing colonic permeability and sensitivity. *Gut.* 2008;57(5):591-599.
69. Lee JW, Park JH, Park DI, Park JH, Kim HJ, Cho YK, Sohn CI, Jeon WK, Kim BI. Subjects with diarrhea-predominant IBS have increased rectal permeability responsive to tryptase. *Dig Dis Sci.* 2010;55(10):2922-2928.
70. Li L, Xiong L, Yao J, Zhuang X, Zhang S, Yu Q, Xiao Y, Cui Y, Chen M. Increased small intestinal permeability and RNA expression profiles of mucosa from terminal ileum in patients with diarrhoea-predominant irritable bowel syndrome. *Dig Liver Dis.* 2016;48(8):880-887.
71. Martinez C, Vicario M, Ramos L, Lobo B, Mosquera JL, Alonso C, Sanchez A, Guilarte M, Antolin M, de Torres I, Gonzalez-Castro AM, Pigrau M, Saperas E, Azpiroz F, Santos J. The jejunum of diarrhea-predominant irritable bowel syndrome shows molecular alterations in the tight junction signaling pathway that are associated with mucosal pathobiology and clinical manifestations. *Am J Gastroenterol.* 2012;107(5):736-746.
72. Zhou Q, Souba WW, Croce CM, Verne GN. MicroRNA-29a regulates intestinal membrane permeability in patients with irritable bowel syndrome. *Gut.* 2010;59(6):775-784.
73. Martinez C, Lobo B, Pigrau M, Ramos L, Gonzalez-Castro AM, Alonso C, Guilarte M, Guila M, de Torres I, Azpiroz F, Santos J, Vicario M. Diarrhoea-predominant irritable bowel syndrome: an organic disorder with structural abnormalities in the jejunal epithelial barrier. *Gut.* 2013;62(8):1160-1168.
74. Wilcz-Villega E, McClean S, O'Sullivan M. Reduced E-cadherin expression is associated with abdominal pain and symptom duration in a study of alternating and diarrhea predominant IBS. *Neurogastroenterol Motil.* 2014;26(3):316-325.
75. Ludidi S, Jonkers D, Elamin E, Pieters HJ, Schaepkens E, Bours P, Kruimel J, Conchillo J, Masclee A. The intestinal barrier in irritable bowel syndrome: subtype-specific effects of the systemic compartment in an in vitro model. *PLoS One.* 2015;10(5):e0123498.

76. Khalif IL, Quigley EM, Konovitch EA, Maximova ID. Alterations in the colonic flora and intestinal permeability and evidence of immune activation in chronic constipation. *Dig Liver Dis.* 2005;37(11):838-849.
77. Mandai K, Nakanishi H, Satoh A, Obaishi H, Wada M, Nishioka H, Itoh M, Mizoguchi A, Aoki T, Fujimoto T, Matsuda Y, Tsukita S, Takai Y. Afadin: A novel actin filament-binding protein with one PDZ domain localized at cadherin-based cell-to-cell adherens junction. *J Cell Biol.* 1997;139(2):517-528.
78. Tanaka-Okamoto M, Hori K, Ishizaki H, Itoh Y, Onishi S, Yonemura S, Takai Y, Miyoshi J. Involvement of afadin in barrier function and homeostasis of mouse intestinal epithelia. *J Cell Sci.* 2011;124(Pt 13):2231-2240.

Chapter 4

The Role of the lncRNA *AFDN-DT* in epithelial permeability

ABSTRACT

Chapter 3 of this dissertation describes the analysis of microarray data from sigmoid mucosal biopsies of patients with irritable bowel syndrome (IBS) and healthy controls (HCs) in order to determine differentially expressed long non-coding RNAs (lncRNAs). In addition, Chapter 3 describes the process and experimentation which led to the decision to further study the lncRNA afadin divergent transcript (*AFDN-DT*). In the microarray data, *AFDN-DT* was up-regulated in IBS with constipation (IBS-C) vs. HCs and was positively correlated with its “head-to-head” sense counter-part, afadin (*AFDN*), which encodes for an actin-binding adherens junction protein. We hypothesized that *AFDN-DT* may play a role in IBS pathophysiology by affecting intestinal permeability. The cell culture experiments described in this chapter address two aims. The first aim was to test the hypothesis that *AFDN-DT* affects epithelial permeability. The second aim was to elucidate the role of *AFDN-DT* in colonic epithelial cells using a systems biology approach to gene expression profiling data.

INTRODUCTION

In Chapter 3 of this dissertation, the long non-coding RNA (lncRNA) *AFDN-DT* was selected for further investigation due to dysregulation in irritable bowel syndrome (IBS) and confirmed gene expression in colonic epithelial cells by *in-situ* hybridization. We hypothesize that *AFDN-DT* affects epithelial permeability, possibly via regulation of afadin (*AFDN*), which encodes an adherens junction protein. The importance of adherens junctions in establishing the epithelial barrier as well as the association between IBS and intestinal permeability are discussed in Chapter 3.

This chapter describes investigations in colonic epithelial cells, which were designed to address the following aims: 1) to test the hypothesis that up-regulation of *AFDN-DT* increased permeability of colonic epithelial cell monolayers; and 2) to determine downstream effectors of *AFDN-DT*.

To perform these investigations, we generated colonic epithelial cells capable of inducible overexpression of *AFDN-DT*. The T84 cell line was used because it is an established model for the study of epithelial barrier function,(1) and one frequently used by our group.(2) Studies were repeated Caco2-BBE (Caco2) cells, another cell line commonly used for the study of epithelial barrier function.(3)

Both T84 and Caco2 cells differentiate and form monolayers of polarized cells when seeded on permeable supports, which allow for separate apical and basolateral media compartments. The study of genetically modified differentiated cells presents a challenge. Differentiation can take one to two weeks and differentiated cells can be difficult to transfect. Therefore, the best approach is to create stably transfected cell lines. We chose to use an inducible expression system in order to control the timing of gene expression. In this way, we could study the effect of *AFDN-DT* after or during differentiation.

Inducible expression of *AFDN-DT* was accomplished through use of the Tet-On Advanced System (Takara Bio) by sequential lentiviral transfection and antibiotic selection of cells

expressing the Tet repressor protein followed by *AFDN-DT* under control of the Tet Responsive Element. The Tet-On Advanced System is a modification of the mechanism for regulation of tetracycline-resistance genes in *Escherichia coli*. In *E. coli*, the Tet repressor protein (TetR) is constitutively bound to the tet operator but dissociates in the presence of tetracycline to allow transcription of resistance genes. The Tet-On Advanced System uses a mutated TetR (rTetR), which binds to tetO but activates transcription in the presence of tetracycline or its analog doxycycline.(4) The promoter controlling expression of *AFDN-DT* in the lentiviral vector (pLVX-Tight-Puro) is a combination of a modified Tet-Responsive Element joined to a modified minimal cytomegalovirus (CMV) promoter which does not bind endogenous mammalian transcription factors, which improves the inducible characteristics of the system. Using this system, we tested the effect of *AFDN-DT* on a) differentiated cells; b) cells undergoing differentiation; and c) differentiated cells treated with pro-inflammatory cytokines known to increase epithelial permeability.

In aim 2 of this chapter, we determined downstream effectors of *AFDN-DT* by investigating changes in gene expression in *AFDN-DT*-overexpressing cells. Differential gene expression and pathway analysis was performed using an approach similar to that described in Chapter 2 on 3' RNA sequencing data, which yields gene level count data.

METHODS

Generation of T84-TR-AFDN-DT and Caco2-TR-AFDN-DT cells

Maintenance of T84 and Caco2-BBE cells

T84 and Caco2-BBE cells, obtained from American Type Culture Collection (ATCC, Manassas, VA), were incubated at 37°C with 5% CO₂ in Dulbecco's Modified Eagle's medium (DMEM, Corning, Fisher Scientific, Hanover Park, IL) supplemented with 10% (v/v) heat-inactivated fetal bovine serum (FBS, Thomas Scientific, Swedesboro, NJ), 10 U/mL penicillin, and 100 µg/mL streptomycin (Gibco, Thermo Fisher Scientific, Waltham, MA). Medium was changed every 2-3 days and cells were passaged at 80% confluency. Cell lines were verified to be free of mycoplasma with the Mycoplasma Detection Kit (ATCC).

Lentiviral particle generation

Lentiviral vectors were generated by transient co-transfection of 293T cells with a three-plasmid packaging combination, as previously described, with minor modifications.⁽⁵⁾ The construct pMD.G was used for the production of the VSV-G viral envelop in combination with the packaging constructs pMDLg/pRRE and pRSV-REV. Briefly, 100-mm dishes of 70-80% confluent 293T cells were co-transfected with 6.5 µg of pMDLg/pRRE, 3.5 µg of pMD.G, 2.5 µg of pRSV-REV and 10 µg of the lentiviral vector (pLVX-Tet-On Advanced, modified or empty pLVX-Tight-Puro) by the calcium phosphate-DNA coprecipitation method.^(6,7) The next day, to obtain high-titer virus production,⁽⁸⁾ cells were washed with phosphate-buffered saline (PBS), and fresh medium containing 10 mM sodium butyrate and 20 mM Hepes was added to the cells. After incubation for 6-8 h, cells were washed and incubated in fresh medium without the sodium butyrate. The virus-containing medium was harvested 16-18 h later and passed through 0.45-µm filters. Aliquots of the virus were stored at -80°C until transductions were performed. Viral titer was determined by assessing viral p24 antigen concentration by ELISA (Coulter Immunotech, Miami, FL).

Lentiviral transfection

For transfection, cells were seeded at 50% confluency in 6-well plates (Olympus Plastics, Genesee Scientific, San Diego, CA). After 24 hours, cells were incubated for 6 hours in lentiviral supernatant diluted with serum-free medium to calculated multiplicities of infection (MOIs) of 10, 30-50, and 100 (not included for empty vector due to lower concentration). Lentivirus was replaced with complete medium which was exchanged for medium containing selection antibiotics after 12 hours.

Generation of T84-TR cells

T84-TR cells were generated by transfection with lentivirus carrying the pLVX-Tet-On Advanced vector (Takara Bio, Mountain View, CA). Following transfection, T84-TR cells were cultured in DMEM with 10% FBS and 0.5 mg/ml G418 (Fisher Scientific) for selection for 14 days.

Generation of T84-TR-AFDN-DT cells

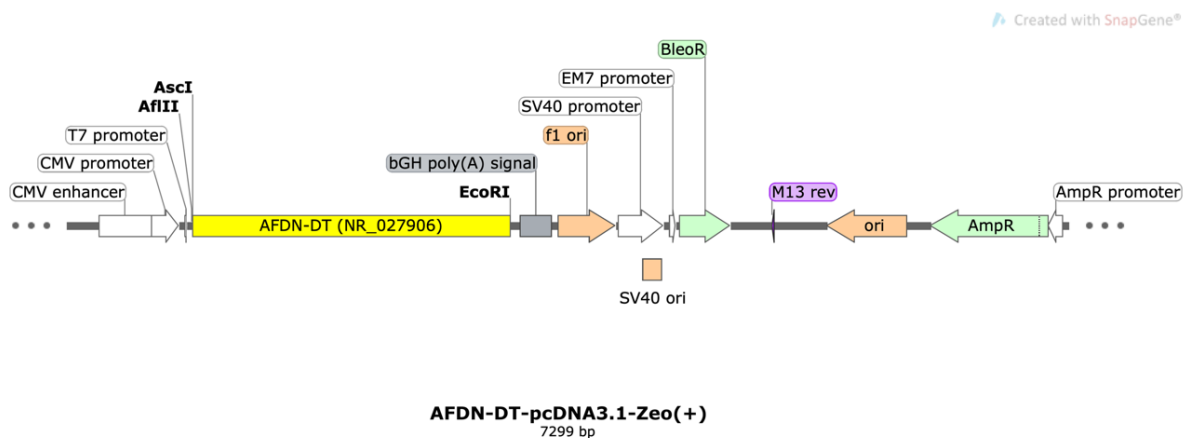
T84-TR-AFDN-DT cells were generated by transfection with lentivirus containing the *AFDN-DT* transcript sequence under control of the Tet Responsive Element. The *AFDN-DT* transcript sequence (NR 027906) flanked by the Afl II and Asc I restriction enzyme sites (5') and the EcoR I restriction enzyme site (3') was synthesized and subcloned (GenScript, Piscataway, NJ) into the pcDNA3.1/Zeo (+) vector at Afl II and EcoR I to generate AFDN-DT-pcDNA3.1-Zeo(+) (Figure 1A). To create the lentiviral vector, the lncRNA sequence was released from the plasmid by digestion with Ascl and EcoRI. The 2.3 kb DNA fragment was isolated from a 1% agarose gel and the DNA purified from the gel using a PCR clean up kit (Promega). The fragment was ligated into a lentiviral vector, pLVX-Tight-PGK-Puro (Takara Bio) that was digested with MluI and EcoRI. The resulting plasmids with the insert (Figure 1B) were sequenced to ensure that the sequence was correct. After transfection, T84-TR-AFDN-DT cells were cultured in DMEM with 10% FBS, 500 µg/ml G418 and 7.5 µg/ml puromycin (Enzo Life Sciences, Farmingdale, NY) for selection for 7 days. Following selection, T84-TR-AFDN-DT cells were maintained in

DMEM with 10% FBS, 100 µg/ml G418 and 250 ng/ml puromycin. T84-TR-EV cells were also generated with the unmodified pLVX-Tight-Puro vector. T84-TR-GFP cells were generated to evaluate for expression in the absence of doxycycline (system leakiness).

Generation of Caco2-TR-AFDN-DT cells

The protocol for generation of Caco2-TR and Caco2-TR-AFDN-DT cells was identical to that of T84 except that a lower concentration of puromycin (2 µg/ml) was used.

A



B

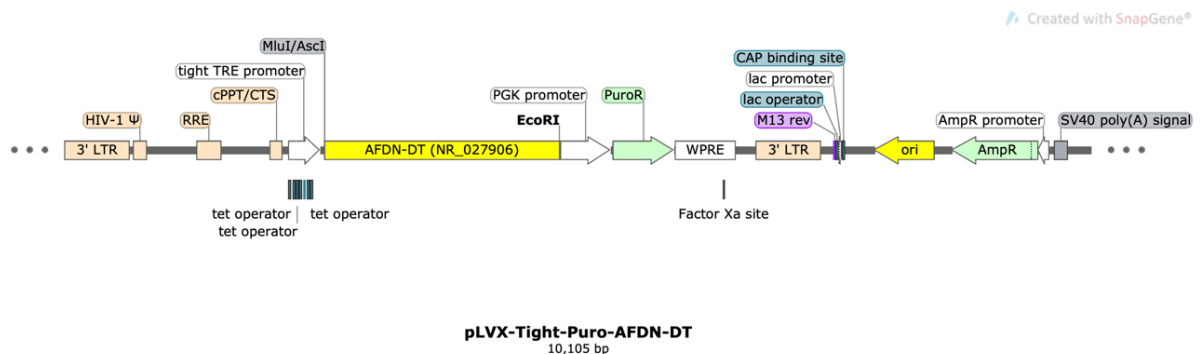


Figure 1: AFDN-DT vector maps

Maps created with SnapGene software (from GSL Biotech; available at snapgene.com).

Permeability measurements

Determination of the effect of *AFDN-DT* following or during differentiation

For growth of differentiated monolayers, 10^5 cells were seeded on Transwell (Corning, Fisher Scientific) permeable supports (6.5-mm diameter, 0.4- μ m pore size) in 24-well plates (Olympus Plastics, Genesee Scientific, San Diego, CA). Resistance (ohms, Ω) was measured using the Millicell ERS-2 Voltohmmeter (Millipore Sigma, Burlington, MA). TEER ($\Omega \cdot \text{cm}^2$) was calculated by multiplying the difference between the resistance of the monolayer and the baseline resistance of the membrane without cells by the surface area of the membrane. TEER readings were performed in triplicate. Doxycycline (1 $\mu\text{g/ml}$) was added once stable TEER was established (4-8 days) or at seeding to determine the effect of *AFDN-DT* during differentiation.

Interferon gamma and TNF alpha response

Recombinant human interferon gamma (IFN- γ) and tumor necrosis factor alpha (TNF- α) were from StemCell Technologies (Cambridge, MA). To ensure an effect on TEER, IFN- γ was used at a concentration of 10 ng/ml as previously described.(9) Medium was exchanged after 48 hours without adding additional IFN- γ . For transcellular permeability measurements, the concentration of IFN- γ was increased to 100 ng/ml and medium was replaced daily with medium containing fresh IFN- γ . For transcellular permeability measurements in Caco2-TR-AFDN-DT, TNF- α (100 ng/ml) was added in combination with IFN- γ . IFN- γ and TNF- α were added to the basal compartment only, as the basal and not the apical region is responsive to IFN- γ .(10)

Paracellular permeability

Following equilibration in Hanks Balanced Salt Solution (HBSS, Corning, Fisher Scientific) with calcium and magnesium for 30 min at 37°C, 1 mg/ml conjugated-dextran in HBSS was added to the apical compartment and fresh HBSS was added to the basal compartment. For T84-TR-AFDN-DT cells, 10-kDa Dextran AlexaFluor 555 (Fisher Scientific) was used, and for Caco2-TR-AFDN-DT cells, 4-kDa dextran-FITC (Millipore Sigma, Burlington, MA) was used. At serial

time points, the HBSS in the basal compartment was removed and replaced with fresh HBSS. The fluorescence of the basal fluid was measured in clear-bottom 96-well plates (Life Technologies, Thermo Fisher Scientific, Waltham, MA) with a microplate reader (Molecular Devices, San Jose, CA) and the concentration of dextran was calculated with a standard curve.

Real-time Quantitative Polymerase Chain Reaction

Wells were washed twice with ice-cold PBS. Total RNA was extracted with TRIzol Reagent (Thermo Fisher Scientific) or with the miRNeasy mini Kit plus on-column DNase treatment (Qiagen, Germantown, MD). RNA was quantified with the NanoDrop spectrophotometer (Thermo Fisher Scientific). Complementary DNA (cDNA) was synthesized from 250-500 ng RNA in a 10- μ L reaction using the iScript™ gDNA Clear cDNA Synthesis Kit (Bio-Rad Laboratories, Hercules, CA) and was diluted 1:2 with nuclease-free water. Real-time quantitative PCR (RT-PCR) was performed in 384-well plates using the CFX384 Touch Real-Time PCR Detection System (Bio-Rad). Each 5- μ L reaction contained 1 μ L diluted cDNA (concentration normalized across samples), 2.5 μ L SYBR® Green Supermix (Bio-Rad), and 0.5 μ L each of 10- μ M forward and reverse primers (Integrated DNA Technologies, Coralville, IA). Primer sequences are shown in Table 1. Resulting data were calculated by using the 2- $^{-\Delta\Delta Ct}$ method, and the level of RNA expression was normalized to *HPRT1* which was constant across samples and conditions.

Table 1: Primer sequences

AFDN-DT	F – TGAGAGATGCCACGTGATCG; R - CACCGGGATGCTCACATAGT
HPRT1	F – GACCAGTCAACAGGGGACAT; R - GCTTGCGACCTTGACCATCT

F, forward; R, reverse

RNA sequencing

Library preparation and sequencing

RNA quantity was verified with the Quant-iT™ RiboGreen™ RNA Assay Kit (Fisher Scientific) and samples were normalized to a concentration of 25 ng/ μ L. RNA quality (RNA integrity

number > 9 for all samples) was confirmed by Agilent TapeStation capillary electrophoresis. RNA sequencing (RNA-seq) was performed in the UCLA Neuroscience Genomics Core Laboratory using Lexogen QuantSeq 3' FWD cDNA library synthesis (11) and multiplex DNA sequencing on an Illumina HiSeq 4000 instrument with single-strand 65 nt sequence reads following standard manufacturer protocols.

Generation of gene count data

Samples yielded >3 million reads. De-multiplexed raw reads (FASTQs) were then subjected to the QuantSeq 3' mRNA-Seq Integrated Data Analysis Pipeline on the Bluebee© Genomics platform (<https://www.bluebee.com/lexogen/>), which uses standard tools but with parameter settings optimized for processing QuantSeq data. This pipeline includes trimming of adapter and polyA sequences, as well as low quality nucleotides using BBDuk (<http://jgi.doe.gov/data-and-tools/bb-tools/>), alignment of trimmed reads against the human genome (GRCh38) with STAR,(12) and determination of gene counts with HTseq.(13)

Statistical analysis

All statistics were performed using the R software version 3.5.2.(14)

Permeability measurements

Percent change in TEER was calculated as the percent of TEER at baseline ($t = 0$). If outliers were present (at least 1.5 times the interquartile range from the maximum/minimum of the interquartile range), values were replaced with the mean from the same group (cell line, treatment).

Group differences

Group differences (gene expression, mean or percent change TEER, flux) were determined with the Wilcoxon signed rank test. A p value threshold of 0.05 was used to determine statistical significance.

Meta-analysis

Meta-analysis of permeability experiments was performed with the meta package for R using a random effects model.(15)

RNA-sequencing

Differential expression

Differential expression was determined using the DESeq2 package in R.(16) The primary outcome was differential expression in T84-TR-AFDN-DT with vs. without doxycycline.

Differential expression was also determined for overall effect of doxycycline and group (T84-TR-AFDN-DT vs. T84-TR-EV) as well as for doxycycline within T84-TR-EV, and T84-TR-AFDN-DT vs. T84-TR-EV without doxycycline. These additional analyses were to assist in determining differences that were specific to doxycycline in T84-TR-AFDN-DT. A Benjamini-Hochberg adjusted p value of 0.1 was used to determine statistical significance.

Weighted gene co-expression network analysis (WGCNA)

WGCNA was performed using the WGCNA package in R.(17) First, according to the WGCNA developer's recommendations, counts were prepared for linear modeling and quantile normalized. Preparation for linear modeling was performed with the variance stabilization function in the DESeq2 package.(16) Quantile normalization was performed using the linear models for microarray (limma) package.(18) WGCNA using "signed" networks was performed to generate both a single network and consensus networks in T84-TR-AFDN-DT and T84-TR-EV. The single network was performed using a soft-thresholding power of 18, a minimum module size of 30, and merging of modules with correlation of 0.75 or greater. The consensus network used the same parameter except a higher threshold for merging (correlation of 0.85 or greater).

Functional annotation

Overrepresented gene ontology (GO) terms in WGCNA modules were determined using the hypergeometric function in the GOstats package in R.(19) For large modules, DAVID was used to assess for clusters of functional annotation terms.(20,21) Gene Set Enrichment Analysis (GSEA) (22) was performed using the Java GSEA Desktop Application (v2.2.4) on the expression data used for the WGCNA analysis (filtered and transformed for linear modeling) using default parameters against the KEGG (23) and GO (24) genes sets obtained from Molecular Signatures Database (v6.0).(25)

RESULTS

Inducible expression of *AFDN-DT* in differentiated colonic intestinal epithelial cells (IECs)

Expression of *AFDN-DT* in differentiated T84 cells was confirmed by RT-PCR. Expression was similar in differentiated and undifferentiated cells ($p = 0.3$). Because epithelial cell differentiation can take 1-2 weeks and transient transfection is not reliably effective in differentiated cells, we elected to use an inducible expression system. Inducible expression of *AFDN-DT* (T84-TR-*AFDN-DT*) was accomplished through use of the Tet-On Advanced System (Takara Bio) by sequential lentiviral transfection and antibiotic selection of the Tet repressor protein followed by *AFDN-DT* under control of the Tet Responsive Element. The Tet-On Advanced System is described in the introduction (page 115).

Verification of overexpression

Overexpression was induced by supplementation of the medium with doxycycline (1 $\mu\text{g/ml}$, Enzo Life Sciences) and was verified by RT-PCR. Since some FBS contains trace tetracycline, T84-TR-GFP was generated and the absence of induced expression without doxycycline was confirmed by absence of fluorescent colonies after 48 h (fluorescent colonies were present 24 h following addition of doxycycline).

Expression of *AFDN-DT* (normalized to *HPRT1*) was increased in differentiated T84-TR-*AFDN-DT*, but not T84-TR-EV cells with vs. without doxycycline at baseline (Figure 2A). Medium with/without doxycycline was added at day six following seeding ($n = 6$ wells per group, 24 wells total). This was also seen in cells treated with IFN- γ (Figure 2B). For IFN- γ -treated cells, there were $n = 7$ wells with doxycycline and $n = 4$ wells without doxycycline for each cell line (22 wells total). Of the 7 wells with doxycycline, doxycycline was added at time of seeding to 3 of the wells and at day 7 to 4 wells. Expression was similar regardless of when doxycycline was added ($p = 1$). *AFDN-DT* expression was greater in T84-TR-*AFDN-DT* wells with vs. without doxycycline at 7 days ($p = 0.029$) and there was a trend for doxycycline at seeding ($p = 0.057$).

The higher p value in this group may be due to fewer wells. A similar trend was seen in undifferentiated T84-TR-AFDN-DT cells (Figure 2C), but this did not reach statistical significance likely due to small sample size.

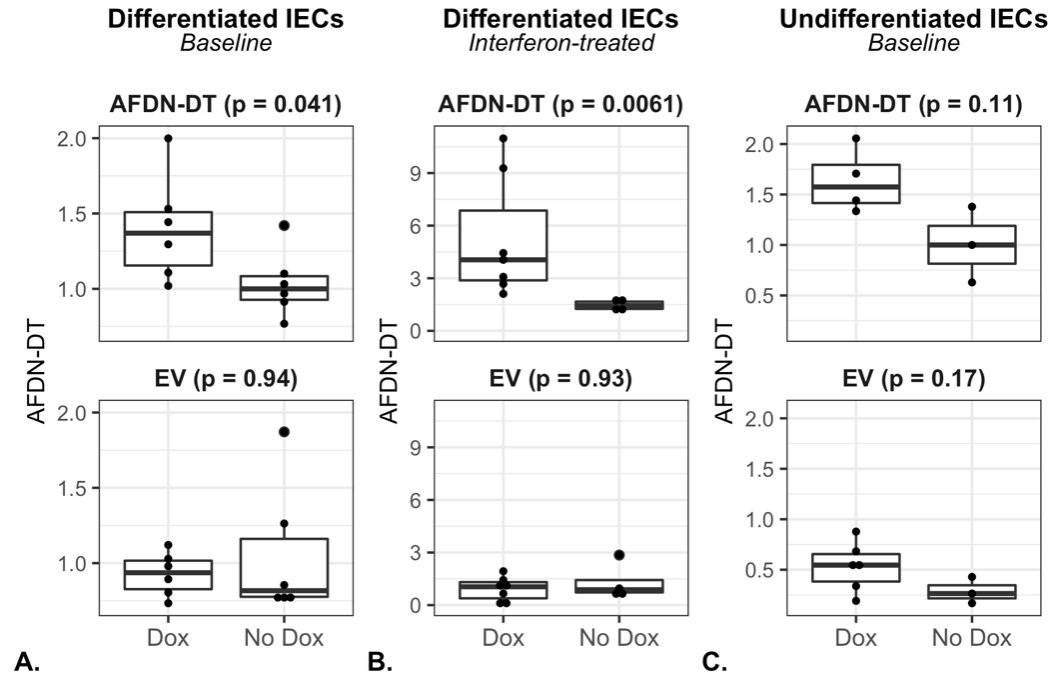


Figure 2: PCR verification of inducible *AFDN-DT* overexpression

Expression of *AFDN-DT* (normalized to *HPRT1*) is increased in differentiated T84-TR-*AFDN-DT* cells with vs. without doxycycline at baseline (A) and following treatment with IFN- γ (B). Panel C shows expression in undifferentiated cells. *IEC*, intestinal epithelial cell; *Dox*, doxycycline

Effect of *AFDN-DT* overexpression on permeability of IEC monolayers

Effect of *AFDN-DT* overexpression on TEER

To determine the effect of *AFDN-DT* overexpression on TEER of IEC monolayers, T84-TR-*AFDN-DT* (n = 16 wells) and T84-TR-EV cells (n = 12 wells) were seeded on permeable supports in 24-well plates. When TEER was > 1000 ohms (day 4-8), medium was replaced with medium +/- 1 μ g/ml doxycycline. Figure 3 shows percent change in TEER following addition of doxycycline for three separate experiments. There was no significant difference in TEER resulting from overexpression of *AFDN-DT* (p > 0.05). Results were similar across varying baseline TEER values, which ranged from 3,000-5,000 (Figure 3A) to 20,000 (Figure 3B).

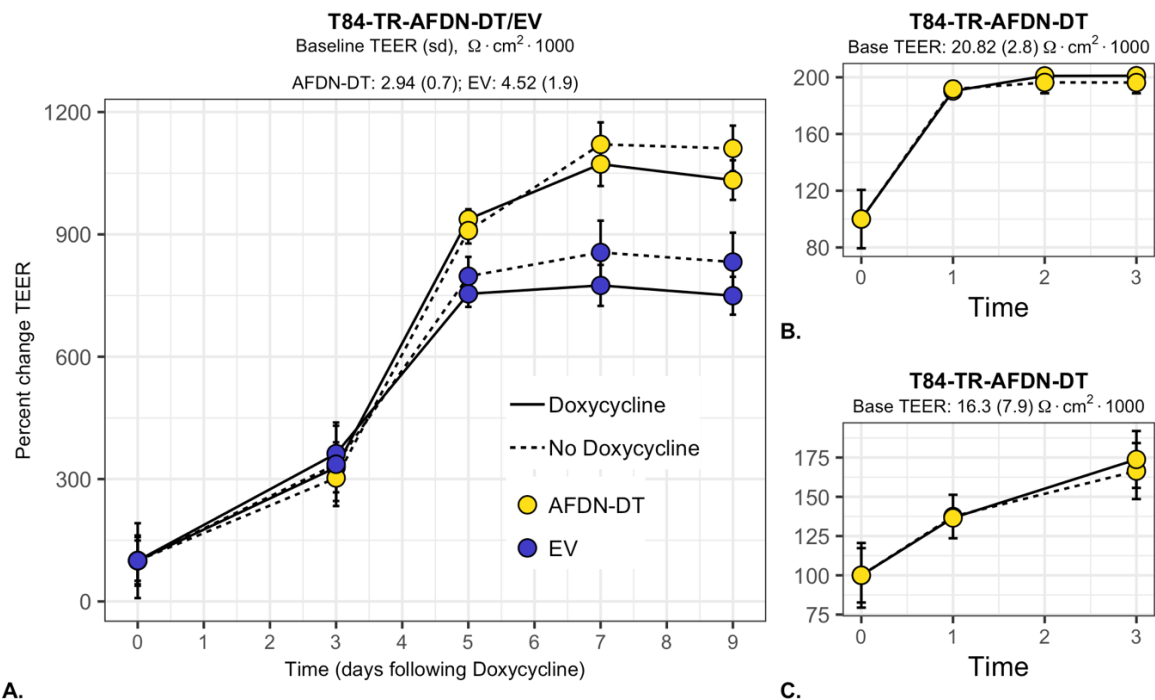


Figure 3: Effect of AFDN-DT overexpression on TEER in differentiated cells

Effect of AFDN-DT overexpression on TEER. A. T84-TR-AFDN-DT (n= 10 and 6 with and without doxycycline) and T84-TR-EV cells (n= 6 and 6 with and without doxycycline) following addition of doxycycline seven days after seeding on permeable supports (Time = 0). B and C show additional experiments with similar results for T84-TR-AFDN-DT (B: n=3 per group, doxycycline added day 6; C: n=4 per group, doxycycline added day 6). *sd*, standard deviation; AFDN-DT, T84-TR-AFDN-DT; EV, T84-TR-EV

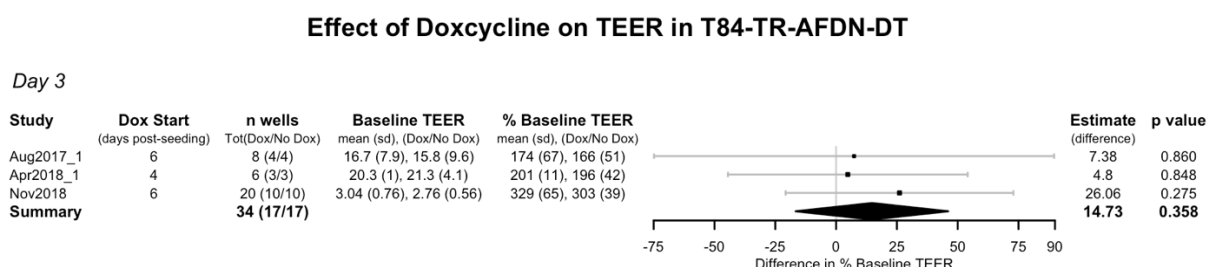


Figure 4: Pooled results for TEER response to AFDN-DT overexpression

Summary result shows the pooled result using random effects model. The data for Aug2017 1, Apr2018 1 and Nov 2018 is shown in Figure 3B, C, and A. *sd*, standard deviation; *Tot*, total; *Dox*, doxycycline

There was a difference between T84-TR-AFDN-DT (n=16) and T84-TR-EV (n=16) with and without doxycycline at days 7 (p = 0.033) and 9 (p = 0.033). Meta-analysis of these experiments also revealed a similar TEER with vs. without overexpression of *AFDN-DT* (Figure 4).

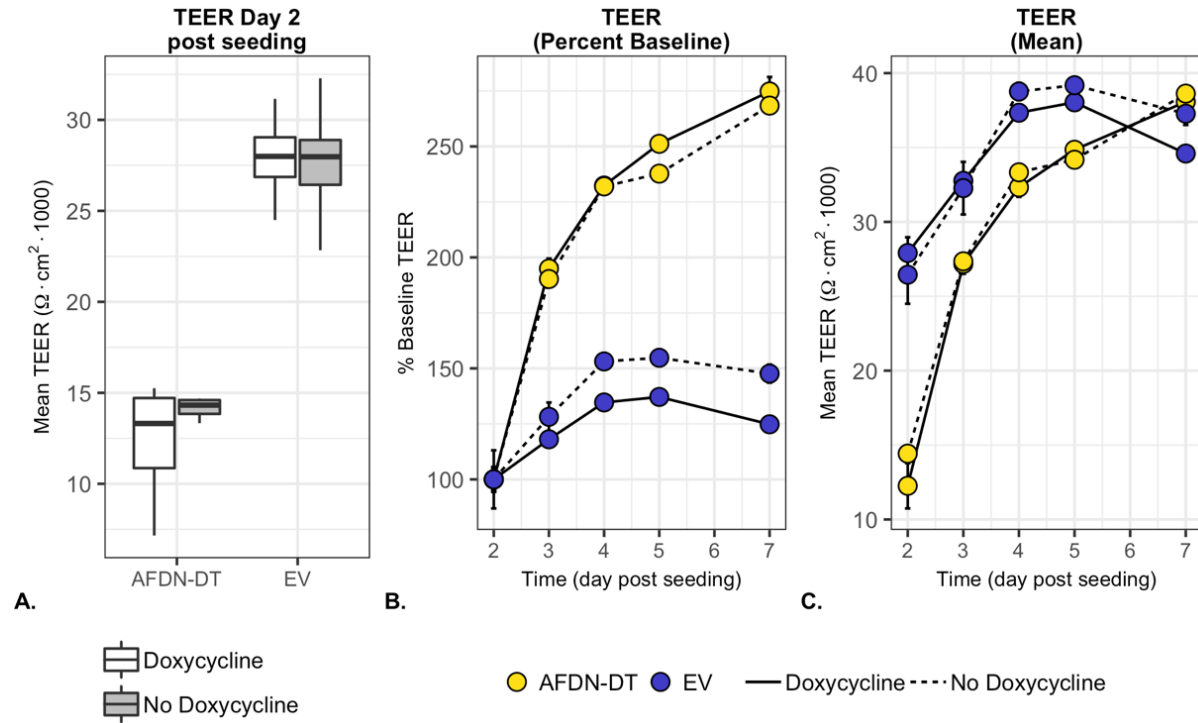


Figure 5: Effect of *AFDN-DT* overexpression during differentiation on TEER

TEER was similar with or without overexpression of *AFDN-DT* during differentiation. A. T84-TR-AFDN-DT and T84-TR-EV (n= 4/8 with/without doxycycline). Doxycycline 1 µg/ml added at time of seeding (Time = 0). B. Percent change in TEER from first recording on day 2 post seeding/doxycycline. C. Mean TEER (same data as in B). *sd*, standard deviation; *AFDN-DT*, T84-TR-AFDN-DT; *EV*, T84-TR-EV

Effect of *AFDN-DT* overexpression during differentiation on TEER

To determine whether overexpression of *AFDN-DT* during the process of differentiation may affect permeability of IEC monolayers, doxycycline was added at the time of seeding (Figure 5). Although TEER at day 2 was higher in T84-TR-EV than T84-TR-AFDN-DT (p < 0.005), there was no difference in either cell line with the addition of doxycycline (p = 0.68 and 1 for T84-TR-AFDN-DT and T84-TR-EV). Subsequent TEER measurements expressed as percent of baseline (day 2) are shown in (Figure 5B). Visualization of raw TEER (Figure 5C) demonstrates

that while differentiation lagged in T84-TR-AFDN-DT (with and without doxycycline), both lines reached similar TEER.

Effect of *AFDN-DT* overexpression on TEER response to IFN- γ

Since TEER did not change with overexpression of *AFDN-DT*, we next sought to determine whether an effect could be seen if the monolayers were treated with the pro-inflammatory cytokine IFN- γ , which results in a reduction in TEER in T84 cells.⁽⁹⁾ T84-TR-AFDN-DT and T84-TR-EV cells were treated with IFN- γ , which was added to the basal compartment at a final concentration of 10 ng/ml. Representative plots and meta-analytic results are shown in Figure 6 and Figure 7. TEER response to IFN- γ was similar with or without overexpression of *AFDN-DT*. In one experiment (Figure 6A-C), there was a small but significant reduction in the response to IFN- γ for T84-TR-AFDN-DT with vs. without doxycycline at day 1 ($p = 0.0061$). Meta-analysis of the differences in percent baseline TEER showed a trend for a reduced response in T84-TR-AFDN-DT with vs. without doxycycline at day 2 ($p = 0.054$, Figure 7A). There was no effect of doxycycline in T84-TR-EV (Figure 7B). The absolute differences were very small (5.4 % at day 1 and 0.19 % at day 3). A difference between the cell lines (T84-TR-AFDN-DT vs. T84-TR-EV) was seen ($p < 0.05$) for days 1-3 (panels A-C) and days 2-3 in panel D.

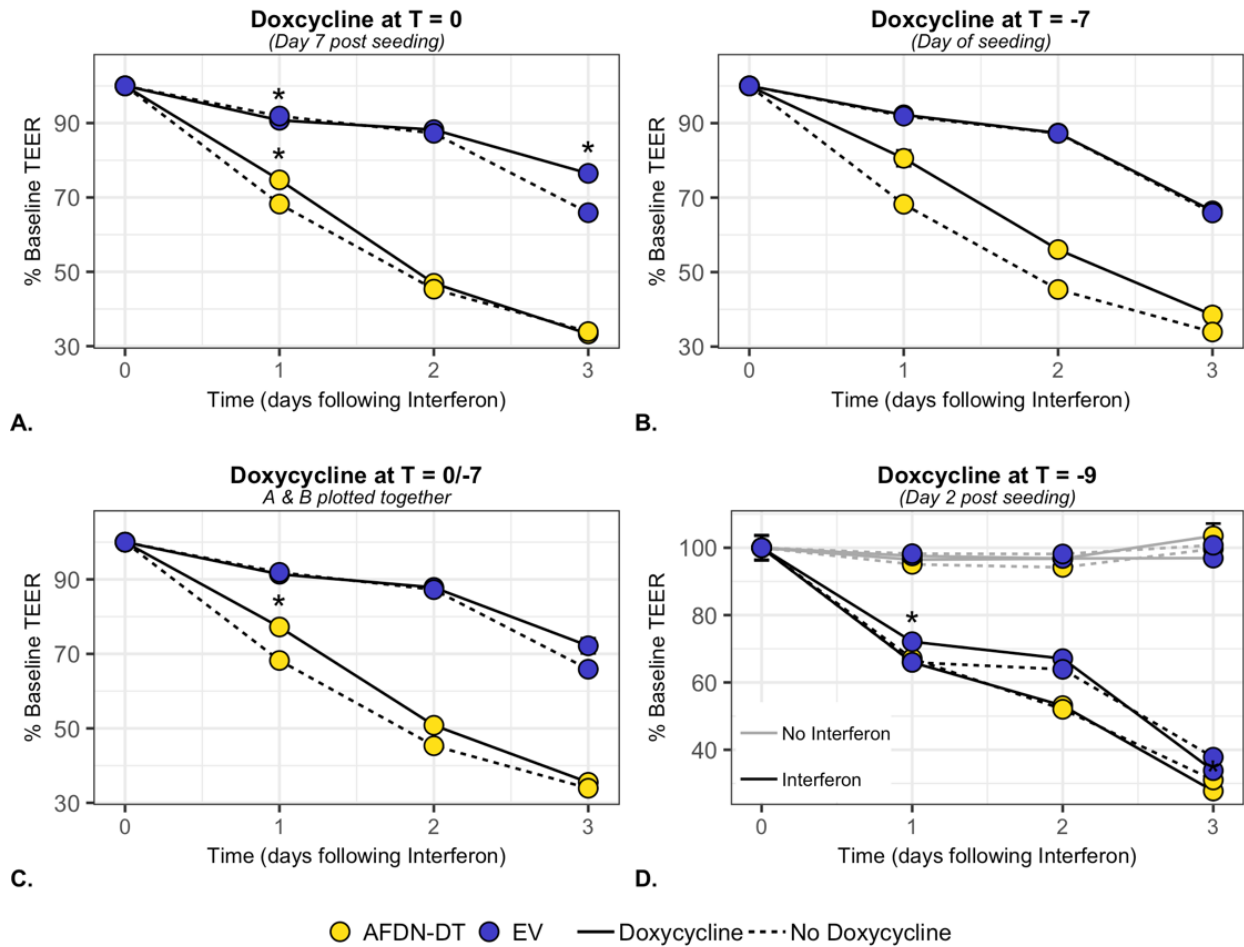


Figure 6: Effect of *AFDN-DT* overexpression on TEER response to IFN- γ

* $p < 0.05$ for doxycycline vs. no doxycycline within cell line (T84-TR-AFDN-DT/EV). A. T84-TR-AFDN-DT and T84-TR-EV ($n = 4/8$ with/without doxycycline). Panels A-C show results of the same experiment which included wells with doxycycline along with IFN- γ (A, $n = 4$ each with/without doxycycline, $n = 16$ total) and at time of seeding (B, $n = 3$ with and 4 without doxycycline, $n = 14$ total). C shows the combination of all doxycycline/no doxycycline wells. Panel D shows a separate experiment ($n = 4$ each with and without doxycycline, $n = 16$ total). *AFDN-DT*, *T84-TR-AFDN-DT*; *EV*, *T84-TR-EV*

Effect of Doxycycline on TEER Response to Interferon Gamma

A. T84-TR-AFDN-DT

Day 1

Study	Start (Dox/Inf) (days post-seeding)	n wells Tot(Dox/No Dox)	Baseline TEER mean (sd), (Dox/No Dox)	% Baseline TEER mean (sd), (Dox/No Dox)	Estimate (difference)	p value
Aug2017_1	6/9	8 (4/4)	25.5 (7.2), 28.7 (2.1)	36.9 (10.1), 26.2 (6.08)	10.67	0.071
Apr2018_2	7*/7	14 (7/7)	38.4 (1.2), 38 (0.42)	77.2 (4.2), 68.2 (2.27)	8.99	0.000
Jul2018	2/11	8 (4/4)	30.3 (0.45), 30.6 (1.3)	66.0 (2.2), 67.1 (0.58)	-1.15	0.305
Summary		30 (15/15)			5.4	0.215

Day 2

Study	Start (Dox/Inf) (days post-seeding)	n wells Tot(Dox/No Dox)	Baseline TEER mean (sd), (Dox/No Dox)	% Baseline TEER mean (sd), (Dox/No Dox)	Estimate (difference)	p value
Aug2017_1	6/9	8 (4/4)	25.5 (7.2), 28.7 (2.1)	23.8 (6.7), 18.7 (3.1)	5.04	0.171
Apr2018_2	7*/7	14 (7/7)	38.4 (1.2), 38 (0.42)	50.8 (5.2), 45.3 (2.1)	5.5	0.010
Jul2018	2/11	8 (4/4)	30.3 (0.45), 30.6 (1.3)	53.1 (2.2), 52.0 (1.6)	1.09	0.413
Summary		30 (15/15)			3.2	0.054

Day 3

Study	Start (Dox/Inf) (days post-seeding)	n wells Tot(Dox/No Dox)	Baseline TEER mean (sd), (Dox/No Dox)	% Baseline TEER mean (sd), (Dox/No Dox)	Estimate (difference)	p value
Aug2017_1	6/9	8 (4/4)	25.5 (7.2), 28.7 (2.1)	14.8 (4.2), 11.2 (3.11)	3.56	0.176
Apr2018_2	7*/7	14 (7/7)	38.4 (1.2), 38 (0.42)	35.5 (3.0), 33.9 (1.66)	1.59	0.218
Jul2018	2/11	8 (4/4)	30.3 (0.45), 30.6 (1.3)	27.8 (1.0), 31.0 (0.57)	-3.27	0.000
Summary		30 (15/15)			0.19	0.931

B. T84-TR-EV

Day 1

Study	Start (Dox/Inf) (days post-seeding)	n wells Tot(Dox/No Dox)	Baseline TEER mean (sd), (Dox/No Dox)	% Baseline TEER mean (sd), (Dox/No Dox)	Estimate (difference)	p value
Apr2018_2	7*/7	14 (7/7)	36.9 (2.9), 34.8 (0.33)	91.4 (0.87), 91.9 (0.47)	-0.51	0.176
Jul2018	2/11	8 (4/4)	26.9 (0.38), 26.9 (0.34)	72.1 (1.22), 66.0 (2.95)	6.13	0.000
Summary		22 (11/11)			2.63	0.427

Day 2

Study	Start (Dox/Inf) (days post-seeding)	n wells Tot(Dox/No Dox)	Baseline TEER mean (sd), (Dox/No Dox)	% Baseline TEER mean (sd), (Dox/No Dox)	Estimate (difference)	p value
Apr2018_2	7*/7	14 (7/7)	36.9 (2.9), 34.8 (0.33)	87.9 (1.2), 87.3 (1.1)	0.6	0.318
Jul2018	2/11	8 (4/4)	26.9 (0.38), 26.9 (0.34)	67.1 (3.7), 63.9 (3.6)	3.18	0.217
Summary		22 (11/11)			0.73	0.211

Day 3

Study	Start (Dox/Inf) (days post-seeding)	n wells Tot(Dox/No Dox)	Baseline TEER mean (sd), (Dox/No Dox)	% Baseline TEER mean (sd), (Dox/No Dox)	Estimate (difference)	p value
Apr2018_2	7*/7	14 (7/7)	36.9 (2.9), 34.8 (0.33)	72.2 (6.2), 65.9 (2.8)	6.28	0.015
Jul2018	2/11	8 (4/4)	26.9 (0.38), 26.9 (0.34)	33.8 (2.2), 37.8 (2.4)	-3.98	0.014
Summary		22 (11/11)			0.95	0.852

Figure 7: Pooled results for the effect of AFDN-DT overexpression on the TEER response to IFN- γ

Summary result shows the pooled difference in percent baseline TEER using a random effects model. The data for Apr2018 2 and Jul2018 is shown in C and D. * For Apr2018 2, doxycycline start was day 7 for 4 wells and day 0 for 3 wells. *sd*, standard deviation; *Tot*, total; *Dox*, doxycycline

Effect of AFDN-DT overexpression on paracellular permeability

To determine the effect of *AFDN-DT* overexpression on paracellular permeability, the flux of 10-kDa Dextran Alexa Fluor 555 (dextran-555) from apical to basolateral compartments was measured. Dextran-555 (1 mg/ml, 100 μ M) was added to the apical compartment of T84-TR-AFDN-DT and T84-TR-EV cells with vs. without doxycycline which had been treated with IFN- γ (10 ng/ml) for three days or had not received IFN- γ (n=4 per group/treatment, n=32 total). Dextran in the basolateral compartment was measured at 1, 3 and 9 hours. TEER values for these samples are shown in Figure 6D. Baseline (day 3 in Figure 6D) TEER (sd) was 9,297 (815) and 28,595 (1,949) for IFN- γ -treated and untreated cells. Cells treated with doxycycline had increased paracellular permeability (Figure 8). This was true for T84-TR-AFDN-DT with IFN- γ and for T84-TR-AFDN-DT and T84-TR-EV without IFN- γ . Although not statistically significant, the same pattern was seen in T84-TR-EV with IFN- γ . Multivariate regression revealed a significant overall effect of IFN- γ ($p < 0.005$) and for doxycycline ($p < 0.005$). There was no significant cell-line by doxycycline interaction effect ($p = 0.87$), which indicates that the effect of doxycycline was similar in T84-TR-AFDN-DT and T84-TR-EV.

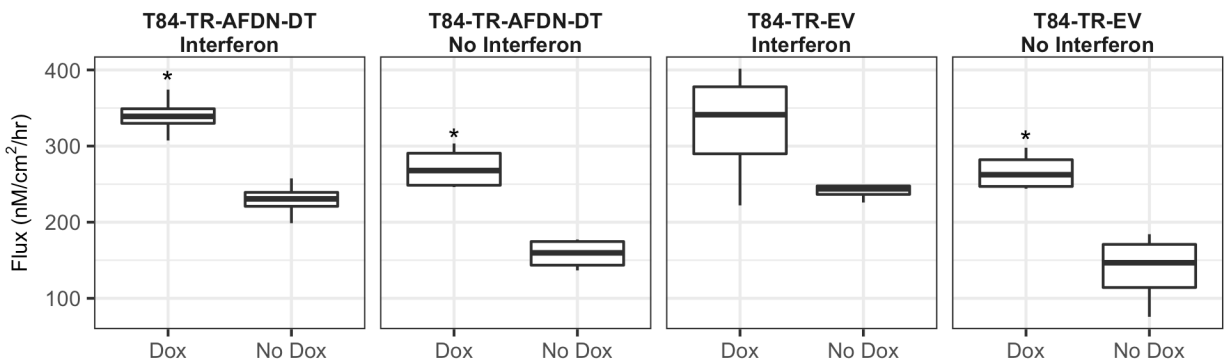


Figure 8: Effect of *AFDN-DT* overexpression on paracellular permeability

Median flux of 10kDa Dextran Alexa Fluor 555 for T84-TR-AFDN-DT and T84-TR-EV (n = 4 wells per group/treatment/IFN- γ n=32 wells total). Flux of control well without cells = 4,735 nM/cm²/hr.

* indicates $p < 0.05$ for doxycycline vs. no doxycycline

Replication in Caco2 cells

In order to determine whether the results were specific to T84 cells, Caco2 cells capable of inducible expression of *AFDN-DT* were generated with the same lentivirus. Permeability experiments were repeated in these cells (Caco2-TR-*AFDN-DT*) and cells transfected with an empty vector (Caco2-TR-EV).

Verification of *AFDN-DT* overexpression

Caco2-TR-*AFDN-DT* cells treated with vs. without doxycycline had increased expression of *AFDN-DT* (normalized to *HPRT1*, Figure 9). The difference was not statistically significant likely due to sample size ($p = 0.2$).

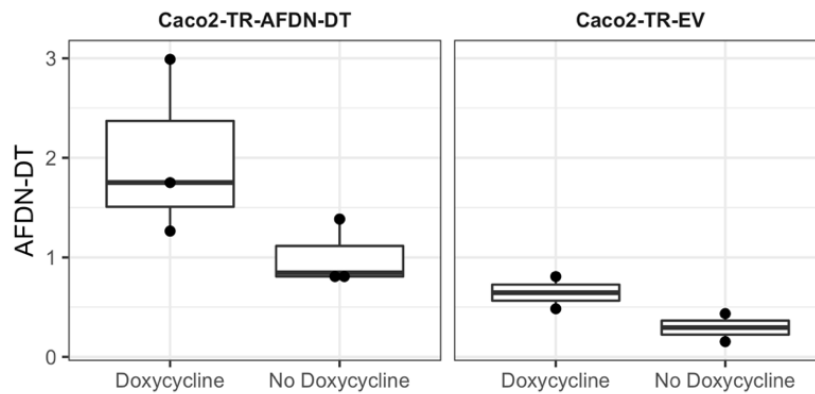


Figure 9: Inducible *AFDN-DT* overexpression in Caco2-TR-*AFDN-DT*

Expression of *AFDN-DT* (normalized to *HPRT1*) in differentiated Caco2-TR-*AFDN-DT* cells with vs. without doxycycline. Doxycycline was added at day 12 following seeding and cells harvested 2 days following doxycycline treatment. $p = 0.2$ (Caco2-TR-*AFDN-DT*), 0.33 (Caco2-TR-EV).

Effect of *AFDN-DT* overexpression on TEER of Caco2 monolayers

TEER of Caco2-TR-*AFDN-DT* and Caco2-TR-EV in response to doxycycline vs. no doxycycline is shown in Figure 10. Baseline TEER (sd) was 6,058 (396) and 6,179 (468) for Caco2-TR-*AFDN-DT* and Caco2-TR-EV. Doxycycline was added on day 12 following seeding on permeable supports. TEER was similar in Caco2-TR-*AFDN-DT* with vs. without doxycycline. There was a small ($< 10\%$) but significant ($p < 0.05$) decrease in TEER with vs. without doxycycline in Caco2-TR-EV.

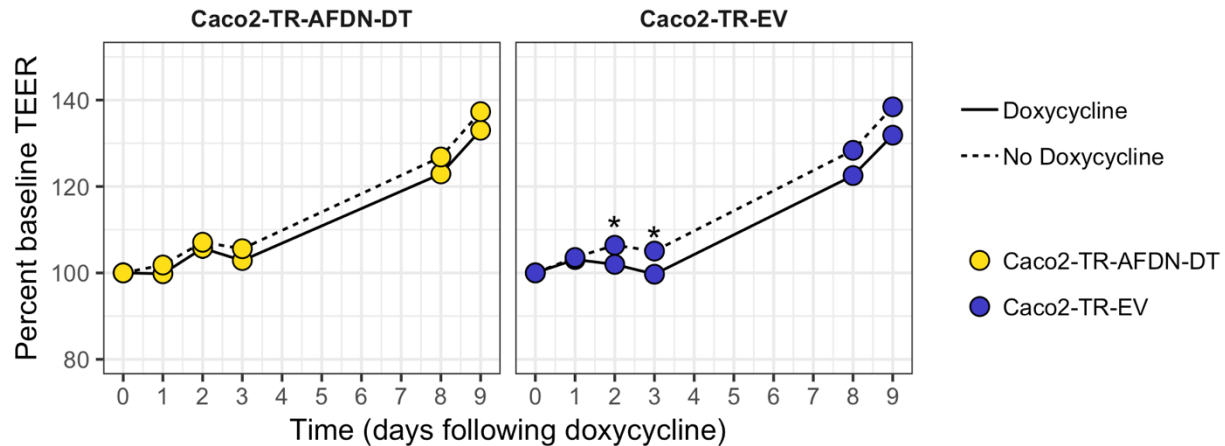


Figure 10: Effect of *AFDN-DT* overexpression on TEER of Caco2 monolayers

Doxycycline was added day 12 after seeding to Caco2-TR-AFDN-DT and Caco2-TR-EV cells (n=11/8 with/without doxycycline).

* indicates $p < 0.05$ for doxycycline vs. no doxycycline.

Effect of *AFDN-DT* overexpression on TEER response to IFN- γ and TNF- α in Caco2 monolayers

The TEER response of Caco2-TR-AFDN-DT and Caco2-TR-EV with vs. without doxycycline to IFN- γ followed by the combination of IFN- γ and TNF- α is shown in Figure 10. Twenty-one days after seeding (9 days after doxycycline treatment), basal compartment medium was exchanged for medium with or without 100 ng/ml IFN- γ (n = 4 per cell line/doxycycline/treatment for IFN- γ -treated cells and n = 3 per cell line/doxycycline/treatment for IFN- γ /TNF- α -treated cells, n = 28 total). Baseline TEER (sd) was 6,058 (396) and 6,179 (468) for Caco2-TR-AFDN-DT and Caco2-TR-EV. TEER response to IFN- γ was similar in Caco2-TR-AFDN-DT and Caco2-TR-EV cells with vs. without doxycycline. There was a small (<5%), but statistically significant ($p < 0.05$) increase in Caco2-TR-AFDN-DT without vs. with doxycycline at day 1 in cells without IFN- γ . Despite the increased concentration of IFN- γ (100 ng/ml vs. 10 ng/ml in T84 experiments), the response to IFN- γ in Caco2 cells, which was a decrease of about 40%, was smaller in comparison to T84 cells (typically >50%). In an attempt to maximize transcellular permeability for subsequent experiments, untreated cells were treated on day 3 with IFN- γ and TNF- α (both 100 ng/ml). Cells treated with IFN- γ on day 0 continued to be treated with IFN- γ . In

all cases the medium was exchanged daily with medium containing fresh cytokines. The response to IFN- γ and TNF- α is shown by the red line in Figure 11. There was no effect of doxycycline on the response to IFN- γ /TNF- α ($p > 0.05$).

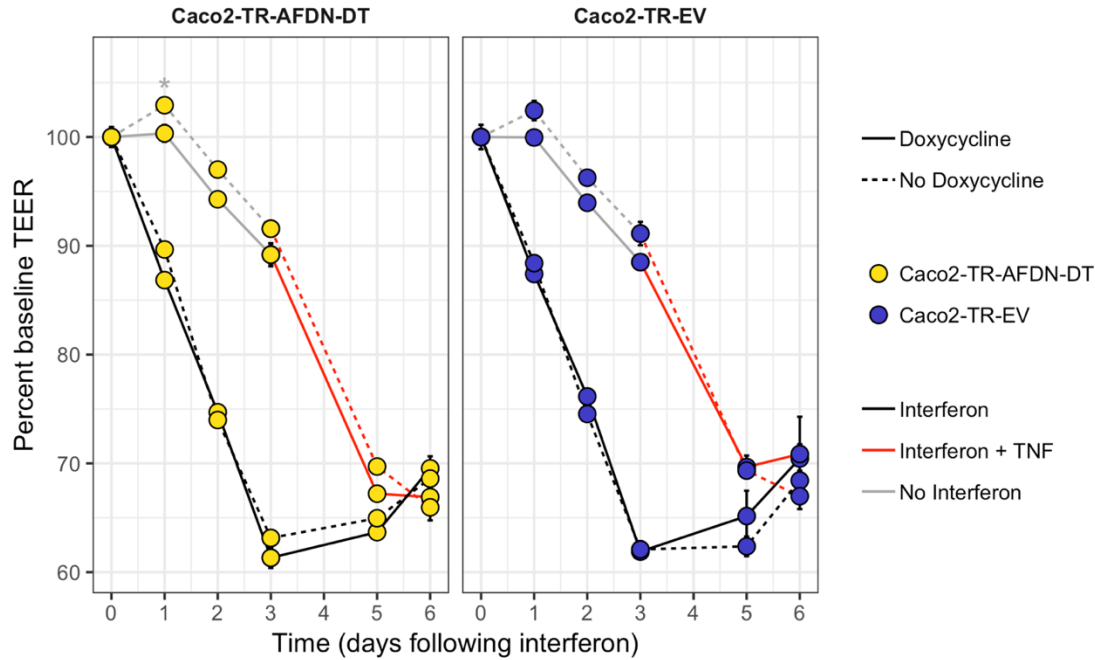


Figure 11: Effect of *AFDN-DT* overexpression on TEER response to IFN- γ and TNF- α in Caco2

IFN- γ (100 ng/ml) was added at day 21 after seeding (day 9 after doxycycline) to Caco2-TR-AFDN-DT and Caco2-TR-EV cells ($n=4$ per cell line/doxycycline/treatment, $n=28$ total). At day 3, untreated cells were exposed to IFN- γ and TNF- α (red line, $n=3$ per cell line/doxycycline/treatment).

* indicates $p < 0.05$ for doxycycline vs. no doxycycline.

Effect of *AFDN-DT* overexpression on paracellular permeability in Caco2 monolayers

To determine the effect of *AFDN-DT* overexpression on paracellular permeability in Caco2 cells, the flux of 4-kDa Dextran FITC (dextran-FITC) from apical to basolateral compartments was measured. Dextran-FITC (1 mg/ml, 250 μ M) was added to the apical compartment of Caco2-TR-AFDN-DT and Caco2-TR-EV cells with/without doxycycline. IFN- γ -treated cells ($n=4$ per cell line/doxycycline, $n=16$ total) were treated with 100 ng/ml for 6 days. IFN- γ + TNF- α treated cells ($n=3$ per cell line/doxycycline, $n=12$ total) were treated with IFN- γ and TNF- α (both 100 ng/ml) for 3 days. Dextran in the basolateral compartment was measured at 8 and 24 hours. There was

no effect of doxycycline on flux in either cell line (Figure 12). TEER values for these samples are shown in Figure 11. Baseline (day 6 in Figure 11) TEER (sd) was 5,736 (154) and 5,368 (272) for IFN- γ and IFN- γ + TNF- α treated cells.

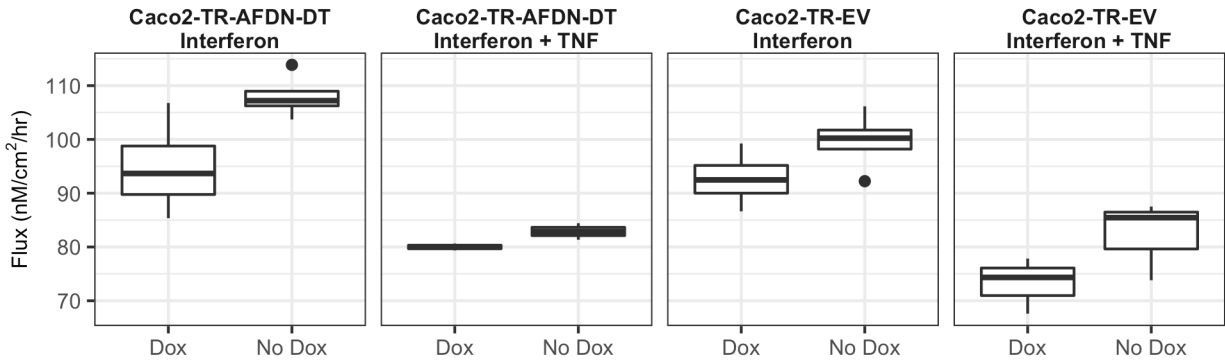


Figure 12: Effect of *AFDN-DT* overexpression on paracellular permeability in Caco2

Median flux of 4-kDa Dextran FITC for Caco2-TR-AFDN-DT and Caco2-TR-EV. IFN- γ -treated cells (n=4 per cell line/doxycycline, n=16 total) were treated with 100 ng/ml for 6 days. IFN- γ + TNF- α treated cells (n=3 per cell line/doxycycline, n=12 total) were treated with IFN- γ and TNF- α (both 100 ng/ml) for 3 days. Flux of control well without cells = 4,381 nM/cm²/hr.

Gene expression profiling of T84-TR-AFDN-DT and T84-TR-EV with and without doxycycline

In order to determine downstream effectors of *AFDN-DT*, we performed RNA sequencing (RNA-seq) on differentiated T84-TR-AFDN-DT and T84-TR-EV cells with and without doxycycline (n=4 per group/doxycycline, n=16 total). The samples were selected from those shown in Figure 2A based on RNA quality and within group uniformity of *AFDN-DT* expression levels. RNA concentration was normalized to 25 ng/ μ l. Sequencing libraries were prepared with the QuantSeq 3'mRNA-Seq Library Prep Kit FWD for Illumina (Lexogen, Vienna, Austria).(11) For traditional mRNA Sequencing library preparation, mRNAs (with polyA tails) are isolated, fragmented, and then reversed transcribed into complementary DNA (cDNA) with random primers. Longer transcripts yield more fragments and thus higher counts, which can introduce a bias with longer transcripts more likely to be differentially expressed.(26) QuantSeq and other 3' library preps use primers targeting the polyA tail. This results in one transcript per sequence. Furthermore, since the 3' end is conserved between isoforms, fewer reads are needed for

confident read alignment. Because of the lower required sequencing depth, libraries can be multiplexed (with addition of primers containing index sequences to allow deconvolution of results) and sequenced in a single lane which reduces costs considerably. This method provides gene-level but not transcript-level counts. Mean RNA quality (RNA integrity number), reads yielded by sequencing, as well as TEER on day of harvest and *AFDN-DT* gene expression (RT-PCR) by group is shown in Table 2. There were no group differences in RNA quality, sequencing depth or TEER. Expression of *AFDN-DT* was increased in T84-TR-AFDN-DT treated with doxycycline compared with each other group. One sample (T84-TR-AFDN-DT with doxycycline) was excluded as it was determined to be an outlier following sample clustering).

Table 2: RNA quality and reads

Cell line	T84-TR-AFDN-DT		T84-TR-EV	
Treatment	-	Doxycycline	-	Doxycycline
N wells	4	3	4	4
RIN	9.9 (0.1)	9.8 (0.3)	9.9 (0.1)	9.8 (0.2)
Reads x 10 ⁶	3.6 (0.3)	3.2 (0.1)	3.5 (0.1)	3.3 (0.2)
TEER x 10 ³	34.3 (0.6)	32.6 (3.2)	34.5 (0.6)	34.2 (1.1)
AFDN-DT expression	0.9 (0.1)	1.6 (0.4)*	0.8 (0)	0.9 (0.1)

All values are mean (standard deviation). * indicates $p < 0.005$ vs. all other groups. *RIN*, RNA integrity number

Differential expression

The numbers of differentially expressed genes (DEGs) for different comparisons and at different significance thresholds is shown in Table 3. There were more DEGs between the two cell lines, both overall and restricted to wells without doxycycline, than between cells within the same group (T84-TR-AFDN-DT/T84-TR-EV) with vs. without doxycycline. The principal components (PCA) plot shown in Figure 13 reveals that cells with vs. without doxycycline did not cluster separately, which is in concordance with the low number of DEGs. The sixteen DEGs in T84-TR-AFDN-DT with vs. without doxycycline ($FDR < 0.1$) are shown in Table 4. There were four

genes for which there was also an effect of doxycycline in T84-TR-EV ($p < 0.05$) indicating that the effect may be due to doxycycline. On visualization of dot plots, two other genes also appeared to have the same pattern of change in T84-TR-EV. These six genes are in italics in Table 4. Expression of the remaining ten genes by group was visualized in dot plots (Figure 14).

Table 3: Summary of differential expression

Comparison	n DEGs all (up/down)		
	FDR < 0.05	FDR < 0.10	P < 0.05
Dox. (All wells with vs. without Dox.)	3 (2/1)	5 (3/2)	506 (237/269)
Group (All T84-TR-AFDN-DT vs. EV)	208 (117/91)	291 (157/134)	797 (424/373)
Dox*Group interaction effect	-	-	575 (276/299)
Dox. (T84-TR-AFDN-DT)	5 (3/2)	16 (7/9)	140 (45/95)
Dox. (EV)	-	-	296 (145/151)
T84-TR-AFDN-DT vs. EV in wells without Dox.	85 (60/25)	148 (91/57)	418 (233/185)

DEGs, differentially expressed genes; Dox., doxycycline; EV, T84-TR-EV

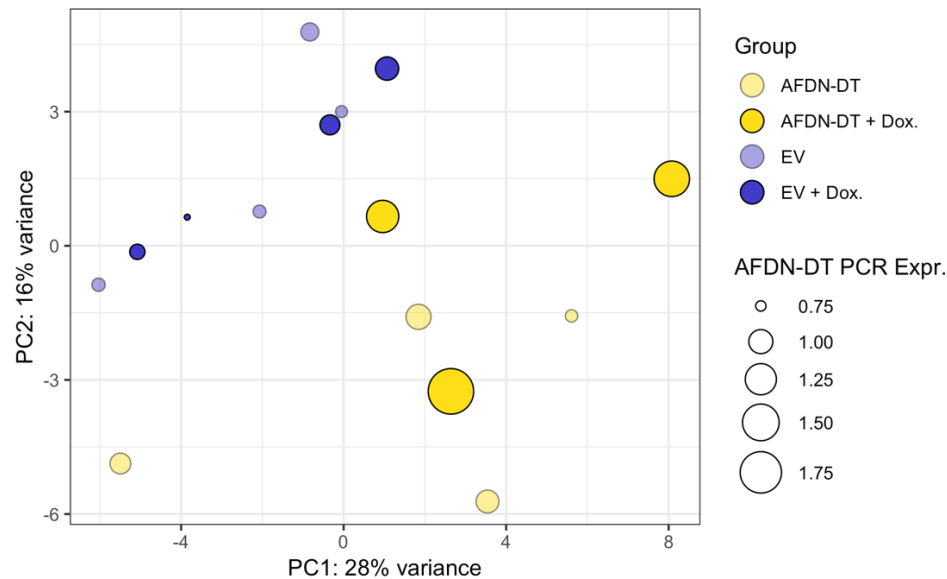


Figure 13: RNA-seq principal components plot

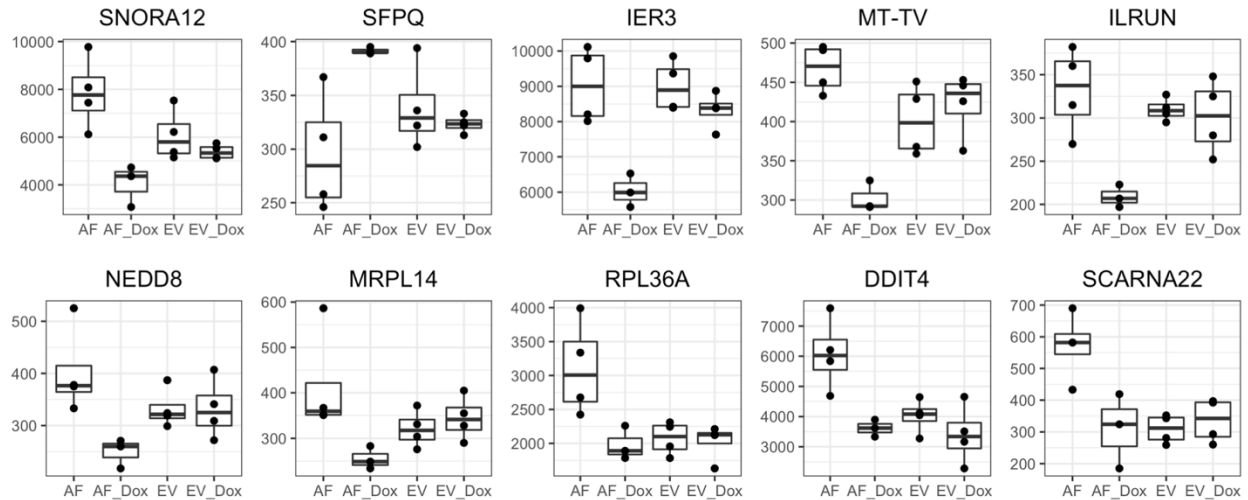
The first two principal components are plotted. Each dot represents a well. Dot size is proportional to expression of *AFDN-DT* by RT-PCR.

Dox., doxycycline

Table 4: Differential gene expression in T84-AFDN-DT with vs. without doxycycline

Symbol, name	Mean	FC	P	Adj. P
Up-regulated				
<i>MT-ATP8, mitochondrially encoded ATP synthase membrane subunit 8</i>	755	1.6	***	0.001
SFPQ , splicing factor proline and glutamine rich	326	1.5	***	0.006
<i>MT-ND1, mitochondrially enc. NADH:ubiquinone oxidoreductase core subunit 1</i>	1403	1.4	***	0.022
<i>MT-CYB, mitochondrially encoded cytochrome b</i>	4795	1.3	***	0.059
<i>MTATP6P1, MT-ATP6 pseudogene 1</i>	297	1.6	**	0.079
<i>MT-CO2, mitochondrially encoded cytochrome c oxidase II</i>	12965	1.2	**	0.079
<i>MT-ND6, mitochondrially enc. NADH:ubiquinone oxidoreductase core subunit 6</i>	707	1.5	*	0.096
Down-regulated				
SNORA12 , small nucleolar RNA, H/ACA box 12	5806	0.6	***	0.002
IER3 , immediate early response 3	8002	0.7	***	0.013
MT-TV , mitochondrially encoded tRNA-Val (GUN)	394	0.7	***	0.063
ILRUN ¹ , inflammation and lipid regulator with UBA-like and NBR1-like domains	284	0.7	***	0.063
RPL36A , ribosomal protein L36a	2250	0.7	**	0.088
NEDD8 , neural precursor cell expressed, developmentally down-regulated 8	324	0.7	**	0.088
MRPL14 , mitochondrial ribosomal protein L14	328	0.7	*	0.096
DDIT4 , DNA damage inducible transcript 4	4203	0.7	*	0.096
SCARNA22 , small Cajal body-specific RNA 22	377	0.6	*	0.096

Italicized entries are likely to be doxycycline-related. *** $p < 0.0005$, ** $p < 0.005$, * $p < 0.05$; FC, fold change; Adj., adjusted (Benjamini-Hochberg); enc., encoded. ¹ILRUN was renamed from C6orf106 in April 2019

**Figure 14: Plots of differentially expressed genes in T84-AFDN-DT with vs. without doxycycline**

AF, T84-TR-AFDN-DT; AF Dox, T84-TR-AFDN-DT + doxycycline; EV, T84-TR-EV; EV Dox, T84-TR-EV + doxycycline

WGCNA

Initial network generation

In order to gain more insight into the pathways and functions of DEGs, a weighted gene co-expression network analysis was performed.⁽¹⁷⁾ The initial WGCNA analysis was performed in all fifteen samples. There were nineteen modules (Figure 15A), though several were quite large. I was unable to find parameters that resulted in a network with a reasonable (<50) number of modules that did not contain several very large modules. For these modules, DAVID (20,21) was used to identify clusters within the modules, which was the case for the “Ribosome/Mitochondria” modules. The correlations of module eigengenes with cell line and treatment is shown in Figure 15B. The first two columns show correlations with cell line for all wells and for wells without doxycycline. Some of the modules were strongly associated with cell line, but in most cases, the association was similar even in wells without doxycycline, indicating that most differences were due to underlying differences in the cells and not overexpression of *AFDN-DT*.

Consensus network analysis

A single network was important in order to view module association with doxycycline for both T84-TR-AFDN-DT and T84-TR-EV. A trial of network generation by cell line revealed that the presence of the large modules in the single network was an effect of clustering by cell line (gene co-expression was strongly related to overall differences in cell lines). This can be seen in Figure 15B as the large modules are correlated with cell line. In order to construct modules more likely to reflect biological pathways rather than group differences between T84-TR-AFDN-DT and T84-TR-EV, a consensus network WGCNA was performed. This analysis performs clustering in two related data sets (each cell line) and finds modules present in both sets. This allows visualization and comparison between the cell lines of T84-related modules, as well as correlation with doxycycline in each cell line. Consensus network intermodular connectivity and

correlations with treatment (doxycycline vs. no doxycycline) and *AFDN-DT* expression by PCR are shown in Figure 16. There were 24 consensus modules. The overall differences between T84-TR-AFDN-DT and T84-TR-EV can be seen by the different patterns of intermodular connectivity in Figure 16A.

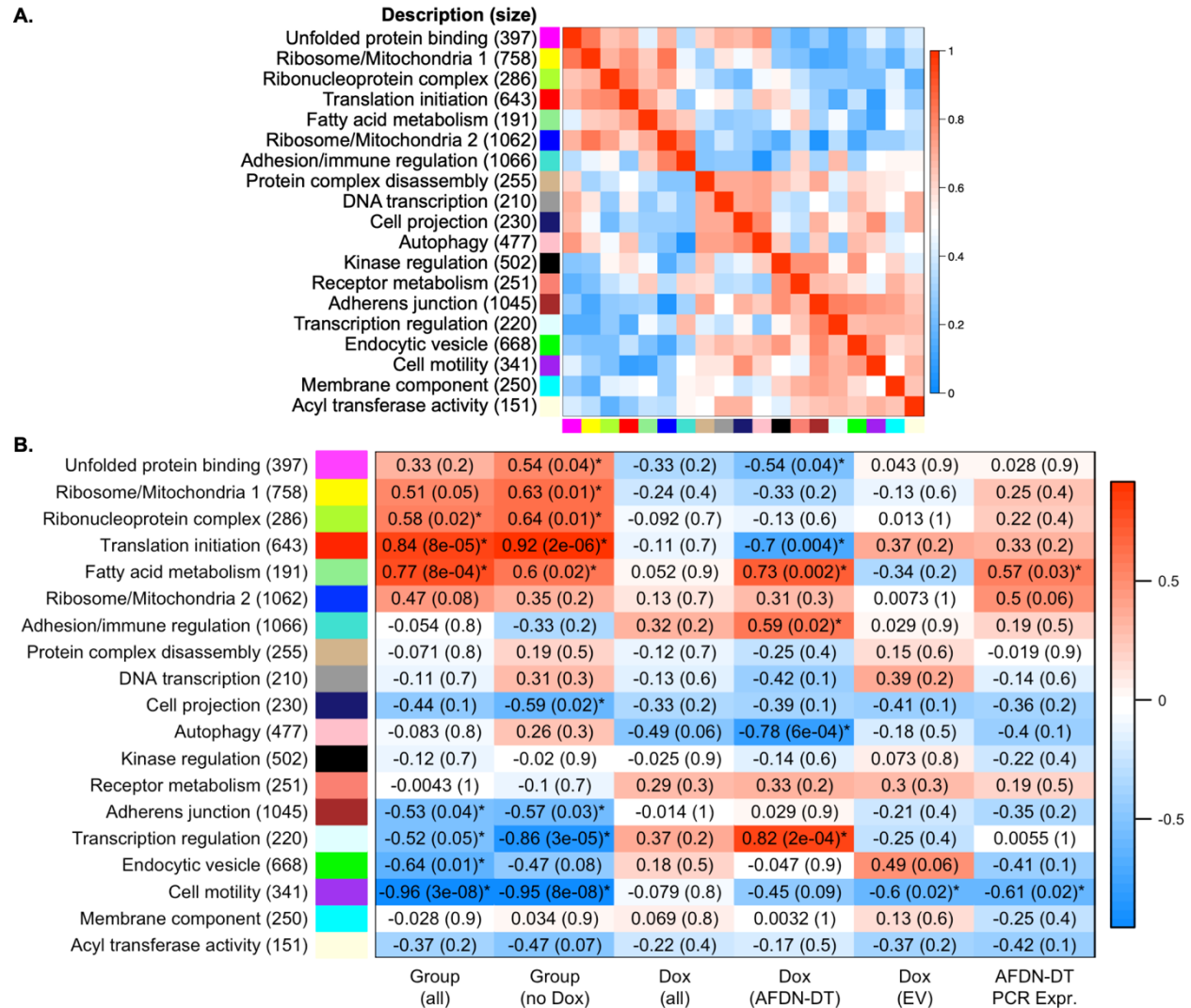


Figure 15: WGCNA, single network

A) Intermodular connectivity. B) Correlations between module eigengenes (adjusted p values) with group (T84-TR-AFDN-DT vs. T84-TR-EV), treatment (doxycycline vs. no doxycycline), and *AFDN-DT* expression by PCR. Group/Dox (all) shows the correlation with cell line or treatment in all wells, whereas Dox (AFDN-DT) and Dox (EV) show the correlation with doxycycline for each cell line (T84-TR-AFDN-DT, T84-TR-EV). Group (no Dox) shows the correlation with cell line only in wells not treated with doxycycline. Dox., doxycycline

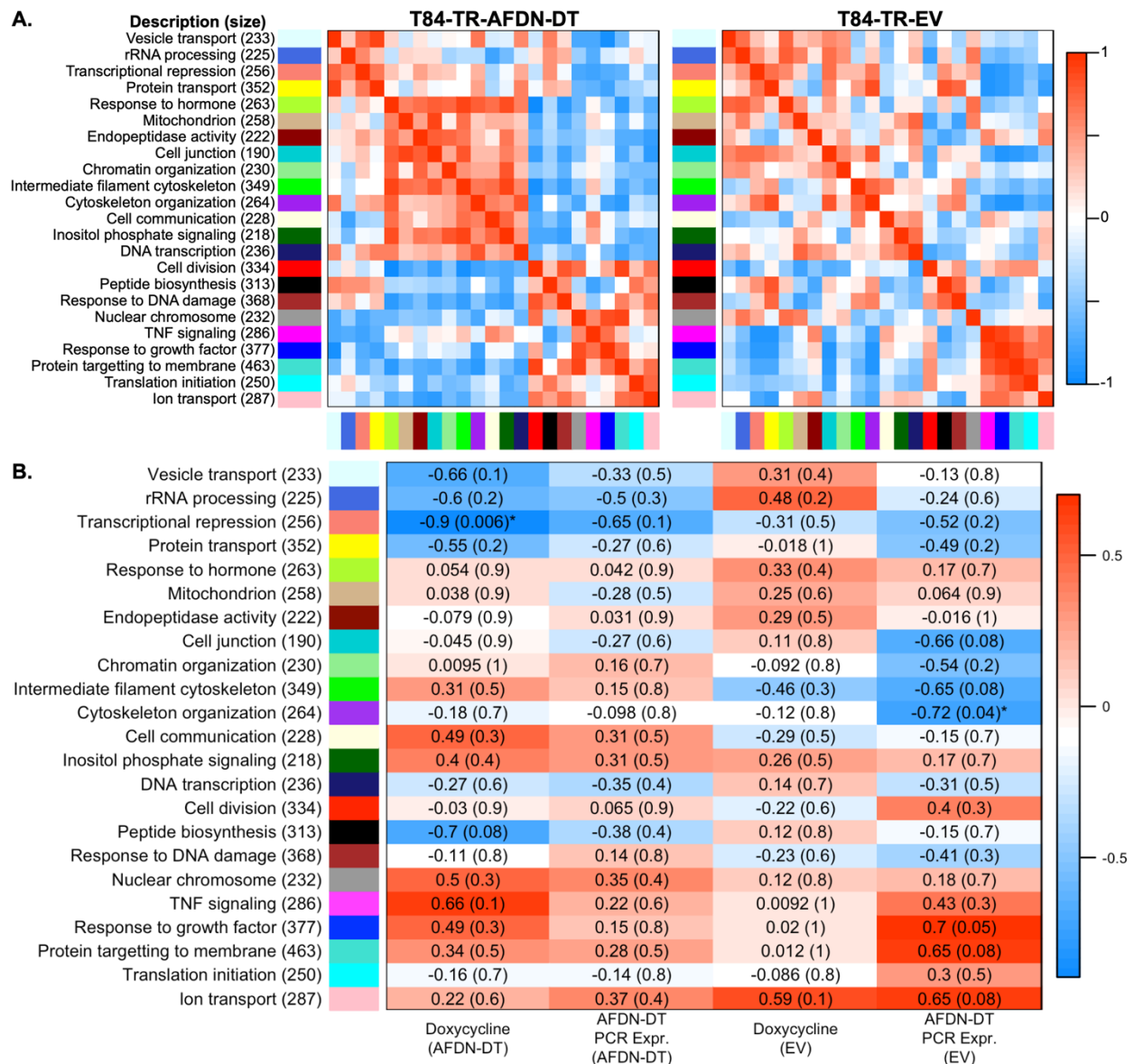


Figure 16: WGCNA, consensus networks

A) Intermodular connectivity. B) Within group (T84-TR-AFDN-DT, T84-TR-EV) correlations for module eigengenes (p values) with treatment (doxycycline vs. no doxycycline) and *AFDN-DT* expression by PCR. *AFDN-DT*, *T84-TR-AFDN-DT*; *EV*, *T84-TR-EV*

Modules correlated with doxycycline in *T84-TR-AFDN-DT*

In the single network WGCNA (Figure 15B), there were six modules correlated with doxycycline in *T84-TR-AFDN-DT*. Three were positively correlated (“Fatty acid metabolism,” “Adhesion/Immune regulation,” and “Transcription regulation”), and three were negatively correlated (“Unfolded protein binding,” “Translation initiation,” and “Autophagy”). There was one

module (“Transcriptional repression”) in the consensus network (Figure 16B) that was negatively correlated with doxycycline in T84-TR-AFDN-DT. The top ten genes with the highest module membership (correlation with module eigengene, MM) in these modules are shown in Table 5.

Module membership of DEGs

Module membership was assessed for the ten DEGs specific to T84-TR-AFDN-DT for both networks. In the single network, only four had MM > 0.6 in any module: *RPL36A* and *SCARNA22* were in the “Translation initiation” module, *IER3* was in the “Adherens junction” module, and *ILRUN* was in the “Autophagy” module. The rest clustered (though with MM < 0.6) in “Adhesion/immune regulation” (*SFPQ*), “Translation initiation” (*NEDD8*, *MRPL14*, *DDIT4*), and “Autophagy” (*MT-TV*). The six doxycycline-related genes (MM) clustered into “Adhesion/immune regulation” (*MT-ATP8* (0.3), *MT-ND1* (0.3), *MT-CYB* (0.3), *MTATP6P1* (0.8)) and “Ribosome/Mitochondria 2” (*MTATP6P1* (0.8), *MT-CO2* (0.35), *MT-ND6* (0.28)). Seven of the ten T84-TR-AFDN-DT specific DEGs had module membership (MM) > 0.6 in the consensus network for T84-TR-AFDN-DT. These genes were in the following modules (module: gene (MM)): “Transcriptional repression”: *SNORA12* (0.74), *ILRUN* (0.91); “Protein transport”: *RPL36A* (0.85); “Peptide biosynthesis”: *NEDD8* (0.89); “DNA transcription”: *DDIT4* (0.68); “rRNA processing”: *SCARNA22* (0.88). *SFPQ* and *MRPL14* were not in the consensus network indicating that they clustered differently in each cell line. *MT-TV* and *IER3* clustered into “Peptide biosynthesis” and “Intermediate filament cytoskeleton” but with MMs < 0.6.

Table 5: Top genes in modules correlated with doxycycline in T84-TR-AFDN-DT

A. Positive correlation

Symbol	Gene (MM)	Consensus mod. (MM)
Fatty acid metabolism		
LLGL2	LLGL scribble cell polarity complex component 2 (0.86)	GF response (0.78)
SCD	stearoyl-CoA desaturase (0.86)	TNF signaling (0.49)
PDIA4	protein disulfide isomerase family A member 4 (0.84)	GF response (0.35)
CDT1	chromatin licensing and DNA replication factor 1 (0.83)	Cell division (0.34)
CCDC85C	coiled-coil domain containing 85C (0.82)	-
STC2	stanniocalcin 2 (0.82)	Cell division (0.37)
SART1	spliceosome associated factor 1, recruiter of U4/U6.U5 tri-snRNP (0.82)	NA
PSAT1	phosphoserine aminotransferase 1 (0.81)	GF response (0.47)
ARG2	arginase 2 (0.8)	Prot. targ. mem. (0.4)
HMGCS2	3-hydroxy-3-methylglutaryl-CoA synthase 2 (0.79)	GF response (0.72)
Adhesion/immune regulation		
HES1	hes family bHLH transcription factor 1 (0.97)	GF response (0.96)
TMSB4X	thymosin beta 4 X-linked (0.96)	GF response (0.96)
ELOB	elongin B (0.96)	GF response (0.96)
COPS7A	COP9 signalosome subunit 7A (0.95)	TNF signaling (0.75)
CYSTM1	cysteine rich transmembrane module containing 1 (0.95)	GF response (0.89)
DCAF5	DDB1 and CUL4 associated factor 5 (0.95)	GF response (0.93)
UBL5	ubiquitin like 5 (0.94)	GF response (0.94)
ID1	inhibitor of DNA binding 1, HLH protein (0.94)	GF response (0.93)
RPL41	ribosomal protein L41 (0.93)	GF response (0.94)
SLC44A1	solute carrier family 44 member 1 (0.93)	GF response (0.97)
Transcription regulation		
SRRM2	serine/arginine repetitive matrix 2 (0.91)	Cell communication (0.74)
CRIM1	cysteine rich transmembrane BMP regulator 1 (0.89)	IF cytoskeleton (0.6)
KIRREL1	kirre like nephrin family adhesion molecule 1 (0.87)	Cell communication (0.54)
SOX4	SRY-box 4 (0.85)	TNF signaling (0.87)
CDCP1	CUB domain containing protein 1 (0.84)	IP signaling (0.88)
CTNND1	catenin delta 1 (0.83)	Cell communication (0.81)
YLP1M1	YLP motif containing 1 (0.82)	IF cytoskeleton (0.9)
SERINC5	serine incorporator 5 (0.81)	IF cytoskeleton (0.8)
EPB41L2	erythrocyte membrane protein band 4.1 like 2 (0.81)	TNF signaling (0.87)
P4HA2	prolyl 4-hydroxylase subunit alpha 2 (0.8)	Cell communication (0.51)

“-“ indicates that the gene did not cluster into a module. NA indicates that the gene was not in the analysis (was not expressed at levels to meet filtering in both cell lines)

MM, module membership; mod., module; GF, growth factor; TNF, tumor necrosis factor; Prot. targ. mem., protein targeting to membrane; IF, intermediated filament; IP, inositol phosphate

Table continued on next page.

B. Negative correlation

Unfolded protein binding		
NACA	nascent polypeptide associated complex subunit alpha (0.88)	Resp. DNA damage (0.84)
MTCH2	mitochondrial carrier 2 (0.86)	Resp. DNA damage (0.72)
HINT1	histidine triad nucleotide binding protein 1 (0.83)	Cell division (0.69)
FASTKD2	FAST kinase domains 2 (0.83)	Protein transport (0.73)
PIR	pirin (0.83)	Resp. DNA damage (0.69)
KIAA1143	KIAA1143 (0.82)	Resp. DNA damage (0.88)
RCN1	reticulocalbin 1 (0.82)	Cell division (0.61)
H2AFZ	H2A histone family member Z (0.82)	Resp. DNA damage (0.84)
ILVBL	ilvB acetolactate synthase like (0.82)	Resp. DNA damage (0.72)
UFD1	ubiquitin recognition factor in ER associated degradation 1 (0.81)	Ion transport (0.54)
Translation initiation		
RPS10	ribosomal protein S10 (0.94)	rRNA processing (0.84)
SLC25A6	solute carrier family 25 member 6 (0.94)	Peptide biosynthesis (0.89)
MPDU1	mannose-P-dolichol utilization defect 1 (0.93)	Protein transport (0.86)
TOMM7	translocase of outer mitochondrial membrane 7 (0.91)	rRNA processing (0.75)
DDT	D-dopachrome tautomerase (0.9)	rRNA processing (0.91)
EIF6	eukaryotic translation initiation factor 6 (0.89)	Transcr. regulation (0.77)
EMP2	epithelial membrane protein 2 (0.88)	rRNA processing (0.34)
NDUFAF8	NADH:ubiquinone oxidoreductase complex assembly factor 8 (0.87)	Vesicle transport (0.67)
TFCP2L1	transcription factor CP2 like 1 (0.87)	Protein transport (0.58)
S100A4	S100 calcium binding protein A4 (0.86)	Transcr. repression (0.82)
Autophagy		
UXS1	UDP-glucuronate decarboxylase 1 (0.92)	Transcr. repression (0.89)
ARRDC3	arrestin domain containing 3 (0.9)	Protein transport (0.86)
TRAPPC6A	trafficking protein particle complex 6A (0.88)	Transcr. repression (0.83)
MRPL27	mitochondrial ribosomal protein L27 (0.88)	Protein transport (0.88)
PARP4	poly(ADP-ribose) polymerase family member 4 (0.88)	Protein transport (0.94)
PRKAG1	protein kinase AMP-activated non-catalytic subunit gamma 1 (0.87)	Transcr. repression (0.96)
GCC1	GRIP and coiled-coil domain containing 1 (0.86)	Transcr. repression (0.8)
TPX2	TPX2, microtubule nucleation factor (0.84)	Protein transport (0.78)
NDUFS2	NADH:ubiquinone oxidoreductase core subunit S2 (0.84)	Transcr. repression (0.8)
PCYOX1	prenylcysteine oxidase 1 (0.83)	Protein transport (0.76)
Transcriptional repression (consensus module)		
Symbol	Gene (consensus module MM)	Single network mod (MM)
CLIC1	chloride intracellular channel 1 (0.98)	Endocytic vesicle (0.71)
RNF24	ring finger protein 24 (0.97)	Kinase regulation (0.65)
PRKAG1	protein kinase AMP-activated non-catalytic subunit gamma 1 (0.96)	Autophagy (0.87)
HSF2	heat shock transcription factor 2 (0.96)	Autophagy (0.78)
TTF1	transcription termination factor 1 (0.95)	Autophagy (0.74)
C1orf74	chromosome 1 open reading frame 74 (0.94)	Autophagy (0.66)
DARS	aspartyl-tRNA synthetase (0.93)	Autophagy (0.5)
SERBP1	SERPINE1 mRNA binding protein 1 (0.91)	Autophagy (0.62)
ILRUN	inflammation and lipid regulator with UBA-like and NBR1-like domains(0.91)	Autophagy (0.69)
URM1	ubiquitin related modifier 1 (0.9)	Autophagy (0.58)
AP1S1	adaptor related protein complex 1 subunit sigma 1 (0.9)	Autophagy (0.82)

MM, module membership; mod., module; Resp., response; rRNA, ribosomal RNA; Transcr., transcriptional

Gene set enrichment analysis (GSEA)

At the GSEA recommended statistical significance threshold of FDR q value < 0.25 , the KEGG pathway “Fc gamma R-mediated phagocytosis” was enriched in T84-TR-AFDN-DT with vs. without doxycycline. This pathway involves formation of phagosomes following crosslinking of Fc gamma receptors. The top genes accounting for this enrichment were: protein kinase C, alpha (*PRCKA*), phosphoinositide-3-kinase, regulatory subunit 2 (p85 beta, *PIK3R2*), vav 2 oncogene (*VAV2*), phosphoinositide-3-kinase, catalytic, beta polypeptide (*PIK3CB*), and v-akt murine thymoma viral oncogene homolog 1 (*AKT1*). There were six gene sets enriched in T84-TR-AFDN-DT vs. T84-TR-EV and 29 in T84-TR-EV vs. T84-TR-AFDN-DT (Table 6).

Table 6: GSEA results for T84-TR-AFDN-DT vs. T84-TR-EV

Gene Set	ES	FDR
Enriched in T84-TR-AFDN-DT (n=6)		
DNA directed RNA polymerase III complex (GO)	0.60	0.21
Pyrimidine containing compound catabolic process (GO)	0.60	0.14
Pre-ribosome (GO)	0.52	0.14
Pyrimidine metabolism (KEGG)	0.48	0.07
ER associated ubiquitin dependent protein catabolic process (GO)	0.42	0.22
Purine metabolism (KEGG)	0.39	0.14
Enriched in T84-TR-EV (n=29*)		
Regulation of phosphatidylinositol 3 kinase activity (GO)	-0.72	0.16
Arrhythmogenic right ventricular cardiomyopathy ARVC (KEGG)	-0.63	0.03
Desmosome (GO)	-0.60	0.13
Dilated cardiomyopathy (KEGG)	-0.60	0.07
Neurological postsynaptic density (GO)	-0.60	0.14
Intercalated disc (GO)	-0.56	0.10
Cell-cell contact zone (GO)	-0.54	0.07
Extracellular matrix (GO)	-0.50	0.14
Fc gamma r mediated phagocytosis (KEGG)	-0.50	0.07
Extrinsic component of cytoplasmic side of plasma membrane (GO)	-0.50	0.08

*Only top 10 non-redundant gene sets shown. *GSEA*, gene set enrichment analysis; *ES*, enrichment score

In order to better understand the patterns of enrichment in the four groups (each cell line with and without doxycycline), I compared gene sets enriched (nominal p value < 0.05) in cells with vs. without doxycycline for each cell line: T84-TR-AFDN-DT + doxycycline (AF+dox) vs. T84-TR-AFDN-DT (AF) and T84-TR-EV + doxycycline (EV+dox) vs. T84-TR-EV (EV). Gene sets enriched in the doxycycline or no doxycycline group for both cell lines would likely be related to

doxycycline. I also compared gene sets enriched in T84-TR-AFDN-DT vs. T84-TR-EV for both treatment groups (AF+dox vs. EV+dox and AF vs. EV). Gene sets enriched in one cell line both with and without doxycycline would likely to be related differences in the cell lines.

There were no gene sets that were enriched with or without doxycycline in both cell lines (i.e. AF+dox vs. AF and EV+dox vs. EV). There were five gene sets (all GO terms) that were enriched with doxycycline in T84-TR-AFDN-DT but were enriched without doxycycline in T84-TR-EV: “embryonic heart tube development”, “chromosome condensation”, “positive regulation of cell cycle”, “neural tube formation”, and “eye development”. Three of the six gene sets enriched in AF vs. EV (FDR < 0.25) were also enriched in AF+dox vs. EV+dox ($p < 0.05$) and seventeen of the 29 gene sets enriched in EV vs. AF (FDR < 0.25) were also enriched in the EV+dox vs. AF+dox ($p < 0.05$). These are shown in Table 7, rows A and B and represent gene sets that reflect the underlying differences between T84-TR-AFDN-DT and T84-TR-EV. There were also gene sets that had divergent patterns of enrichment in the doxycycline-treated groups. These gene sets (Table 7, rows C and D) were enriched in one cell line without doxycycline, but the other cell line with doxycycline. The five gene sets listed above, which were enriched with doxycycline in T84-TR-AFDN-DT but without doxycycline in T84-TR-EV also represent divergent effects of doxycycline in the cell lines.

Table 7: Comparison of cell line gene set enrichment by doxycycline treatment

Enriched cell line by treatment category		-dox	+dox
A	Pre-ribosome, pyrimidine metabolism (KEGG), ER associated ubiquitin dependent protein catabolic process	AF	AF
B	Cell-cell contact zone, dilated cardiomyopathy (KEGG), intercalated disc, desmosome, cell leading edge, extracellular matrix, transcriptional repressor complex, regulation of phosphatidylinositol 3 kinase activity, side of membrane*, clathrin coated vesicle, ciliary membrane, vacuolar membrane, transporter complex* *also enriched in AF+dox vs. AF ($p < 0.05$)	EV	EV
C	Negative regulation of histone modification*, striated muscle cell differentiation*, Mll1 2 complex *also enriched in EV+dox vs. EV ($p < 0.05$)	EV	AF
D	Transcription from RNA polymerase III promoter (also enriched in AF vs. AF+dox ($p < 0.05$))	AF	EV

Gene sets enriched in AF or EV (FDR < 0.25) that were also enriched in the same cell line with doxycycline are in rows A and B. Gene sets that in one cell line without doxycycline but in the other with doxycycline are in rows C and D. All gene sets are gene ontology terms except where noted.

DISCUSSION

Summary of results

The main results of the investigations in the chapter are the following:

1. *AFDN-DT* was successfully overexpressed in differentiated epithelial cells.
2. *AFDN-DT* overexpression did not affect TEER or paracellular permeability in T84 or Caco2 cells under the following conditions:
 - a. Differentiated
 - b. During differentiation
 - c. Treated with pro-inflammatory cytokines that increase permeability

RNA-seq revealed few genes differentially expressed with overexpression of *AFDN-DT* and pathway analyses supports that this differential expression is likely to be related to the increased transcriptional burden of *AFDN-DT* vs. an empty vector.

Up-regulation of *AFDN-DT* is unlikely to affect epithelial barrier function

There was no effect of *AFDN-DT* overexpression on TEER of differentiated T84 cells. This was confirmed in three experiments with a total of 34 wells (17 each with and with doxycycline, Figure 4), and was also true in Caco2 cells (Figure 10). Furthermore, *AFDN-DT* overexpression had no effect on TEER during differentiation (Figure 5) or on the increase in permeability following treatment with IFN- γ (Figure 6 and Figure 7) or IFN- γ + TNF- α for Caco2 cells (Figure 11). While there was a significant effect at one time point in one experiment (Figure 6), the difference in TEER was small (5.4%) and there was no significant effect in pooled meta-analysis of multiple experiments (30 wells, 15 each with and without doxycycline, Figure 7). There was also no effect on paracellular permeability in T84 (Figure 8) and Caco2 cells (Figure 12), though doxycycline resulted in a small increase in flux in both cell lines for T84. If *AFDN-DT* affected epithelial permeability, differences are likely to have been seen in this experimental system as the expression of *AFDN-DT* was confirmed to be in colonic epithelial cells by *in situ*

hybridization (Chapter 3, Figure 13), overexpression in differentiated cells was confirmed by PCR (T84: Figure 2, Caco2: Figure 9), and T84 and Caco2 cells are both good models of epithelial barrier function.(3,9) This conclusion is supported by the RNA-seq results. A difference in TEER was seen between T84-TR-AFDN-DT and T84-TR-EV both with and without doxycycline, particularly in response to IFN- γ (Figure 6) where a greater decrease was seen in T84-TR-AFDN-DT. The top three differentially expressed genes (fold change, all with adjusted p value $< 1 \times 10^{-20}$) were membrane-associated: aquaporin 8 (*AQP8*, 6.6), lipocalin 15 (*LCN15*, 2.7), and filamin A (*FLNA*, 0.49). *AQP8* is a water channel expressed both intracellularly and on the apical surface of colonic epithelial cells.(27) It is increased in biopsies from IBS-D patients and could account for the observed permeability differences between the cell lines.(28) In addition, adhesion-related gene sets were enriched in T84-TR-EV vs. T84-TR-AFDN-DT (Table 6). These differences in permeability and gene expression between the cell lines indicate that the experimental design was adequate to detect differences in permeability and associated gene expression. While it is possible that these underlying differences in T84-TR-AFDN-DT and T84-TR-EV could have changed the phenotype such that an effect of *AFDN-DT* on permeability was not seen, a functional effect of *AFDN-DT* (even if masked) is not supported by the RNA-seq results, which revealed little effect of *AFDN-DT* on gene expression.

Differentially expressed pathways are likely to be non-specific effects of the overexpression model

T84-TR-AFDN-DT and T84-TR-EV differ in cell adhesion and translational burden

As discussed above, T84-TR-AFDN-DT and T84-TR-EV are different with regard to permeability, which is reflected in gene expression and gene set enrichment. This difference was also seen in the WGCNA analysis, in which the “Adherens junction,” and “Cell motility” (actin-binding proteins) modules were negatively correlated with T84-TR-AFDN-DT vs. T84-TR-EV with and without doxycycline (Figure 15). The heatmap in Figure 15 also shows a group of

modules positively correlated with T84-TR-AFDN-DT vs. T84-TR-EV which include “Unfolded protein binding,” “Ribosome/Mitochondria 1,” “Ribonucleoprotein complex,” and “Translation initiation.”

These findings suggest a difference in translational burden between the cell lines. Additional protein synthesis that is anticipated with this model includes expression of rTetR with associated G418 antibiotic resistance, and puromycin antibiotic resistance present on the vector with the Tet-Responsive Element. The same T84-TR cells were used to create T84-TR-AFDN-DT and T84-TR-EV, so similar expression of rTetR would be expected; however, all cell lines were polyclonal and it is possible that differential “drift” (29) in T84-TR-AFDN-DT vs. T84-TR-EV resulted in different levels of rTetR /G418-resistance translation. It is more likely that the puromycin resistance was transfected at different levels due to the variability in efficiency of lentiviral infection and insertion of viral DNA into the genome. The likelihood of variability in transfection of T84-TR cells with *AFDN-DT/EV* is further increased by the fact that we based our multiplicity of infection (MOI) calculations on an indirect lentiviral measure (p24 assay) rather than direct lentiviral titer. It has been observed that longer inserts decrease lentiviral titers,(29) which would predict increased efficiency and translational burden with T84-TR-EV, but this may not always be the case.

Another source of variability is that the lentiviral genome is transcribed to dsDNA and inserted into the host genome randomly,(30) which can result in off-target effects and differences between cells lines. As mentioned, above, polyclonal populations were used, which mitigates these differences somewhat, but the population can become less diverse after repeated passages due to “drift”.(29) Gene set enrichment of “DNA directed RNA polymerase III complex” and gene sets related to nucleotide metabolism (Table 6) suggest an increase in RNA transcription even without doxycycline in T84-TR-AFDN-DT vs. T84-TR-EV.

Doxycycline increases transcriptional burden and cellular stress in T84-TR-AFDN-DT

Differential gene expression in T84-TR-AFDN-DT cells with vs. without doxycycline may be related to the transcriptional burden associated with *AFDN-DT* (2,311 nucleotides), which can increase metabolic demand on the cell.(29) This is supported by positive correlation of doxycycline treatment in T84-TR-AFDN-DT with the WGCNA modules “Transcriptional regulation” (Figure 15) and negative correlation with “Transcriptional repression” (Figure 16). Review of the DEGs is also supportive. Five are directly related to gene expression: *SCARNA22* and *SNORA12* are both small nucleolar RNAs (snoRNAs), which contribute to ribosomal biosynthesis; *RPL36A* encodes a component of the 60S subunit of cytosolic ribosomes (31); *SFPQ* is a DNA- binding protein involved in RNA splicing and transcriptional repression; and *MT-TV* is a transfer RNA for the amino acid valine.(32,33) The other five DEGs have some role in the response to DNA damage or cellular stress. *NEDD8* is an ubiquitin-like protein. Ubiquitins and similar proteins are linked to other proteins to target them for degradation but also serve a variety of other cellular functions.(34) *NEDD8* is involved in the response to DNA damage,(35) and has also been shown to modify the actin cytoskeleton.(36) *DDIT4* is expressed in response to cellular stress and affects cellular growth and proliferation through negative regulation of the mTOR pathway.(37,38) *IER3* is induced by cellular stress and can promote apoptosis.(39,40) *ILRUN* is a newly characterized antiviral protein that suppresses synthesis of interferon TNF- α in response to double-stranded RNA.(39-41) Of note, endoplasmic reticulum (ER) stress can induce autocrine pro-apoptotic TNF- α signaling,(42) which can result in increased reactive oxygen species and mitochondrial damage.(43)

Increased cellular stress in T84-TR-AFDN-DT cells with doxycycline may also account for divergent effect of doxycycline on the enrichment of certain gene sets. These were gene sets enriched in either T84-TR-AFDN-DT or T84-TR-EV in untreated cells, but enriched in the opposite cell line in doxycycline-treated cells (Table 7, rows C and D). For example, while without doxycycline, “Transcription from RNA polymerase III promoter” was enriched in T84-TR-

AFDN-DT, treatment with doxycycline may have led to compensatory down-regulation of this gene set resulting in enrichment in T84-TR-EV.

Effects of doxycycline are evident in gene expression profiling data

Several of the DEGs in T84-TR-AFDN-DT with vs. without doxycycline were transcripts of mitochondrial origin (mitochondrially encoded ATP synthase 8, mitochondrially encoded cytochrome c oxidase II, mitochondrially encoded NADH dehydrogenase 1 and 6, and mitochondrially encoded cytochrome b). These were also differentially expressed in T84-TR-EV with vs. without doxycycline (although at $p < 0.05$) or had the same trend of expression when plotted. Impairment of mitochondrial function and increased cellular stress is a well-documented effect of tetracyclines.(44-46) These changes are even seen for the doses commonly used in doxycycline-induced expression systems, such as 1 $\mu\text{g/ml}$, which was used in this experiment.(46) While some of the genes were up-regulated with doxycycline, which does not fit with the observed dysfunction of mitochondria, directionality of change for mitochondrial genes should be interpreted with caution when using polyA-based library generation, as polyA-shortening is an important mechanism for transcript degradation in mitochondria.(47,48)

Limitations

Challenges of LncRNA expression models

The most important limitation to the experimental design is that the exogenous expression of *AFDN-DT* may not replicate the function of the lncRNA in its genomic context. For some lncRNAs, the effect of their transcription is entirely based on their proximity to other genes. For example, transcription may change the chromatin structure exposing nearby genes for transcription, or lncRNAs may bind and thereby recruit enhancers or other gene regulatory proteins.(49) The fact that *AFDN* expression was unaffected by *AFDN-DT* overexpression (fold change 0.9) could be due to this out-of-context transcription. This is a considerable limitation to

our design; however, there are limitations to most experimental methods of manipulating lncRNA gene expression, including CRISPR/Cas9 methods.(50)

Selection of controls

The empty vector control was both a strength and a limitation in this study. Since the differential transcriptional burden of the 2,311 nucleotide *AFDN-DT* vs. the vector with no insert was evident in gene expression profiling, a better control may have been a scrambled nucleotide sequence of the same length as *AFDN-DT*; however, it would be difficult to confidently conclude that there were no functional effects of the random sequence that occurred by chance. The choice of an inducible expression system was also a strength as the internal control was not associated with differences in cell populations arising from non-specific or off-target effects of lentiviral transfection. These differences between T84-TR-*AFDN-DT* and T84-TR-EV turned out to be quite significant and incidentally provide some evidence that up-regulation of *AQP8* may increase epithelial permeability. This chance finding demonstrates the effectiveness of approaches that use random mutagenesis and phenotype screening to determine gene function.

Conclusions

In conclusion, exogenous expression of *AFDN-DT* did not affect permeability of colon epithelial cell monolayers. The possibility that *AFDN-DT* could affect other aspects of cellular physiology is unlikely in the absence of gene expression changes associated with *AFDN-DT* expression.

REFERENCES:

1. Annahazi A, Ferrier L, Bezirard V, Leveque M, Eutamene H, Ait-Belgnaoui A, Coeffier M, Ducrotte P, Roka R, Incze O, Gecse K, Rosztoczy A, Molnar T, Ringel-Kulka T, Ringel Y, Piche T, Theodorou V, Wittmann T, Bueno L. Luminal cysteine-proteases degrade colonic tight junction structure and are responsible for abdominal pain in constipation-predominant IBS. *Am J Gastroenterol*. 2013;108(8):1322-1331.
2. Chen ML, Pothoulakis C, LaMont JT. Protein kinase C signaling regulates ZO-1 translocation and increased paracellular flux of T84 colonocytes exposed to *Clostridium difficile* toxin A. *J Biol Chem*. 2002;277(6):4247-4254.
3. Peterson MD, Mooseker MS. Characterization of the enterocyte-like brush border cytoskeleton of the C2BBE clones of the human intestinal cell line, Caco-2. *J Cell Sci*. 1992;102 (Pt 3):581-600.
4. Urlinger S, Baron U, Thellmann M, Hasan MT, Bujard H, Hillen W. Exploring the sequence space for tetracycline-dependent transcriptional activators: novel mutations yield expanded range and sensitivity. *Proc Natl Acad Sci U S A*. 2000;97(14):7963-7968.
5. Naldini L, Blomer U, Gage FH, Trono D, Verma IM. Efficient transfer, integration, and sustained long-term expression of the transgene in adult rat brains injected with a lentiviral vector. *Proc Natl Acad Sci U S A*. 1996;93(21):11382-11388.
6. Sakoda T, Kaibuchi K, Kishi K, Kishida S, Doi K, Hoshino M, Hattori S, Takai Y. smg/rap1/Krev-1 p21s inhibit the signal pathway to the c-fos promoter/enhancer from c-Ki-ras p21 but not from c-raf-1 kinase in NIH3T3 cells. *Oncogene*. 1992;7(9):1705-1711.
7. Chen C, Okayama H. High-efficiency transformation of mammalian cells by plasmid DNA. *Mol Cell Biol*. 1987;7(8):2745-2752.
8. Sakoda T, Kasahara N, Hamamori Y, Kedes L. A high-titer lentiviral production system mediates efficient transduction of differentiated cells including beating cardiac myocytes. *J Mol Cell Cardiol*. 1999;31(11):2037-2047.
9. Donato RP, El-Merhibi A, Gundsambuu B, Mak KY, Formosa ER, Wang X, Abbott CA, Powell BC. Studying permeability in a commonly used epithelial cell line: T84 intestinal epithelial cells. *Methods Mol Biol*. 2011;763:115-137.
10. Adams RB, Planchon SM, Roche JK. IFN-gamma modulation of epithelial barrier function. Time course, reversibility, and site of cytokine binding. *J Immunol*. 1993;150(6):2356-2363.
11. Moll P, Ante M, Seitz A, Reda T. QuantSeq 3' mRNA sequencing for RNA quantification. *Nature Methods*. 2014;11:972.
12. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15-21.

13. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015;31(2):166-169.
14. R Core Team. R: A language and environment for statistical computing. 2013. Vienna, Austria: R Foundation for Statistical Computing. Available from: <http://www.R-project.org/>.
15. Schwarzer G. meta: An R package for meta-analysis. *R News*. 2007;7(3):40-45.
16. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550.
17. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008;9:559.
18. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol*. 2004;3:Article3.
19. Falcon S, Gentleman R. Using GOstats to test gene lists for GO term association. *Bioinformatics*. 2007;23(2):257-258.
20. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009;4(1):44-57.
21. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res*. 2009;37(1):1-13.
22. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102(43):15545-15550.
23. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28(1):27-30.
24. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet*. 2000;25(1):25-29.
25. Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J, Puigserver P, Carlsson E, Ridderstrale M, Laurila E, Houstis N, Daly MJ, Patterson N, Mesirov JP, Golub TR, Tamayo P, Spiegelman B, Lander ES, Hirschhorn JN, Altshuler D, Groop LC. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet*. 2003;34(3):267-273.
26. Oshlack A, Wakefield MJ. Transcript length bias in RNA-seq data confounds systems biology. *Biol Direct*. 2009;4:14.

27. Pelagalli A, Squillacioti C, Mirabella N, Meli R. Aquaporins in Health and Disease: An Overview Focusing on the Gut of Different Species. *Int J Mol Sci.* 2016;17(8).
28. Camilleri M, Carlson P, Chedid V, Vijayvargiya P, Burton D, Busciglio I. 1024 - Expression of Aquaporin Channels in Mucosal Biopsies from Patients with Ibs-Diarrhea and Ibs-Constipation Subgroups Compared to Healthy Controls. *Gastroenterology.* 2018;154(6):S-195-S-196.
29. Shearer RF, Saunders DN. Experimental design for stable genetic manipulation in mammalian cell lines: lentivirus and alternatives. *Genes Cells.* 2015;20(1):1-10.
30. Ramezani A, Hawley RG. Overview of the HIV-1 Lentiviral Vector System. *Curr Protoc Mol Biol.* 2002;Chapter 16:Unit 16 21.
31. Davies MS, Henney A, Ward WH, Craig RK. Characterisation of an mRNA encoding a human ribosomal protein homologous to the yeast L44 ribosomal protein. *Gene.* 1986;45(2):183-191.
32. Lee M, Sadowska A, Bekere I, Ho D, Gully BS, Lu Y, Iyer KS, Trehwella J, Fox AH, Bond CS. The structure of human SFPQ reveals a coiled-coil mediated polymer essential for functional aggregation in gene regulation. *Nucleic Acids Res.* 2015;43(7):3826-3840.
33. Xu J, Zhong N, Wang H, Elias JE, Kim CY, Woldman I, Pifl C, Gygi SP, Geula C, Yankner BA. The Parkinson's disease-associated DJ-1 protein is a transcriptional co-activator that protects against neuronal apoptosis. *Hum Mol Genet.* 2005;14(9):1231-1241.
34. Sun L, Chen ZJ. The novel functions of ubiquitination in signaling. *Curr Opin Cell Biol.* 2004;16(2):119-126.
35. Ma T, Chen Y, Zhang F, Yang CY, Wang S, Yu X. RNF111-dependent neddylation activates DNA damage-induced ubiquitination. *Mol Cell.* 2013;49(5):897-907.
36. Leck YC, Choo YY, Tan CY, Smith PG, Hagen T. Biochemical and cellular effects of inhibiting Nedd8 conjugation. *Biochem Biophys Res Commun.* 2010;398(3):588-593.
37. Sofer A, Lei K, Johannessen CM, Ellisen LW. Regulation of mTOR and cell growth in response to energy stress by REDD1. *Mol Cell Biol.* 2005;25(14):5834-5845.
38. Vega-Rubin-de-Celis S, Abdallah Z, Kinch L, Grishin NV, Brugarolas J, Zhang X. Structural analysis and functional implications of the negative mTORC1 regulator REDD1. *Biochemistry.* 2010;49(11):2491-2501.
39. Sebens Muerkoster S, Rausch AV, Isberner A, Minkenberg J, Blaszczyk E, Witt M, Folsch UR, Schmitz F, Schafer H, Arlt A. The apoptosis-inducing effect of gastrin on colorectal cancer cells relates to an increased IEX-1 expression mediating NF-kappa B inhibition. *Oncogene.* 2008;27(8):1122-1134.
40. Arlt A, Schafer H. Role of the immediate early response 3 (IER3) gene in cellular stress response, inflammation and tumorigenesis. *Eur J Cell Biol.* 2011;90(6-7):545-552.

41. Ambrose RL, Liu YC, Adams TE, Bean AGD, Stewart CR. C6orf106 is a novel inhibitor of the interferon-regulatory factor 3-dependent innate antiviral response. *J Biol Chem*. 2018;293(27):10561-10573.
42. Hu P, Han Z, Couvillon AD, Kaufman RJ, Exton JH. Autocrine tumor necrosis factor alpha links endoplasmic reticulum stress to the membrane death receptor pathway through IRE1alpha-mediated NF-kappaB activation and down-regulation of TRAF2 expression. *Mol Cell Biol*. 2006;26(8):3071-3084.
43. Suematsu N, Tsutsui H, Wen J, Kang D, Ikeuchi M, Ide T, Hayashidani S, Shiomi T, Kubota T, Hamasaki N, Takeshita A. Oxidative stress mediates tumor necrosis factor-alpha-induced mitochondrial DNA damage and dysfunction in cardiac myocytes. *Circulation*. 2003;107(10):1418-1423.
44. Ahler E, Sullivan WJ, Cass A, Braas D, York AG, Bensinger SJ, Graeber TG, Christofk HR. Doxycycline alters metabolism and proliferation of human cell lines. *PLoS One*. 2013;8(5):e64561.
45. Chatzispyrou IA, Held NM, Mouchiroud L, Auwerx J, Houtkooper RH. Tetracycline antibiotics impair mitochondrial function and its experimental use confounds research. *Cancer Res*. 2015;75(21):4446-4449.
46. Moullan N, Mouchiroud L, Wang X, Ryu D, Williams EG, Mottis A, Jovaisaite V, Frochaux MV, Quiros PM, Deplancke B, Houtkooper RH, Auwerx J. Tetracyclines Disturb Mitochondrial Function across Eukaryotic Models: A Call for Caution in Biomedical Research. *Cell Rep*. 2015;10(10):1681-1691.
47. Wang DD, Guo XE, Modrek AS, Chen CF, Chen PL, Lee WH. Helicase SUV3, polynucleotide phosphorylase, and mitochondrial polyadenylation polymerase form a transient complex to modulate mitochondrial mRNA polyadenylated tail lengths in response to energetic changes. *J Biol Chem*. 2014;289(24):16727-16735.
48. Levy S, Schuster G. Polyadenylation and degradation of RNA in the mitochondria. *Biochem Soc Trans*. 2016;44(5):1475-1482.
49. Schmitz SU, Grote P, Herrmann BG. Mechanisms of long noncoding RNA function in development and disease. *Cell Mol Life Sci*. 2016;73(13):2491-2509.
50. Goyal A, Myacheva K, Gross M, Klingenberg M, Duran Arque B, Diederichs S. Challenges of CRISPR/Cas9 applications for long non-coding RNA genes. *Nucleic Acids Res*. 2016, 10.1093/nar/gkw883:gkw883.

Chapter 5:

Gene Expression Profiling of the Gut-Brain Axis in Healthy Deceased Donors

ABSTRACT

While the preceding chapters of this dissertation have focused on changes in the colonic mucosa in irritable bowel syndrome (IBS), alterations in the gut-brain axis is another key concept in IBS pathophysiology. This chapter of my dissertation describes a proof-of-concept analysis in which I have used a systems biology approach in the analysis of gene expression profiling data from the gut and the brain. This approach can be used to explore changes in the gut-brain axis in animal models of IBS. It also identified significant correlations between gene modules (groups of highly correlated genes) in the gut and the brain. The genes in these modules could be important in the gut-brain axis. Interestingly, while we found most differential gene expression in IBS with constipation vs. healthy controls, genes that were in these gut-brain modules were more likely to be differentially expressed in IBS with diarrhea vs. healthy controls. A version of this chapter will be submitted for publication.

INTRODUCTION

The "gut-brain axis" describes the bidirectional communication between the enteric, autonomic and central nervous systems (Figure 1). A unifying theory in the pathogenesis of irritable bowel syndrome (IBS) is that there is a dysregulation in the gut-brain axis that results in altered sensation and motility.(1)

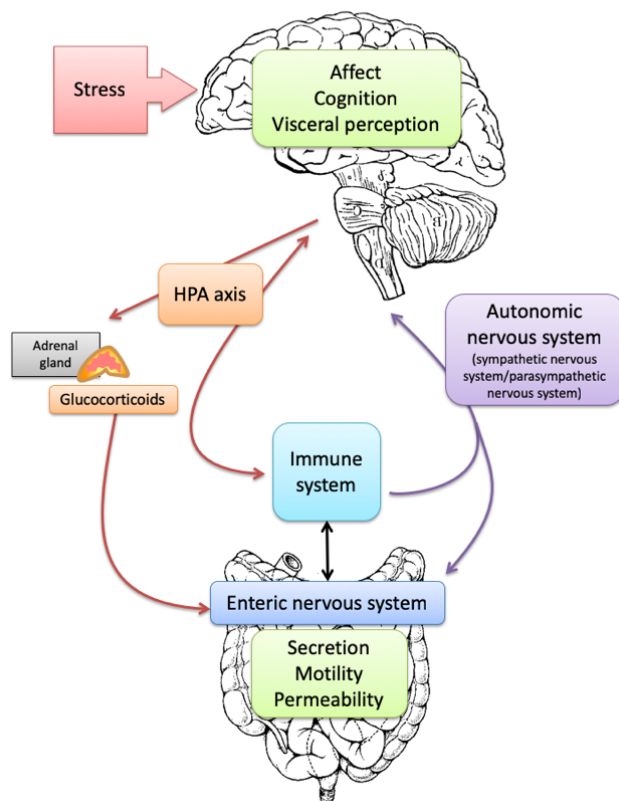


Figure 1: The gut-brain axis

HPA, hypothalamic pituitary adrenal. Modified from Neuroscience & Biobehavioral Reviews, Vol 36, Issue 1, Kennedy P, et. al.(2), *Gut memories: Towards a cognitive neurobiology of irritable bowel syndrome*, pages 310-340, Copyright 2012 with permission from Elsevier.

Dysregulation of the gut-brain axis results in enhanced visceral perception (decreased pressure thresholds to pain and other sensations, and/or increased sensory ratings and viscerosomatic referral areas to barostat distension), which is a reproducible finding in a subset of patients with IBS.(3) Enhanced perception of visceral stimuli can occur in IBS patients either as a result of

greater sensitivity of visceral afferent pathways or as central amplification of visceral afferent input.(4) Gut-brain axis dysregulation may be due to dysfunction in any of the components involved in gut sensation, such as signal transmission from the gut, signal transmission to the central nervous system (CNS), and modulation/processing of sensation by the CNS.(5)

One of the better understood mechanisms for alterations in the gut-brain axis relates to the effect of psychological stress on the gut. There is strong support for a link between chronic stress and IBS.(6-8) Studies in IBS patients have shown that experimental stress increases visceral sensitivity,(9,10) and preclinical studies support stress-induced changes at the mucosal level.(11,12) Additional triggers that can affect the gut-brain axis include infection (post-infection IBS),(13,14) diet and changes in the microbiome.(15) The cellular and molecular mechanisms involved in modulation of the gut-brain axis in health and disease, including IBS, are poorly understood.

I hypothesized that a systems biology approach to the gut-brain axis would yield insight into the integrated biology of the two organs. Furthermore, diseases of the gut-brain axis, such as IBS, could be better understood by comparing the gut-brain axis at the molecular level between disease and control, an approach that would be feasible in animal models.

The goal of the project described in this chapter was to perform this analysis using publicly available data as a proof of concept. While there are multiple publicly available gene expression profiling data-sets from the brain and the colon, these cannot be easily used for cross-tissue comparisons as observed differences are likely to result from experimental and batch effects. Brain and colon tissues should be collected using the same protocol and should not be processed or sequenced in separate batches. Ideally, the tissues should be from the same organisms (each organism contributes both brain and colon). One data-set that meets these requirements is the data from the genotype-tissue expression (GTEx) project.(16) Designed to study the relationship between genetic variation and gene expression across tissues, this project has resulted in a large and publicly available repository of RNA-sequencing (RNA-seq)

samples from healthy deceased donors. The advantage of this data is that it was designed for comparison across tissues. Samples were distributed by tissue and donor to prevent batch effects.

Using data from the GTEx project, I performed several weighted gene co-expression network analyses (WGCNA) using donors with both sigmoid colon (SC) and cerebral cortex (CC) gene expression data. The analyses and the aims addressed by each analysis are outlined in Table 1 and Figure 2.

Table 1: Analyses and aims

Analysis	Description	Aims
Tissue-specific	WGCNA of SC and CC	To identify correlations between SC and CC modules
Consensus	Consensus network of SC/CC	To identify modules present in both SC and CC
Unified	Single WGCNA using merged SC and CC data	To identify modules containing both SC and CC genes and to explore relationships between SC-predominant and CC-predominant modules.

SC, sigmoid colon; CC, cerebral cortex

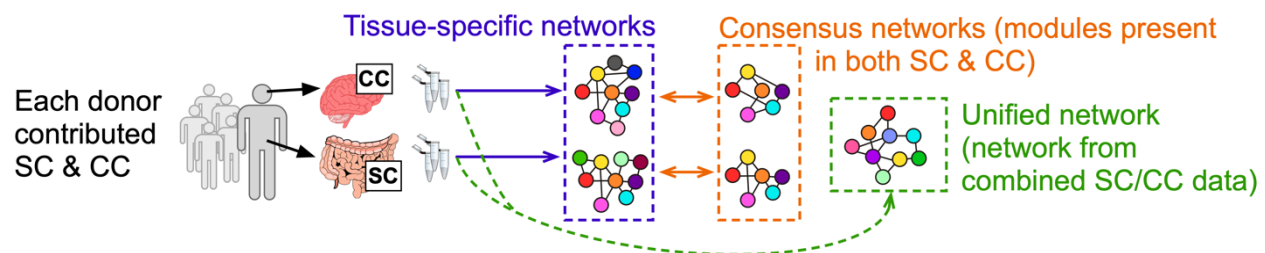


Figure 2: Outline of network analyses

SC, sigmoid colon; CC, cerebral cortex

METHODS

All analysis was performed in R version 3.4 for Mac.(17)

Data acquisition

GTEx data in the form of gene-level counts was obtained from the recount2 database.(18)

Donors with data for both SC and CC were included. CC was used because it resulted in a good sample size and is readily available (in contrast to smaller brain regions). Retrieved counts were scaled with the recount package for R.(18-20) Data were filtered to include genes with counts of 10 or more in at least 30% of samples.

Weighted gene co-expression network analysis (WGCNA)

WGCNA was performed using the WGCNA package in R.(21) First, according to the WGCNA developer's recommendations, counts were prepared for linear modeling and quantile normalized. Preparation for linear modeling was performed with the variance stabilization function in the DESeq2 package.(22) Quantile normalization was performed using the linear models for microarray (limma) package.(23) Genes on the X and Y chromosomes were excluded. Sample clustering was performed with the WGCNA package and outliers were excluded. Parallel analyses were performed with data from lung and CC for comparison. WGCNA using a "signed" network was performed with a soft-thresholding power of 18 and a minimum module size of 80. Highly correlated ($\rho > 0.85$) modules were merged. For the unified analysis, the top 10,000 genes after sorting by variance were combined into one matrix ($n = 20,000$ genes). Prior to combination, probe names were appended with "c" or "b" so that the tissue of origin could be determined. A soft-thresholding power of 12 was used for the unified analysis.

Functional enrichment and visualization

Overrepresented gene ontology (GO) terms were identified with the GOstats package.(24)

Visualization was performed using Cytoscape version 3.5.(25) Edge weights were derived from the Topology Overlap Matrix used in the WGCNA.

Statistical analysis

Association between module eigengenes in the tissue-specific networks was determined with Pearson correlations. P values were adjusted using the Benjamini-Hochberg method and an adjusted-p value threshold of 0.05 was used to determine statistical significance.

Comparison with colonic mucosal biopsy data from IBS and controls

Genes in the CC-correlated SC modules (gut-brain-axis genes) were compared with differential gene expression results from our microarray study (Chapter 2 (26), Chapter 3) and the sigmoid colon biopsy microarray data from the Mayo Clinic, which was used for external comparison in Chapter 2.(26,27) Statistical significance was determined by adjustment of p values (Benjamini-Hochberg) for the gut-brain-axis genes present in the microarray data.

RESULTS

Donor characteristics

There were 34 donors with data from SC and CC. After excluding outliers in either tissue, 29 donors (17 men, 12 women) were included. Distribution of age range and death classifications are shown in Figure 3. Most donors were ages 50-69 with deaths classified as “Fast and natural,” which corresponds to a terminal phase of less than one hour, such as death due to myocardial infarction. RNA quality was good with a median RNA integrity number of 6.4.

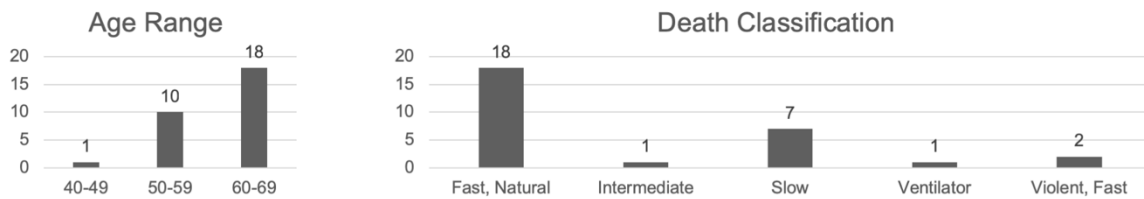


Figure 3: Donor characteristics

Tissue-specific networks

WGCNA uses unsupervised clustering to identify modules of highly correlated genes.(21) Each module can be represented by its first principal component which is referred to as the module eigengene. Genes that are highly correlated with the module eigengene are referred to as hub genes. Clustering of hub genes in the SC and CC networks are shown in Figure 4A and B. Different colors represent different modules and module names are derived from overrepresented GO terms. Hub genes cluster in modules but related modules also cluster, such as the immune related modules (“Response to cytokines”, “NF-kappaB signaling” and “Immune Response”) in the SC network (Figure 4A). There is more intermodular connectivity in the brain likely due to increased homogeneity of cell type. In Figure 4C, the dots represent module eigengenes from the SC (green) and CC (pink) networks. The lines are correlations that are significant ($p < 0.05$) after adjusting for multiple comparisons. Yellow highlighting shows positive correlations between SC and CC module eigengenes.

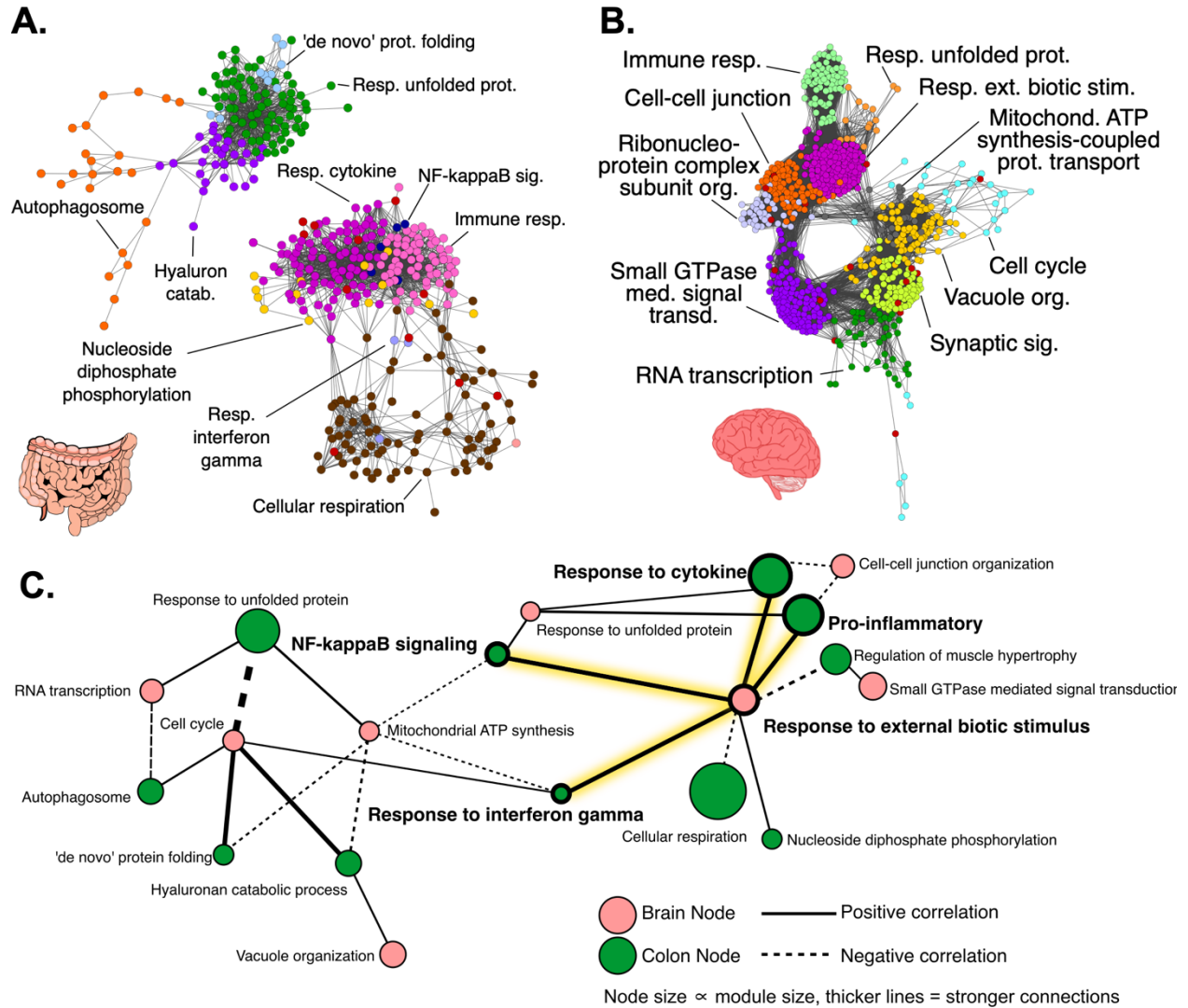


Figure 4: Tissue-specific networks

A and B show clustering of hub genes in sigmoid colon and cerebral cortex. Different modules are in different colors and annotations are overrepresented GO terms. Panel C shows significant (Benjamini-Hochberg adjusted p value <0.05) correlations between module eigengenes in the SC (green) and CC (pink) organ-specific networks.
Prot., protein; *Resp.*, response; *sig.*, signaling; *catab.*, catabolism; *ext.*, external; *stim.*, stimulus; *Mitochond.*, mitochondrial; *org.*, organization; *med.*, mediated; *transd.*, transduction

Consensus Networks

There were fifteen SC-CC consensus modules (present in both tissues). Heatmaps showing correlations between modules (intermodular connectivity) are shown in Figure 5A. Similar to the tissue-specific analysis, there is increased intermodular connectivity in the CC with two main clusters of positively-correlated modules. While the connectivity in the SC is quite different than in the CC, there are modules with similar relationships in both tissues (highlighted by dashed outline). Figure 5B shows consensus modules for lung and CC (n = 19). Comparison of these analyses reveals that consensus modules are specific to tissue pairs. There are more consensus modules in the lung-CC analysis. The reason for this is unclear but perhaps relates to conserved microtubule-related biology in axons and ciliated cells.

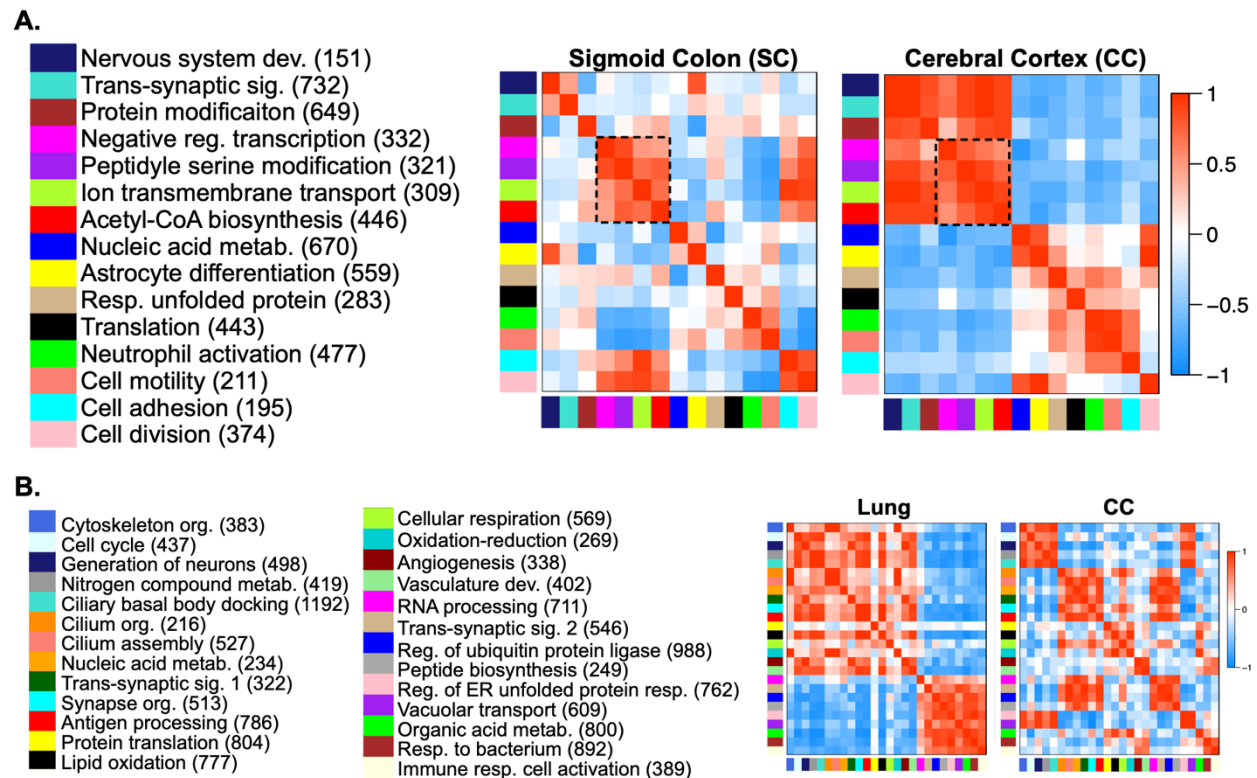


Figure 5: Consensus networks

Intermodular connectivity of SC and CC (A) and lung and CC (B).

SC, sigmoid colon; CC, cerebral cortex; dev., development; sig., signaling; reg., regulation; metab., metabolism; resp., response; org., organization

Unified network

The SC-CC unified network (generated from merged SC and CC data) contained 16 modules of which 5 were comprised predominantly of genes expressed in SC ($\geq 98\%$). Intermodular connectivity is shown in Figure 6A. There was a positive correlation between SC and CC modules with neuro-glial and inflammation-related themes. The lung-CC analysis ($n = 19$) did not reveal the same level of cross-tissue integration in related pathways (Figure 6B). The strongest cross-tissue correlations were between inflammation-related modules. Connectivity of hub genes in these modules is shown in Figure 7. CC expression of the interleukin 4 receptor (IL4R) in the CC inflammatory response module (yellow) was highly correlated with genes in the SC inflammatory response module (purple).

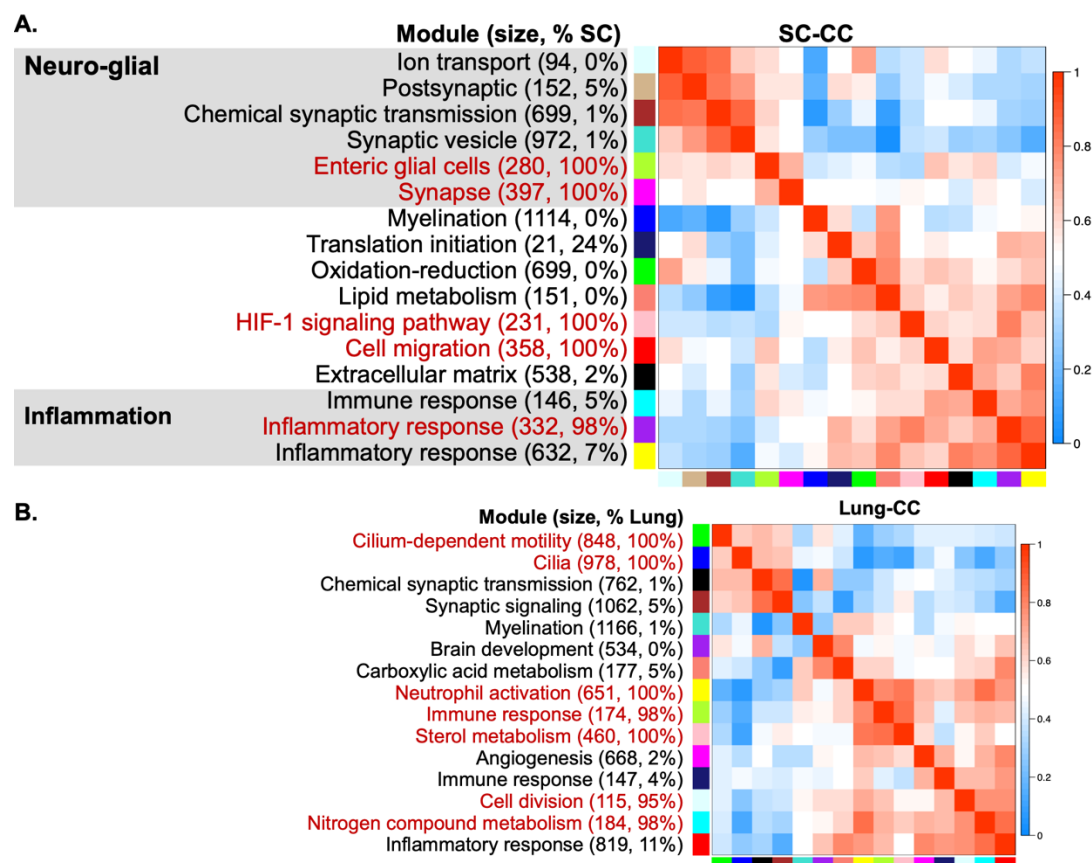


Figure 6: Unified network

SC/lung modules are in red. Module size and % SC or lung (size, %) refer to the number of genes with at least 0.6 correlation with module eigengene. Module names are overrepresented GO terms or KEGG pathways. SC, *sigmoid colon*; CC, *cerebral cortex*

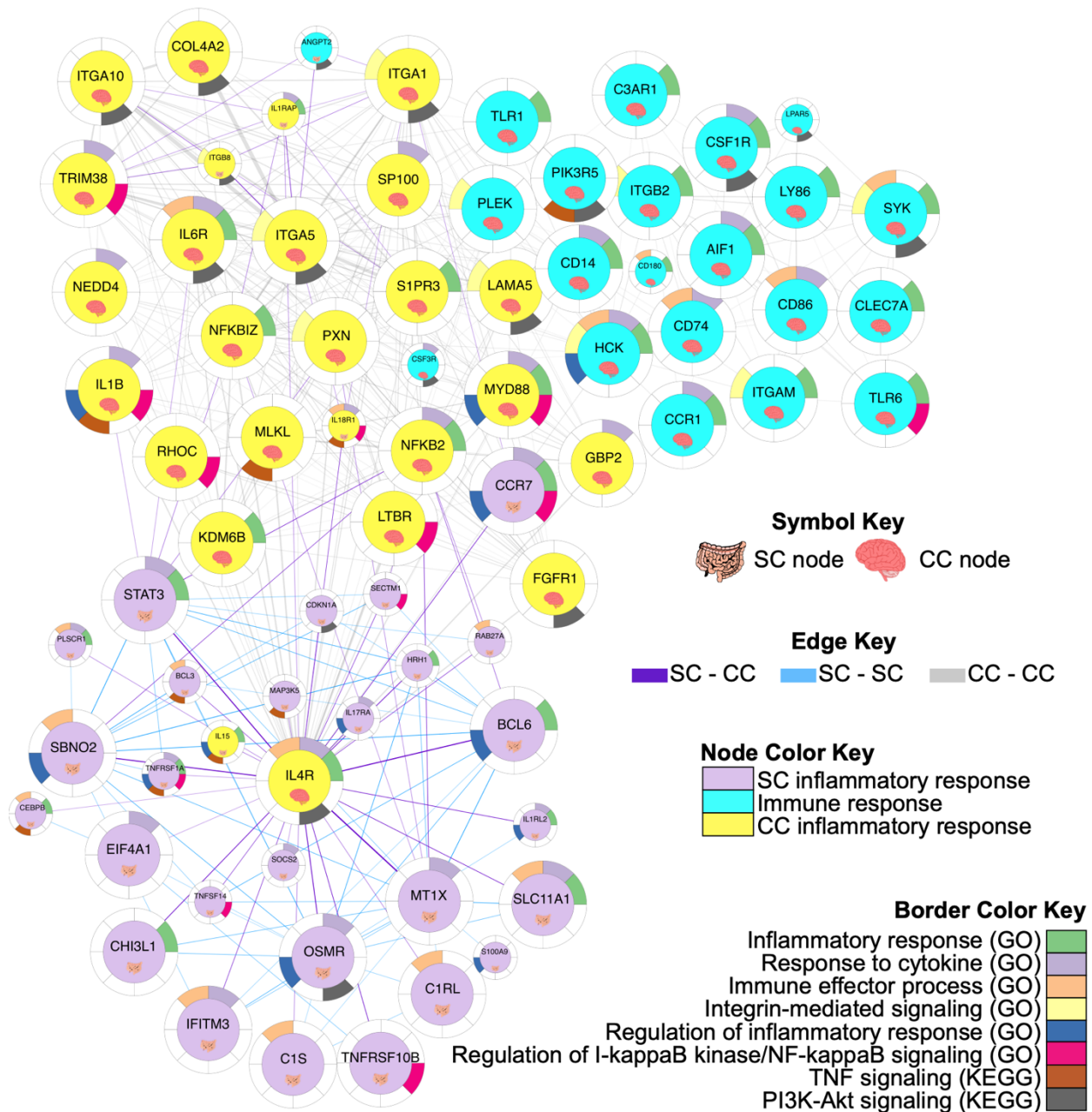


Figure 7: Inflammation hub genes

Networks of hub genes in inflammation-related modules in the unified SC-CC network (Figure 6). Each node represents a single gene. Node colors correspond to modules. Functional annotation, identified by node border colors, was performed using the StringApp in Cytoscape.(28) Larger nodes have higher correlations with the module eigengene. The tissue origin of each node is identified by the colon or brain graphic. Purple edges highlight gut-brain (SC-CC) connectivity. SC-SC and CC-CC edges are blue and grey, respectively. Thicker edges have stronger weight.

SC, sigmoid colon; CC, cerebral cortex

Differential expression of gut-brain-axis genes in IBS

As dysregulation of the gut-brain axis is hypothesized to be central to IBS pathophysiology, differentially expressed genes in IBS present in CC-correlated SC modules would be potential targets for further study. I focused on the inflammation-related modules, as correlation of inflammation-related genes was consistent across analyses. Genes with module membership (correlation with module eigengene) > 0.6 in the SC inflammation-related modules (tissue-specific analysis, Figure 4, “NF-kappaB signaling,” “Response to interferon gamma,” “Response to cytokine,” and “Pro-inflammatory”) were classified as immune gut-brain-axis (GBA) genes. Differential expression was determined for immune GBA genes that were present in our microarray study (Chapter 2 (26), Chapter 3) and/or the sigmoid colon biopsy microarray data from the Mayo Clinic, which was used for external comparison in Chapter 2.(26,27) P values were adjusted for the number of immune GBA genes. Only one immune GBA gene was differentially expressed in our microarray data. The zinc transporter, solute carrier family 39 member 1 (*SLC39A1*) was up-regulated in IBS with diarrhea (IBS-D) vs. healthy controls (fold change = 2.9, adjusted p value = 0.024). In the Mayo clinic data, there were fifteen differentially expressed immune GBA genes. Fourteen were down-regulated in IBS-D vs. healthy controls (Table 2). One gene, heparan sulfate-glucosamine 3-sulfotransferase 2 (*HS3ST2*) was up-regulated in IBS vs. healthy controls (fold change = 1.3, adjusted p value = 0.02). The IBS-D genes (Table 2) were all in three of the four modules (“Response to interferon gamma,” “Response to cytokine,” “Pro-inflammatory”).

Table 2: Immune gut-brain-axis genes differentially expressed in IBS-D (Mayo Clinic cohort)

Symbol	Name	FC (p value)	MM
<i>Response to cytokine</i>			
ADGRE2	adhesion G protein-coupled receptor E2	0.72 (0.02)	0.61
CDK2	cyclin dependent kinase 2	0.87 (0.03)	0.83
HAT1	histone acetyltransferase 1	0.86 (0.04)	0.73
RPL22L1	ribosomal protein L22 like 1	0.83 (<0.005)	0.71
SAMSN1	SAM domain, SH3 domain and nuclear localization signals 1	0.7 (0.02)	0.84
ZNF367	zinc finger protein 367	0.76 (0.02)	0.63
<i>Response to IFN-γ</i>			
CARD16*	caspase recruitment domain family member 16	0.77 (<0.005)	0.64
CASP1	caspase 1	0.71 (<0.005)	0.80
PSMB8-AS1	PSMB8 antisense RNA 1 (head to head)	0.78 (0.03)	0.83
WARS	tryptophanyl-tRNA synthetase	0.8 (0.03)	0.91
<i>Pro-inflammatory</i>			
C1orf162	chromosome 1 open reading frame 162	0.79 (0.04)	0.80
GGCT	gamma-glutamylcyclotransferase	0.89 (0.02)	0.76
RFC2*	replication factor C subunit 2	0.81 (<0.005)	0.79
RPN1	ribophorin I	0.91 (0.03)	0.64

*indicates also differentially expressed in IBS vs. healthy controls. P values are adjusted.
FC, fold change; *MM*, module membership

DISCUSSION

Inflammation-related genes are correlated in the gut and brain

My hypothesis was that a systems biology approach to gene expression profiling data from the gut and the brain could yield insight into the gut-brain axis. In order to test this hypothesis, I used three approaches for analysis of sigmoid colon (SC) and cerebral cortex (CC) gene expression profiling data to address the following aims (Table 1): 1) To identify correlations between SC and CC modules in tissue-specific networks; 2) To identify modules present in both SC and CC using a consensus network; and 3) To identify modules containing both SC and CC genes, and to explore relationships between SC-predominant and CC-predominant modules in a unified network of genes from both tissues. Both the tissue-specific (Aim 1) and unified (Aim 3) analyses highlighted positive correlations between inflammation-related genes in SC and CC. Association between systemic (29) and colonic (30-32,33) inflammation and central nervous system (CNS) neuroinflammation has been well-documented in animal models (described further below). Thus, the identification of positive correlation in inflammation-related genes in the current analyses supports the validity of the analytic approach and confirms the hypothesis that a systems biology approach to gut-brain gene expression data can be used to study the gut-brain axis.

Central effects of peripheral inflammation

There is strong evidence in animal models for the effect of systemic and peripheral inflammation on brain neuroinflammation. In one study, both intra-peritoneal injection of endotoxin as well as skin inflammation induced by Imiquimod resulted in increased transcription of interferon-stimulated genes in the brain.(29) There have also been several studies on CNS effects of intestinal inflammation. Experimental colitis in rodent models resulted in microglial activation and increased cytokine production in the brain.(30-32,33) This was seen in both of the established chemically-induced experimental colitis models: trinitrobenzene sulfonic acid

(TNBS) (30,31) and dextran sodium sulfate (DSS).(32,33) The presence of intestinal inflammation following traumatic brain injury in animal models demonstrates the bidirectionality of this communication.(34)

The effect of peripheral inflammation on the CNS is supported by data in humans. For example, progression of multiple sclerosis can be triggered by infection,(35) and interferon alpha treatment causes symptoms of depression,(36) which has been linked to microglial activation in animal models.(37) In addition, the phenomenon of “sickness behavior,” which describes the psychological correlates of systemic illness, including insomnia, impaired cognition, fatigue, anhedonia and anorexia, is mediated by circulating cytokines, predominantly tumor necrosis factor alpha (TNF- α) and interleukin 1 beta (IL-1 β), and to a lesser extent interleukin 6 (IL-6).(38)

There are three main routes whereby peripheral immune events are communicated to the CNS: neural, humoral and cell-mediated.(39) The neural route is mediated by vagal afferents, which can respond to local cytokines (“vago-vagal inflammatory reflex”).(40) In the humoral route, circulating cytokines signal to the brain via receptors on cerebral endothelial cells or access the brain via regions without a blood brain barrier, such as the circumventricular organs and choroid plexus.(39,41) Cell-mediated communication is triggered by adhesion of circulating monocytes to cerebral endothelial cells, which leads to proinflammatory signaling as well as migration of monocytes to the brain.(39) Interestingly, natalizumab which targets monocyte adhesion by blocking alpha 4 integrin, reduced fatigue in multiple sclerosis (42) and improved health-related quality of life in Crohn’s Disease.(43) Gut microbes also contribute to the gut-brain axis in the setting of inflammation. Increased epithelial permeability facilitates gut-brain signaling via circulating microbial metabolites.(44) There is also evidence that the effect of microbes on the CNS is mediated by the vagus nerve.(45)

Interleukin 4 signaling

CC expression of the interleukin 4 receptor (IL4R) was highly correlated with SC genes in the inflammatory response module. Interleukin-4 (IL-4) is a cytokine produced by Th2 subtype T helper cells, as well as by basophils and mast cells.(46) Th2 cells are a subset of CD4 lymphocytes that mediate the humoral immune response directed against extracellular pathogens.(47) Th2 cells produce cytokines associated with allergic (IL-4, IL-5, IL-13) and anti-inflammatory (IL-10) responses, and which counteract Th1 T helper cells (phagocytic response).(47)

IL-4 plays an important role in mediating the allergic response by inducing B cells to produce IgE immunoglobulins (class-switching).(48) Secreted by Th2 cells, IL-4 propagates the Th2 response by promoting differentiation of lymphocytes into Th2-like cells (48) and up-regulating IL4R, which encodes the alpha subunit of the IL-4 receptor (IL-4R α).(49) IL-4 also promotes recruitment of T cells and eosinophils through up-regulation of vascular cell adhesion molecule - 1,(48) and down-regulation of E-selectin, which is mediated by competitive binding of signal transducer and activator of transcription 6 (STAT6) to NF- κ B sites in the E-selectin promoter.(50) In the intestine, IL-4 is central to allergic response and defense against helminths.(51) It has also been shown to impair epithelial barrier function.(52-54)

IL-4/IL-13 signaling is important in inflammatory bowel disease

IL-13 has been implicated in the barrier dysfunction associated with ulcerative colitis (UC). (55) Both IL-13R and IL-4R are expressed on colonic epithelial cells,(55) and have shared biology due to the use of a common receptor, the IL-13 receptor alpha 1 chain (IL-13R α 1), which heterodimerizes with IL-4R α and forms a complex that can bind both IL-4 or IL-13.(56) IL-4 can also signal through heterodimerization IL-4R α with a gamma chain, which is a component of the IL-2 receptor.(48) Downstream signaling is predominantly via STAT6, but also can involve PI3 kinase.(48) Lamina propria mononuclear cells from UC patients vs. controls and patients with

Crohn's Disease produce increased IL-13.(55) IL-13 producing cells in UC are positive for CD161, a marker of natural killer (NK) T cells.(57) In cell culture models, IL-13 decreased transepithelial resistance and increased paracellular permeability, an effect that involved apoptosis of epithelial cells and dysregulation of tight junctions characterized by increased mRNA and protein expression of claudin-2.(55) This effect was only minimally reduced by addition of an antibody neutralizing IL-4R, however the effect was not induced by IL-4.(55) Tralokinumab, a human monoclonal antibody targeting IL-13, did not provide improve clinical response rates over placebo in UC (though clinical remission and Mayo score were improved),(58) but improved atopic dermatitis symptoms in a Phase 2 trial,(59) and Phase 3 trials are ongoing (NCT03131648, NCT03160885, NCT03363854). Likewise, anrukinzumab, a humanized IgG1 antibody targeting the binding of IL-13 to IL-4R α , was not effective over placebo in UC.(60) While there is not an establish role for epithelial IL-4 signaling in Crohn's disease, one study found IL-4 to be increased in the muscularis externa of Crohn's disease samples, and showed that IL-4 and IL-13 increased intestinal smooth muscle cell contractility via STAT6.(61)

IL-4 is immunoregulatory in CNS and may contribute to CNS effects of peripheral inflammation

In the CNS, both microglial cells and astrocytes produce IL-4 and express IL-4R α .(62) Microglial cells expressing IL-4R α have an M2 (immune-regulatory) phenotype and were found to regulate autoimmune inflammation in an animal model of multiple sclerosis.(63) IL-4 signaling is neuroprotective in CNS injury and ischemia through IL-4R α expressed on neurons,(64) and through neuronal production of IL-4.(65) Conversely, increased expression of IL-4R α has also been associated with neurodegeneration due to increased oxidative stress.(62)

There are several studies in animal models that suggest a role for IL-4 signaling in the CNS response to peripheral inflammation. Peripheral lipopolysaccharide increased microglial expression of IL-4R α ,(66) and depressive behavior in mice treated with interferon alpha was

associated with down-regulation of IL-4R α .(67) In addition, while IL-4 knockout mice, have increased depressive behavior, it is not further modified by interferon alpha.(67)

Additional genes in the SC inflammatory response are linked to IL-4 signaling

Several of the genes in the SC inflammatory response module (bolded text below) are related to IL-4 signaling, which provides additional support for the presence of integrated gut-brain inflammatory signaling. Expression of B-cell CLL/lymphoma 6 (**BCL6**) is induced by IL-4 via STAT6.(68) **BCL6**, in turn, can repress transcription of IL-4/STAT6 induced genes.(69) Signal transducer and activator of transcription 3 (**STAT3**) also increases **BCL6** expression and is involved in T cell repression of IL-4 expression by **BCL6**.(70) Strawberry notch homolog 2 (**SBNO2**) is a transcriptional corepressor of NF- κ B in macrophages, and its expression is induced by the Th2 cytokine IL-10 via **STAT3**.(71) Finally, Chitinase 3 like 1 (**CHI3L1**), which is increased in the circulation in asthma, is secreted by respiratory epithelial cells and macrophages in response to IL-13.(72)

Comparison of WGCNA methods

Both the tissue-specific and unified networks identified the correlation between inflammation-related genes in SC and CC. In the unified network, the modules were almost entirely tissue-specific (modules consisted of either SC or CC genes), so the results of these analyses were quite similar. The advantage of the unified network is that it calculated connectivity between SC and CC genes and allowed the visualization in Figure 7. The consensus network is important because it confirmed that consensus modules were present. This analysis would be particularly useful in the comparison of disease and control groups in gut-brain axis disorders as it would highlight the conserved biology in the gut and brain and allow for comparison of consensus networks in disease and controls.

Dysregulation of immune gut-brain-axis genes in IBS-D

Differential expression of immune gut-brain-axis (GBA) genes was primarily seen in IBS-D vs. healthy controls. Overall, our IBS-D microarray samples were more similar to controls than in other cohorts, which explains why there was only one immune GBA gene differentially expressed in our data while there were fifteen in the Mayo Clinic data. The dysregulation of immune GBA genes in IBS-D, could lead one to hypothesize that gut-brain communication of immune-related phenomena is altered in IBS-D, which could potentially be linked to the increased intestinal permeability commonly seen in this subtype.(73)

Interestingly, two dysregulated immune GBA genes (*CASP1*, *CARD16*) are related to protease activity which is a potential factor in visceral sensitivity.(74) Although not dysregulated in the microarray data, two other genes involved in this pathway, histamine receptor H1 and F2R like thrombin or trypsin receptor 3 (encodes protease activated receptor 4), were among the colon genes closely linked to brain IL4R in the unified network.

Conclusions

A systems biology approach to the analysis of gene expression profiling of the gut and brain is feasible. The ability to do this relies upon careful experimental design that avoids batch effects, such as is the case with data from the GTEx project, which was specifically designed to compare expression across tissues. In this proof of concept analysis in healthy deceased donors, the known connection between peripheral and CNS inflammatory signaling was illustrated. Comparison with gene expression profiling data in IBS, revealed dysregulation of gut-brain axis immune-related genes, specifically in IBS-D. Furthermore, IL-4 signaling was identified as potential target for further investigation of the mechanisms of gut-brain immune-related communication. Interestingly, there is already some experimental support for IL-4 signaling as a mediator of the CNS response to peripheral inflammation.(37,66,67)

REFERENCES:

1. Mayer EA, Tillisch K. The brain-gut axis in abdominal pain syndromes. *Annu Rev Med.* 2011;62:381-396.
2. Kennedy PJ, Clarke G, Quigley EM, Groeger JA, Dinan TG, Cryan JF. Gut memories: towards a cognitive neurobiology of irritable bowel syndrome. *Neurosci Biobehav Rev.* 2012;36(1):310-340.
3. Ludidi S, Conchillo JM, Keszthelyi D, Van Avesaat M, Kruimel JW, Jonkers DM, Masclee AA. Rectal hypersensitivity as hallmark for irritable bowel syndrome: defining the optimal cutoff. *Neurogastroenterol Motil.* 2012;24(8):729-733, e345-726.
4. Mayer EA. The neurobiology of stress and gastrointestinal disease. *Gut.* 2000;47(6):861-869.
5. Grundy D, Al-Chaer ED, Aziz Q, Collins SM, Ke MY, Tache Y, Wood JD. Fundamentals of neurogastroenterology: Basic science. *Gastroenterology.* 2006;130(5):1391-1411.
6. Bennett EJ, Tennant CC, Piesse C, Badcock CA, Kellow JE. Level of chronic life stress predicts clinical outcome in irritable bowel syndrome. *Gut.* 1998;43(2):256-261.
7. Dunlop SP, Jenkins D, Neal KR, Spiller RC. Relative importance of enterochromaffin cell hyperplasia, anxiety, and depression in postinfectious IBS. *Gastroenterology.* 2003;125(6):1651-1659.
8. Gwee KA, Leong YL, Graham C, McKendrick MW, Collins SM, Walters SJ, Underwood JE, Read NW. The role of psychological and biological factors in postinfective gut dysfunction. *Gut.* 1999;44(3):400-406.
9. Posserud I, Agerforz P, Ekman R, Bjornsson ES, Abrahamsson H, Simren M. Altered visceral perceptual and neuroendocrine response in patients with irritable bowel syndrome during mental stress. *Gut.* 2004;53(8):1102-1108.
10. Dickhaus B, Mayer EA, Firooz N, Stains J, Conde F, Olivas TI, Fass R, Chang L, Mayer M, Naliboff BD. Irritable bowel syndrome patients show enhanced modulation of visceral perception by auditory stress. *Am J Gastroenterol.* 2003;98(1):135-143.
11. Hattay P, Prusator DK, Tran L, Greenwood-Van Meerveld B. Psychological stress-induced colonic barrier dysfunction: Role of immune-mediated mechanisms. *Neurogastroenterol Motil.* 2017;29(7).
12. Tache Y, Martinez V, Wang L, Million M. CRF1 receptor signaling pathways are involved in stress-related alterations of colonic function and viscerosensitivity: implications for irritable bowel syndrome. *Br J Pharmacol.* 2004;141(8):1321-1330.
13. Sung J, Morales W, Kim G, Pokkunuri V, Weitsman S, Rooks E, Marsh Z, Barlow GM, Chang C, Pimentel M. Effect of repeated *Campylobacter jejuni* infection on gut flora and mucosal defense in a rat model of post infectious functional and microbial bowel changes. *Neurogastroenterol Motil.* 2013;25(6):529-537.

14. Zanini B, Ricci C, Bandera F, Caselani F, Magni A, Laronga AM, Lanzini A, San Felice del Benaco Study I. Incidence of post-infectious irritable bowel syndrome and functional intestinal disorders following a water-borne viral gastroenteritis outbreak. *Am J Gastroenterol*. 2012;107(6):891-899.
15. Mayer EA, Savidge T, Shulman RJ. Brain-gut microbiome interactions and functional bowel disorders. *Gastroenterology*. 2014;146(6):1500-1512.
16. Carithers LJ, Ardlie K, Barcus M, Branton PA, Britton A, Buia SA, Compton CC, DeLuca DS, Peter-Demchok J, Gelfand ET, Guan P, Korzeniewski GE, Lockhart NC, Rabiner CA, Rao AK, Robinson KL, Roche NV, Sawyer SJ, Segre AV, Shive CE, Smith AM, Sobin LH, Undale AH, Valentino KM, Vaught J, Young TR, Moore HM, Consortium GT. A Novel Approach to High-Quality Postmortem Tissue Procurement: The GTEx Project. *Biopreserv Biobank*. 2015;13(5):311-319.
17. R Core Team. R: A language and environment for statistical computing. 2013. Vienna, Austria: R Foundation for Statistical Computing. Available from: <http://www.R-project.org/>.
18. Collado-Torres L, Nellore A, Kammers K, Ellis SE, Taub MA, Hansen KD, Jaffe AE, Langmead B, Leek JT. Reproducible RNA-seq analysis using recount2. *Nat Biotechnol*. 2017;35(4):319-321.
19. Collado-Torres L, Nellore A, Kammers K, Ellis SE, Taub MA, Hansen KD, Jaffe AE, Langmead B, Leek JT. Explore and download data from the recount project. doi:10.18129/B9.bioc.recount, <https://github.com/leekgroup/recount> - R package version 1.8.2. <http://www.bioconductor.org/packages/recount>. 2017.
20. Collado-Torres L, Nellore A, Jaffe AE. recount workflow: Accessing over 70,000 human RNA-seq samples with Bioconductor. *F1000Res*. 2017;6:1558.
21. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008;9:559.
22. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550.
23. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol*. 2004;3:Article3.
24. Falcon S, Gentleman R. Using GOstats to test gene lists for GO term association. *Bioinformatics*. 2007;23(2):257-258.
25. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498-2504.
26. Videlock EJ, Mahurkar-Joshi S, Hoffman JM, Iliopoulos D, Pothoulakis C, Mayer EA, Chang L. Sigmoid colon mucosal gene expression supports alterations of neuronal signaling in irritable bowel syndrome with constipation. *Am J Physiol Gastrointest Liver Physiol*. 2018;315(1):G140-G157.

27. Aerssens J, Camilleri M, Talloen W, Thielemans L, Gohlmann HW, Van Den Wyngaert I, Thielemans T, De Hoogt R, Andrews CN, Bharucha AE, Carlson PJ, Busciglio I, Burton DD, Smyrk T, Urrutia R, Coulie B. Alterations in mucosal immunity identified in the colon of patients with irritable bowel syndrome. *Clin Gastroenterol Hepatol*. 2008;6(2):194-205.
28. Doncheva NT, Morris JH, Gorodkin J, Jensen LJ. Cytoscape StringApp: Network Analysis and Visualization of Proteomics Data. *J Proteome Res*. 2019;18(2):623-632.
29. Thomson CA, McColl A, Cavanagh J, Graham GJ. Peripheral inflammation is associated with remote global gene expression changes in the brain. *J Neuroinflammation*. 2014;11:73.
30. Riazi K, Galic MA, Kuzmiski JB, Ho W, Sharkey KA, Pittman QJ. Microglial activation and TNFalpha production mediate altered CNS excitability following peripheral inflammation. *Proc Natl Acad Sci U S A*. 2008;105(44):17151-17156.
31. Wang K, Yuan CP, Wang W, Yang ZQ, Cui W, Mu LZ, Yue ZP, Yin XL, Hu ZM, Liu JX. Expression of interleukin 6 in brain and colon of rats with TNBS-induced colitis. *World J Gastroenterol*. 2010;16(18):2252-2259.
32. Zonis S, Pechnick RN, Ljubimov VA, Mahgerefteh M, Wawrowsky K, Michelsen KS, Chesnokova V. Chronic intestinal inflammation alters hippocampal neurogenesis. *J Neuroinflammation*. 2015;12:65.
33. Do J, Woo J. From Gut to Brain: Alteration in Inflammation Markers in the Brain of Dextran Sodium Sulfate-induced Colitis Model Mice. *Clin Psychopharmacol Neurosci*. 2018;16(4):422-433.
34. Sundman MH, Chen NK, Subbian V, Chou YH. The bidirectional gut-brain-microbiota axis as a potential nexus between traumatic brain injury, inflammation, and disease. *Brain Behav Immun*. 2017;66:31-44.
35. Buljevac D, Flach HZ, Hop WC, Hijdra D, Laman JD, Savelkoul HF, van Der Meche FG, van Doorn PA, Hintzen RQ. Prospective study on the relationship between infections and multiple sclerosis exacerbations. *Brain*. 2002;125(Pt 5):952-960.
36. Zdilar D, Franco-Bronson K, Buchler N, Locala JA, Younossi ZM. Hepatitis C, interferon alfa, and depression. *Hepatology*. 2000;31(6):1207-1211.
37. Wachholz S, Esslinger M, Plumper J, Manitz MP, Juckel G, Friebe A. Microglia activation is associated with IFN-alpha induced depressive-like behavior. *Brain Behav Immun*. 2016;55:105-113.
38. Dantzer R, O'Connor JC, Freund GG, Johnson RW, Kelley KW. From inflammation to sickness and depression: when the immune system subjugates the brain. *Nat Rev Neurosci*. 2008;9(1):46-56.
39. D'Mello C, Swain MG. Immune-to-Brain Communication Pathways in Inflammation-Associated Sickness and Depression. *Curr Top Behav Neurosci*. 2017;31:73-94.

40. Cailotto C, Costes LM, van der Vliet J, van Bree SH, van Heerikhuize JJ, Buijs RM, Boeckxstaens GE. Neuroanatomical evidence demonstrating the existence of the vagal anti-inflammatory reflex in the intestine. *Neurogastroenterol Motil.* 2012;24(2):191-200, e193.
41. van Bree SH, Cailotto C, Di Giovangiulio M, Jansen E, van der Vliet J, Costes L, Depoortere I, Gomez-Pinilla PJ, Matteoli G, Boeckxstaens GE. Systemic inflammation with enhanced brain activation contributes to more severe delay in postoperative ileus. *Neurogastroenterol Motil.* 2013;25(8):e540-549.
42. Penner IK, Sivertsdotter EC, Celius EG, Fuchs S, Schreiber K, Berko S, Svenningsson A, investigators Tt. Improvement in Fatigue during Natalizumab Treatment is Linked to Improvement in Depression and Day-Time Sleepiness. *Front Neurol.* 2015;6:18.
43. Targan SR, Feagan BG, Fedorak RN, Lashner BA, Panaccione R, Present DH, Spehlmann ME, Rutgeerts PJ, Tulassay Z, Volfova M, Wolf DC, Hernandez C, Bornstein J, Sandborn WJ, International Efficacy of Natalizumab in Crohn's Disease R, Remission Trial G. Natalizumab for the treatment of active Crohn's disease: results of the ENCORE Trial. *Gastroenterology.* 2007;132(5):1672-1683.
44. Mayer EA, Tillisch K, Gupta A. Gut/brain axis and the microbiota. *J Clin Invest.* 2015;125(3):926-938.
45. Bercik P, Park AJ, Sinclair D, Khoshdel A, Lu J, Huang X, Deng Y, Blennerhassett PA, Fahnestock M, Moine D, Berger B, Huizinga JD, Kunze W, McLean PG, Bergonzelli GE, Collins SM, Verdu EF. The anxiolytic effect of Bifidobacterium longum NCC3001 involves vagal pathways for gut-brain communication. *Neurogastroenterol Motil.* 2011;23(12):1132-1139.
46. Bradding P, Feather IH, Howarth PH, Mueller R, Roberts JA, Britten K, Bews JP, Hunt TC, Okayama Y, Heusser CH, et al. Interleukin 4 is localized to and released by human mast cells. *J Exp Med.* 1992;176(5):1381-1386.
47. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Watson JD. Helper T Cells and Lymphocyte Activation. 2002. In: Molecular Biology of the Cell [Internet]. New York: Garland Science. 4. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK26827/>.
48. Nelms K, Keegan AD, Zamorano J, Ryan JJ, Paul WE. The IL-4 receptor: signaling mechanisms and biologic functions. *Annu Rev Immunol.* 1999;17:701-738.
49. Ohara J, Paul WE. Up-regulation of interleukin 4/B-cell stimulatory factor 1 receptor expression. *Proc Natl Acad Sci U S A.* 1988;85(21):8221-8225.
50. Bennett BL, Cruz R, Lacson RG, Manning AM. Interleukin-4 suppression of tumor necrosis factor alpha-stimulated E-selectin gene transcription is mediated by STAT6 antagonism of NF-kappaB. *J Biol Chem.* 1997;272(15):10212-10219.
51. Van Dyken SJ, Locksley RM. Interleukin-4- and interleukin-13-mediated alternatively activated macrophages: roles in homeostasis and disease. *Annu Rev Immunol.* 2013;31:317-343.

52. Colgan SP, Resnick MB, Parkos CA, Delp-Archer C, McGuirk D, Bacarra AE, Weller PF, Madara JL. IL-4 directly modulates function of a model human intestinal epithelium. *J Immunol.* 1994;153(5):2122-2129.
53. Zund G, Madara JL, Dzus AL, Awtrey CS, Colgan SP. Interleukin-4 and interleukin-13 differentially regulate epithelial chloride secretion. *J Biol Chem.* 1996;271(13):7460-7464.
54. Ceponis PJ, Botelho F, Richards CD, McKay DM. Interleukins 4 and 13 increase intestinal epithelial permeability by a phosphatidylinositol 3-kinase pathway. Lack of evidence for STAT 6 involvement. *J Biol Chem.* 2000;275(37):29132-29137.
55. Heller F, Florian P, Bojarski C, Richter J, Christ M, Hillenbrand B, Mankertz J, Gitter AH, Burgel N, Fromm M, Zeitz M, Fuss I, Strober W, Schulzke JD. Interleukin-13 is the key effector Th2 cytokine in ulcerative colitis that affects epithelial tight junctions, apoptosis, and cell restitution. *Gastroenterology.* 2005;129(2):550-564.
56. Mori S, Maher P, Conti B. Neuroimmunology of the Interleukins 13 and 4. *Brain Sci.* 2016;6(2).
57. Fuss IJ, Heller F, Boirivant M, Leon F, Yoshida M, Fichtner-Feigl S, Yang Z, Exley M, Kitani A, Blumberg RS, Mannon P, Strober W. Nonclassical CD1d-restricted NK T cells that produce IL-13 characterize an atypical Th2 response in ulcerative colitis. *J Clin Invest.* 2004;113(10):1490-1497.
58. Danese S, Rudzinski J, Brandt W, Dupas JL, Peyrin-Biroulet L, Bouhnik Y, Kleczkowski D, Uebel P, Lukas M, Knutsson M, Erlandsson F, Hansen MB, Keshav S. Tralokinumab for moderate-to-severe UC: a randomised, double-blind, placebo-controlled, phase IIa study. *Gut.* 2015;64(2):243-249.
59. Wollenberg A, Howell MD, Guttman-Yassky E, Silverberg JI, Kell C, Ranade K, Moate R, van der Merwe R. Treatment of atopic dermatitis with tralokinumab, an anti-IL-13 mAb. *J Allergy Clin Immunol.* 2019;143(1):135-141.
60. Reinisch W, Panes J, Khurana S, Toth G, Hua F, Comer GM, Hinz M, Page K, O'Toole M, Moorehead TM, Zhu H, Sun Y, Cataldi F. Anrukinzumab, an anti-interleukin 13 monoclonal antibody, in active UC: efficacy and safety from a phase IIa randomised multicentre study. *Gut.* 2015;64(6):894-900.
61. Akiho H, Lovato P, Deng Y, Ceponis PJ, Blennerhassett P, Collins SM. Interleukin-4- and -13-induced hypercontractility of human intestinal muscle cells-implication for motility changes in Crohn's disease. *Am J Physiol Gastrointest Liver Physiol.* 2005;288(4):G609-615.
62. McCormick SM, Heller NM. Commentary: IL-4 and IL-13 receptors and signaling. *Cytokine.* 2015;75(1):38-50.
63. Ponomarev ED, Maresz K, Tan Y, Dittel BN. CNS-derived interleukin-4 is essential for the regulation of autoimmune inflammation and induces a state of alternative activation in microglial cells. *J Neurosci.* 2007;27(40):10714-10721.

64. Walsh JT, Hendrix S, Boato F, Smirnov I, Zheng J, Lukens JR, Gadani S, Hechler D, Golz G, Rosenberger K, Kammertons T, Vogt J, Vogelaar C, Siffrin V, Radjavi A, Fernandez-Castaneda A, Gaultier A, Gold R, Kanneganti TD, Nitsch R, Zipp F, Kipnis J. MHCII-independent CD4⁺ T cells protect injured CNS neurons via IL-4. *J Clin Invest*. 2015;125(2):699-714.
65. Zhao X, Wang H, Sun G, Zhang J, Edwards NJ, Aronowski J. Neuronal Interleukin-4 as a Modulator of Microglial Pathways and Ischemic Brain Damage. *J Neurosci*. 2015;35(32):11281-11291.
66. Fenn AM, Henry CJ, Huang Y, Dugan A, Godbout JP. Lipopolysaccharide-induced interleukin (IL)-4 receptor-alpha expression and corresponding sensitivity to the M2 promoting effects of IL-4 are impaired in microglia of aged mice. *Brain Behav Immun*. 2012;26(5):766-777.
67. Wachholz S, Knorr A, Mengert L, Plumper J, Sommer R, Juckel G, Friebe A. Interleukin-4 is a participant in the regulation of depressive-like behavior. *Behav Brain Res*. 2017;326:165-172.
68. Chevrier S, Kratina T, Emslie D, Tarlinton DM, Corcoran LM. IL4 and IL21 cooperate to induce the high Bcl6 protein level required for germinal center formation. *Immunol Cell Biol*. 2017;95(10):925-932.
69. Dent AL, Shaffer AL, Yu X, Allman D, Staudt LM. Control of inflammation, cytokine expression, and germinal center formation by BCL-6. *Science*. 1997;276(5312):589-592.
70. Wu H, Xu LL, Teuscher P, Liu H, Kaplan MH, Dent AL. An Inhibitory Role for the Transcription Factor Stat3 in Controlling IL-4 and Bcl6 Expression in Follicular Helper T Cells. *J Immunol*. 2015;195(5):2080-2089.
71. El Kasmi KC, Smith AM, Williams L, Neale G, Panopoulos AD, Watowich SS, Hacker H, Foxwell BM, Murray PJ. Cutting edge: A transcriptional repressor and corepressor induced by the STAT3-regulated anti-inflammatory signaling pathway. *J Immunol*. 2007;179(11):7215-7219.
72. Komi DE, Kazemi T, Bussink AP. New Insights Into the Relationship Between Chitinase-3-Like-1 and Asthma. *Curr Allergy Asthma Rep*. 2016;16(8):57.
73. Camilleri M, Gorman H. Intestinal permeability and irritable bowel syndrome. *Neurogastroenterol Motil*. 2007;19(7):545-552.
74. Cenac N, Andrews CN, Holzhausen M, Chapman K, Cottrell G, Andrade-Gordon P, Steinhoff M, Barbara G, Beck P, Bunnett NW, Sharkey KA, Ferraz JG, Shaffer E, Vergnolle N. Role for protease activity in visceral pain in irritable bowel syndrome. *J Clin Invest*. 2007;117(3):636-647.

Chapter 6:

Summary and Future Directions

Summary

The overall goal of my doctoral research was to investigate molecular changes in colon biopsies of IBS patients in comparison to healthy controls using a translational approach. This was accomplished through analysis of gene expression profiling data by microarray, selection of a target gene differentially expressed in IBS and further study of that gene in cellular models.

I began analysis of the microarray data by determining differentially expressed genes and pathways among the protein-coding genes on the array. This analysis (Chapter 2) yielded the most significant findings of my doctoral research, which were published in *The American Journal of Physiology-Gastrointestinal and Liver Physiology*.⁽¹⁾ We found that most differences were in IBS with constipation (IBS-C) vs. healthy controls and that these changes were supportive of alterations in neuronal signaling.

Following the analysis of the mRNAs, I used both data-driven and hypothesis-based approaches to select a lncRNA for further study in cellular models (Chapter 3). The decision to study a lncRNA was based on the presence of epigenetic changes in IBS and our group's experience in studying non-coding RNA. Initial analysis identified ghrelin opposite strand (*GHRLOS*) as a promising target; however, it was not expressed in epithelial cells, and would therefore not serve to provide me with experience in epithelial cell culture models. Another lncRNA, afadin divergent transcript (*AFDN-DT*) was up-regulated in IBS and expressed in colonic epithelial cells and was therefore selected for further study.

Due to the proximity of *AFDN-DT* to afadin, which encodes an adherens junction protein, and was also up-regulated in IBS, we hypothesized that *AFDN-DT* might regulate cell junctions and affect epithelial permeability, which is consistently altered in IBS. Chapter 4 describes the study of *AFDN-DT* in colonic epithelial cells. I developed cell lines capable of inducible overexpression of *AFDN-DT*, and I tested the effect of *AFDN-DT* overexpression on transepithelial electrical resistance and transcellular permeability. I also used RNA sequencing to determine the downstream effectors of *AFDN-DT*.

Despite successful overexpression of *AFDN-DT* in differentiated epithelial cells, there was no effect on permeability or gene expression. This may have been due to a principal challenge inherent in genetic manipulation of lncRNAs, which is that the genomic context of transcription is not fully replicated in the model system. It is also possible, that *AFDN-DT* is a transcriptional by-product without an independent function or that the effects are only evident in a tissue setting where different cell types are present.

Even though we were not able to identify a functional role for *AFDN-DT* in epithelial cells, my primary educational objective to complete a translational project were met. Specifically, I identified a molecular target in clinical samples and studied it in cell culture models. The fact that *AFDN-DT* did not have the anticipated effect actually increased the educational value of this project as it unmasked changes in cellular physiology due to off-target effects of common cell culture methodologies, such as lentiviral transfection and antibiotic treatment. It also highlighted the challenges in using transgenic approaches to model dysregulation of lncRNAs. One of the most valuable aspects of this project was the opportunity to become proficient in the analysis of high-throughput gene expression profiling data.

Following my doctoral studies, I intend to continue to study disorders of the gut-brain axis. In particular, I am interested in studying the gut-brain axis at the molecular level. In Chapter 5, I analyzed publicly-available RNA-sequencing data from brain and colon tissue of healthy deceased donors, which highlighted the correlation of inflammation-related genes in the gut and the brain. These findings will support future independent research into the role of colonic inflammation in Parkinson's Disease.

REFERENCES:

1. Videlock EJ, Mahurkar-Joshi S, Hoffman JM, Iliopoulos D, Pothoulakis C, Mayer EA, Chang L. Sigmoid colon mucosal gene expression supports alterations of neuronal signaling in irritable bowel syndrome with constipation. *Am J Physiol Gastrointest Liver Physiol.* 2018;315(1):G140-G157.

Appendix:

R code

Chapter 2

Microarray Differential Expression

Data set-up

Expression data

Download: <https://www.ebi.ac.uk/arrayexpress/files/E-MTAB-5811/E-MTAB-5811.processed.3.zip>

Expression data is pre-processed by Agilent feature extraction.

```
dat01<-read.delim("Normalized_unfiltered_mRNAonly.txt",row.names = 1)
```

Sample data

Download: <https://www.ebi.ac.uk/arrayexpress/files/E-MTAB-5811/E-MTAB-5811.sdrf.txt>

```
pheno<-read.delim("E-MTAB-5811.sdrf.txt",row.names = 1)
```

reorder/reformat

```
pheno<-pheno[names(dat01),]  
pheno$Sex<-ifelse(pheno$Characteristics.sex=="male","M","F")  
pheno$IBS<-ifelse(pheno$Characteristics.disease=="irritable bowel syndrome","IBS","HC")  
pheno$Group[pheno$Characteristics.clinical.history=="healthy control"]<-"HC"  
pheno$Group[pheno$Characteristics.clinical.history=="constipation"]<-"C"  
pheno$Group[pheno$Characteristics.clinical.history=="diarrhea"]<-"D"  
pheno<-pheno[,c("IBS","Group","Sex")]
```

Array Data

Download: <https://www.ebi.ac.uk/arrayexpress/files/A-MTAB-619/A-MTAB-619.adf.txt>

remove lines above column names

```
datArray<-read.delim("A-MTAB-619.adf.txt",row.names = 1)  
names(datArray)<-c("Seq","EntrezID")
```

match expression data, add symbol and gene name

```
datArray<-datArray[rownames(dat01),]
```

```
library(annotate)  
library(org.Hs.eg.db)
```

```
datArray$Symbol<-NA;datArray$Name<-NA  
datArray2<-datArray[!is.na(datArray$EntrezID),]  
datArray2$Symbol <- getSYMBOL(as.character(datArray2$EntrezID),"org.Hs.eg.db")  
datArray2$Name <- as.character(lookup(as.character(datArray2$EntrezID),  
                                     "org.Hs.eg.db", "GENENAME"))
```

```
datArray3<-rbind(datArray[!rownames(datArray)%in%rownames(datArray2),],datArray2)
```

```
# match expression data
```

```
datArray3<-datArray3[rownames(dat01),c("Symbol","Name","EntrezID")]
```

compile expression set

```
library(Biobase)  
phenoData<-new("AnnotatedDataFrame", data=pheno)  
UCmRNAeset<- new("ExpressionSet", exprs=as.matrix(dat01), phenoData=phenoData)  
fData(UCmRNAeset)<-datArray3
```

Differential expression

```
library(limma)
```

convenience function to edit topTable (limma output)

```
# adds means, direction, FC, variable that is direction + Entrez ID  
top.revision<-function(toptable,eset,Group1,Group2){  
  dat1<-toptable  
  dat1$Direction[dat1$logFC>0]<-"Up"; dat1$Direction[dat1$logFC<0]<-"Down"  
  dat1$FC <- 2^(dat1$logFC); dat1$DirID<-paste(dat1$Direction,dat1$EntrezID)  
  eset1<-eset[rownames(dat1),eset$Group%in%Group1]  
  eset2<-eset[rownames(dat1),eset$Group%in%Group2]  
  dat1$mean1 <- rowMeans(exprs(eset1));dat1$mean2 <- rowMeans(exprs(eset2))
```

```

    return(dat1)
}

```

Run limma

```

f<-as.factor(UCmRNAeset$Group)
design <- model.matrix(~ 0 + f) #factorial design
colnames(design)<-levels(f)
contrasts <- makeContrasts((D+C)/2-HC, D-HC, C-HC, levels=design)
fitUC<-lmFit(UCmRNAeset,design)
fitUC <- contrasts.fit(fitUC, contrasts)
fitUC <- eBayes(fitUC) #empirical Bayes moderated t-statistics test
coefficients<-colnames(fitUC$coefficients)

```

IBS vs HC

```

Group1<-c("D", "C");Group2<-c("HC")
Cont<-coefficients[1]
UC.IBSvHCall <- topTable(fitUC, coef=Cont, adjust.method="BH", number=Inf)
UC.IBSvHCall<-top.revision(UC.IBSvHCall,UCmRNAeset,Group1,Group2)
#filter by raw mean > 200 in either group
filter<-UC.IBSvHCall$mean1>=log2(200)|UC.IBSvHCall$mean2>=log2(200)
UC.IBSvHC.200<-UC.IBSvHCall[filter,]

```

repeated for IBS-D vs HC, IBS-C vs HC

Where there are redundant matches, retain by probe with highest variance

ReAnnotate was used to align probes to exome ranges Process output from reAnnotate

```

options(stringsAsFactors = F)
reAnnoAgil<-read.delim("probesAgil_exome_readAnnotation.txt",stringsAsFactors = FALSE)
hits.agil<-strsplit(reAnnoAgil$HIT,"\\;")
names(hits.agil)<-reAnnoAgil$X.PROBE_ID
for (Probe in names(hits.agil)) {hits.agil[[Probe]]$AgilProbe<-Probe}
hit.agilTable<-do.call(rbind,hits.agil)

```

review of alignments revealed that most probes were specific to one or more transcript

```

hit.conflicts<-read.csv("hitAgilTableConflicts.csv",row.names = 1)

ConflictProbes<-as.character(na.omit(
  hit.conflicts$AgilProbe[hit.conflicts$NonSpecific==1])) #43

AgilMatched<-data.frame(row.names = hit.conflicts$AgilProbe,
  RefSeq_ReAnno=hit.conflicts$RefSeq2,
  Symb_ReAnno=hit.conflicts$Symb2)
for (Probe in names(hits.agil[!names(hits.agil)%in%rownames(AgilMatched)])) {
  df.probe<-hits.agil[[Probe]]
  AgilMatched[Probe,"RefSeq_ReAnno"]<-paste(unique(df.probe$RefSeq),collapse = "/")
  AgilMatched[Probe,"Symb_ReAnno"]<-paste(unique(df.probe$Symb),collapse = "/")
}

```

Retain by highest variance among nonspecific probes

```

library(WGCNA)
options(stringsAsFactors = FALSE)

# subset expression data
datExpr<-exprs(UCmRNAeset)[rownames(exprs(UCmRNAeset))%in%ConflictProbes,]

# collapse rows by max variance----
ProbeKey<-read.csv("Nonspecific.csv")
datExpr2<-collapseRows(datExpr,rowGroup = ProbeKey$Symb,
  rowID = ProbeKey$AgilProbe, method = "maxRowVariance")
ExcludeProbes<-names(datExpr2$selectedRow[datExpr2$selectedRow==F])
# Load("UC_CvHCToptables.rda")
UCfilt<-UC.CvHCall$adj.P.Val<=0.05 &
  (UC.CvHCall$mean1>=log2(200)|UC.CvHCall$mean2>=log2(200))
UC<-UC.CvHCall[UCfilt,] #1270
UC.maxVar<-UC[!rownames(UC)%in%ExcludeProbes,] #1248

for (Probe in rownames(UC.maxVar)) {
  if(Probe%in%rownames(AgilMatched)){

```

```

    UC.maxVar[Probe,"RefSeq_ReAnno"]<-AgilMatched[Probe,"RefSeq_ReAnno"]
    UC.maxVar[Probe,"Symb_ReAnno"]<-AgilMatched[Probe,"Symb_ReAnno"]
  }
}

# 2 of top 100 only mapped to non-coding sequences
UC.filteredVar_noNC<-UC.filteredVar[
  !UC.filteredVar$RefSeq_ReAnno%in%c("NR_046338","NR_147125"),]

```

Heatmap

```

library(gplots)
load("Top100Desc.rda") #contains gene descriptions for row labels

esetTop100<-UCmRNAeset[rownames(Top100Desc),]
dim(esetTop100) #100 x 30

# re-order subjects by clustering within group
ReorderSubjects<-function(eset){
  hc<-eset$Group=="HC"; d<-eset$Group=="D"; c<-eset$Group=="C"
  eset.hc<-eset[,hc]; eset.d<-eset[,d]; eset.c<-eset[,c]
  getInd<-function(eset.group){
    data1<-t(exprs(eset.group))
    dend<-as.dendrogram(hclust(dist(data1)))
    dend <- reorder(dend, rowMeans(data1))
    Ind<-rev(order.dendrogram(dend))
    return(Ind)
  }
  subs.c<-rownames(pData(eset.c[,getInd(eset.c)]))
  subs.d<-rownames(pData(eset.d[,getInd(eset.d)]))
  subs.hc<-rownames(pData(eset.hc[,getInd(eset.hc)]))
  subs<-c(subs.c,subs.d,subs.hc)
  eset2<-eset[,subs]
}

UC.grouped<-ReorderSubjects(esetTop100)
colors<-pData(UC.grouped)
colors$color[colors$Group=="HC"]<-"black"
colors$color[colors$Group=="C"]<-"red"
colors$color[colors$Group=="D"]<-"green"
patientcolorsUC_grouped<-as.character(colors$color)

exprs.dat<-exprs(UC.grouped)

pdf("Top100.pdf", width = 3.5, height = 7.7)
heatmap.2(exprs.dat, scale = "row", Colv = NA, dendrogram = "none",
  ColSideColors=patientcolorsUC_grouped,
  trace="none", col = topo.colors(50),
  lmat=rbind(c(0,4),c(0,1),c(3,2),c(0,5)),
  lwid = c(0.01,.5), lhei = c(0.01,.03, 2,.08),
  labRow = paste(Top100Desc$Symbol,Top100Desc$Description,sep = ", " ) ,
  labCol = FALSE, cexRow = .7 ,
  margins = c(0,17), key.title = NA, density.info = "none",
  offsetRow = 0, keysize = 0.25, key.par = list(mar=c(3.2,0,0.5,37.5),cex=0.3),
  key.xtickfun = function() {
    breaks = pretty(parent.frame())$breaks
    breaks = seq(breaks[1],breaks[length(breaks)],2)
    list(at = parent.frame()$scale01(breaks),
         labels = breaks)
  }
)
dev.off()

```

Sensitivity analysis for sex

```

xy<-fData(UCmRNAeset)
xy<-xy[!is.na(xy$EntrezID),]
xy$chr<-as.character(lookup(as.character(xy$EntrezID), "org.Hs.eg.db", "CHR"))
xy<-xy[xy$chr%in%c("X","Y"),]

```

```
noxy<-setdiff(rownames(exprs(UCmRNAeset)),rownames(xy))
eset21<-UCmRNAeset[noxy,]
```

Repeat fit excluding x and y

```
f<-as.factor(eset21$Group)
design <- model.matrix(~ 0 + f)
colnames(design)<-levels(f)
contrasts <- makeContrasts((D+C)/2-HC, D-HC, C-HC, levels=design)
fitUC2<-lmFit(eset21,design)
fitUC2 <- contrasts.fit(fitUC2, contrasts)
fitUC2 <- eBayes(fitUC2)
coefficients<-colnames(fitUC2$coefficients)
Group1<-c("C");Group2<-c("HC")
Cont<-coefficients[3]
UC.CvHCall2 <- topTable(fitUC2, coef=Cont, adjust.method="BH", number=Inf)
UC.CvHCall2<-top.revision(UC.CvHCall2,eset21,Group1,Group2)
#filter by raw mean > 200 in either group
filter<-UC.CvHCall2$mean1>=log2(200)|UC.CvHCall2$mean2>=log2(200)
UC.CvHC.200.2<-UC.CvHCall2[filter,]
UC.CvHC.filtered2<-UC.CvHC.200.2[UC.CvHC.200.2$adj.P.Val<=0.05,] #1146 rows
```

Collapse by max variance

```
datExpr<-exprs(UCmRNAeset)[rownames(exprs(UCmRNAeset))%in%ConflictProbes,]
# collapse rows by max variance----
ProbeKey<-read.csv("Nonspecific.csv")
datExpr2<-collapseRows(datExpr,rowGroup = ProbeKey$Symb,
                        rowID = ProbeKey$AgilProbe, method = "maxRowVariance")
ExcludeProbes<-names(datExpr2$selectedRow[datExpr2$selectedRow==F])
UC.maxVar2<-UC.CvHC.filtered2[!rownames(UC.CvHC.filtered2)%in%ExcludeProbes,] #1204
```

Make expression set just of these 1204 DETs and control for sex within these DEGs

```
esetsig<-eset21[rownames(UC.maxVar2),]
BH<-factor(esetsig$Group,levels = c("HC","C","D"))
Sex<-factor(esetsig$Sex,levels = c("M","F"))
design <- model.matrix(~Sex*BH) # Sex+BH+Sex*BH
colnames(design)[2:4]<-c("SexF","C","D")
fit_BH.sex<-lmFit(esetsig,design)
fit_BH.sex <- eBayes(fit_BH.sex)
```

5 with significant sex effect

```
SigSex <- topTable(fit_BH.sex, coef="SexF", adjust.method="BH",p.value = 0.05) # 5 rows
```

2 with significant interaction effect (these overlap with the 5 with significant sex effect)

```
SigInt <- topTable(fit_BH.sex, coef="SexF:BHC", adjust.method="BH",p.value = 0.05) # 2
```

For two with significant Sex and interaction effects, C Vs HC effect remains significant controlling for Sex interaction

```
SigC<-topTable(fit_BH.sex, coef="C", adjust.method="BH",number = Inf)
SigC[rownames(SigInt),]
```

Exclude interaction effect for remainder (was not significant in other 3)

```
design <- model.matrix(~Sex+BH)
colnames(design)[2:4]<-c("SexF","C","D")
fit_BH.sex<-lmFit(esetsig,design)
fit_BH.sex <- eBayes(fit_BH.sex)
```

For all 5 with significant sex effect, effect of CvHC remains significant after controlling for sex

```
SigC2 <- topTable(fit_BH.sex, coef="C", adjust.method="BH",number = Inf)
SigC2[rownames(SigSex),]
```

WGCNA

Download processed data: <https://www.ebi.ac.uk/arrayexpress/files/E-MTAB-5811/E-MTAB-5811.processed.2.zip>

Set-up

```
dat01<-read.delim("Normalized_SignalFiltered_noXY.txt",row.names = 1)
mean.filter<-rowMeans(dat01)>=log2(200)
dat01<-dat01[mean.filter,]
datExpr<-as.data.frame(t(dat01))
```

Create modules

Choose soft thresholding power

see: <https://labs.genetics.ucla.edu/horvath/CoexpressionNetwork/Rpackages/WGCNA>

```
library(WGCNA)
options(stringsAsFactors = FALSE)
allowWGCNAThreads() #if computer allows multi-threading
adjacency = adjacency(datExpr, power = 14)
TOM = TOMsimilarity(adjacency) #takes very Long (30-60 min depending on RAM)
dissTOM = 1-TOM
geneTree = hclust(as.dist(dissTOM), method = "average");
rm(adjacency)
```

Plot clustering tree (dendrogram)

```
plot(geneTree, xlab="", sub="",
     main = "Gene clustering on TOM-based dissimilarity", labels = FALSE, hang = 0.04)
```

Identify modules using dynamic tree cut:

```
minModuleSize = 30;
dynamicMods = cutreeDynamic(dendro = geneTree,
                           distM = dissTOM, deepSplit = 2,
                           pamRespectsDendro = FALSE,
                           minClusterSize = minModuleSize)
```

Convert numeric labels to colors

```
dynamicColors = labels2colors(dynamicMods)
```

Plot the dendrogram and colors

```
plotDendroAndColors(geneTree, dynamicColors,
                   "Dynamic Tree Cut", dendroLabels = FALSE, hang = 0.03,
                   addGuide = TRUE, guideHang = 0.05,
                   main = "Gene dendrogram and module colors")
```

Merging similar modules

```
MEList = moduleEigengenes(datExpr, colors = dynamicColors)
MEs = MEList$eigengenes
```

Calculate dissimilarity of module eigengenes

```
MEDiss = 1-cor(MEs)
```

Cluster module eigengenes

```
METree = hclust(as.dist(MEDiss), method = "average");
```

Plot the result

```
plot(METree, main = "Clustering of module eigengenes", xlab = "", sub = "")
```

height cut of 0.25, corresponding to correlation of 0.75:

```
MEDissThres = 0.25
plot(METree, main = "Clustering of module eigengenes", xlab = "", sub = "")+
abline(h=MEDissThres, col = "red")
```

Merge correlated modules

```
merge = mergeCloseModules(datExpr, dynamicColors, cutHeight = MEDissThres, verbose = 3)
mergedColors = merge$colors; mergedMEs = merge$newMEs

plotDendroAndColors(geneTree, cbind(dynamicColors, mergedColors),
  c("Dynamic Tree Cut", "Merged dynamic"), dendroLabels = FALSE,
  hang = 0.03, addGuide = TRUE, guideHang = 0.05)
```

Rename to moduleColors

```
moduleColors = mergedColors
```

Create numerical labels for colors

```
colorOrder = c("grey", standardColors(50));
moduleLabels = match(moduleColors, colorOrder)-1;
MEs = mergedMEs
```

Relate modules to clinical traits

```
load("UCmRNAeset.rda")
datTraits<-pData(UCmRNAeset)
datTraits$IBS<-ifelse(datTraits$IBS=="IBS",1,0)
datTraits$Sex<-ifelse(datTraits$Sex=="F",1,0)
datTraits$C[datTraits$Group=="HC"]<-0; datTraits$C[datTraits$Group=="C"]<-1
datTraits$D[datTraits$Group=="HC"]<-0; datTraits$D[datTraits$Group=="D"]<-1
datTraits<-datTraits[,c("IBS","C","D","Sex")]
nGenes = ncol(datExpr); nSamples = nrow(datExpr)
MEs0 = moduleEigengenes(datExpr, moduleColors)$eigengenes
MEs = orderMEs(MEs0)
moduleTraitCor = cor(MEs, datTraits, use = "p")
moduleTraitPvalue = corPvalueStudent(moduleTraitCor, nSamples)
textMatrix = paste(signif(moduleTraitCor, 2),
  "\n(", signif(moduleTraitPvalue, 1), ")", sep = "")
dim(textMatrix) = dim(moduleTraitCor)

pdf(file = "wgcnaplot.pdf", height = 16)
par(mar = c(6, 8.5, 3, 3));
labeledHeatmap(Matrix = moduleTraitCor, textMatrix = textMatrix,
  xLabels = names(datTraits), yLabels = names(MEs),
  ySymbols = names(MEs), colorLabels = FALSE,
  colors = blueWhiteRed(50), setStdMargins = FALSE,
  cex.text = 0.5, zlim = c(-1,1),
  main = paste("Module-trait relationships"))
dev.off()
```

Comparison with external data

Data Acquisition: Mayo

ten samples are in triplicate with one done 3 months later. Per reference, these were well correlated (no change in expression over 3 months). I will use all available, because there is no annotation as to which were the ones done later

```
library(ArrayExpress)
library(affy)
library(hgu133plus2.db)
Mayo <- ArrayExpress("E-TABM-176")
```

Sample data

```
phMayo <- pData(Mayo)

# remove extra columns
phMayo <- phMayo[,c(3,5:6)]

# rename (IBS is IBS/HC, Group is C/D/HC)
colnames(phMayo)<-c("ID", "IBS", "Group")

# rename
phMayo$IBS[phMayo$IBS=="irritable bowel syndrome"] <- "IBS"
```



```

phMayo$IBS[phMayo$IBS=="normal"] <- "HC"
phMayo$Group[phMayo$Group=="diarrhea-predominant"] <- "D"
phMayo$Group[phMayo$Group=="constipation-predominant"] <- "C"
phMayo$Group[phMayo$Group=="not applicable"] <- "HC"
Mayo@phenoData@data <- phMayo # add back new pheno table

# shorten row and column names
rownames(pData(Mayo))<- gsub(".Hybridization","",
                             gsub("jaerssen.prdbe.jnj.com.Tab2MAGE","",
                                   rownames(pData(Mayo))))

colnames(exprs(Mayo))<- gsub(".Hybridization","",
                             gsub("jaerssen.prdbe.jnj.com.Tab2MAGE","",
                                   colnames(exprs(Mayo))))

```

Data are not preprocessed

```

dim(Mayo)
# Features Samples
# 1354896      131

```

Process, normalize, log transform with RMA

```

MayoNorm <- rma(Mayo)
dim(MayoNorm)
# Features Samples
# 54675      131

genes<-data.frame(row.names = rownames(exprs(MayoNorm)))
library(annotate)
genes$Symbol <- getSYMBOL(as.character(rownames(genes)), "hgu133plus2.db")
genes$Name <- as.character(lookup(as.character(rownames(genes)),
                                "hgu133plus2.db", "GENENAME"))
genes$EntrezID <- as.character(lookup(as.character(rownames(genes)),
                                    "hgu133plus2.db", "ENTREZID"))
fData(MayoNorm)<-genes

```

Data Acquisition: Nottingham

```

Nott <- ArrayExpress("E-GEOD-36701")
phNott <- Nott@phenoData@data
phNott <- phNott[,33:36]
colnames(phNott) <- c("Batch", "Group", "Sex", "ID")
phNott$Group[phNott$Group=="HV"]<-"HC"
phNott$Group[phNott$Group=="IBS-D"]<-"D"
phNott$Group[phNott$Group=="IBS-C"]<-"C"
phNott$Group[phNott$Group=="PIBD"]<-"PIBSCamp"

# The data are not pre-processed
dim(Nott)
#Features Samples
#1354896      221

```

Process, normalize, log transform with RMA

```

NottNorm <- rma(Nott)
dim(NottNorm)
#Features Samples
#54675      221

```

The experiment is described as two batches. Compare mean expression from the batches to see if experimental groups across batches can be compared. (Code not included but I also did without combining batches and not much different)

```

dataAll<-exprs(NottNorm)
rowmeans <- rowMeans(t(dataAll))
ph1 <- cbind(phNott,rowmeans)
library(ggplot2)
ggplot(ph1,aes(Batch,rowmeans))+geom_bar(stat = "summary", fun.y = "mean")
# Means are similar

```

Make expression set

```
# there is a punctuation mark difference in the sample names between expression data
# and pheno data (they are the same but this makes them identical)
colnames(dataAll)<-rownames(phNott)
genes<-data.frame(row.names = rownames(dataAll))
genes$Symbol <- getSYMBOL(as.character(rownames(genes)), "hgu133plus2.db")
genes$Name <- as.character(lookup(as.character(rownames(genes)),
                                "hgu133plus2.db", "GENENAME"))
genes$EntrezID <- as.character(lookup(as.character(rownames(genes)),
                                    "hgu133plus2.db", "ENTREZID"))
pData(NottNorm)<-phNott; fData(NottNorm)<-genes
```

Differential expression: Mayo

this uses the duplicateCorrelation function to average replicates. I also used means with similar results. note that duplicateCorrelation can take ~30 min

```
f <- as.factor(MayoNorm$Group)
design <- model.matrix(~ 0 + f)
colnames(design)=levels(f)
subID<-MayoNorm$ID
cor <- duplicateCorrelation(MayoNorm, design, block = subID)
fitMayo <- lmFit(MayoNorm, design, block = subID, correlation = cor$consensus.correlation)
contrasts <- makeContrasts((D+C)/2-HC, D-HC, C-HC, levels=design)
fitMayo <- contrasts.fit(fitMayo, contrasts)
fitMayo <- eBayes(fitMayo)
coefficients<-colnames(fitMayo$coefficients)

Group1<-c("C");Group2<-c("HC");Cont<-coefficients[3]
MayoCvHCdupCorAll<- topTable(fitMayo, coef=Cont, adjust.method="BH", number=Inf)
MayoCvHCdupCorAll<-top.revision(MayoCvHCdupCorAll,MayoNorm,Group1,Group2)
```

Differential Expression: Nottingham

this averages healthy controls over two batches. also did using only batch 1 which compared IBS-C, IBS-D and HC and similar results

```
f <- as.factor(NottNorm$Group)
design <- model.matrix(~ 0 + f)
colnames(design)=levels(f)
subID<-NottNorm$ID
cor <- duplicateCorrelation(NottNorm, design, block = subID)
fitNott <- lmFit(NottNorm, design, block = subID,
                correlation = cor$consensus.correlation)
contrasts <- makeContrasts((D+C+PIBS+PIBSCamp)/4-HC, (D+C)/2-HC,D-HC,C-HC,levels=design)
fitNott <- contrasts.fit(fitNott, contrasts)
fitNott <- eBayes(fitNott)
coefficients<-colnames(fitNott$coefficients)

Group1<- "C";Group2<- "HC";Cont<-coefficients[4]
NottCvHCdupCorAll <- topTable(fitNott, coef=Cont, adjust.method="BH", number=Inf)
NottCvHCdupCorAll<-top.revision(NottCvHCdupCorAll,NottNorm,Group1,Group2)
```

Find common genes

ReAnnotate was used to align UC probes to exome ranges. Affymetrix annotation (2016) downloaded from ThermoFisher

```
options(stringsAsFactors = F)

affy<-read.csv("HG-U133_Plus_2.na36.annot copy.csv")
affy<-affy[!affy$RefSeq.Transcript.ID=="---",]
hits.affy0<-strsplit(affy$RefSeq.Transcript.ID, " /// ")
names(hits.affy0)<-affy$Probe.Set.ID
hits.affy<-list()
for (Probe in names(hits.affy0)){
  x<-hits.affy0[[Probe]]
  Remove<-c(x[grep("XM",x)],x[grep("XR",x)],x[grep("NR",x)])
  x<-x[!x%in%Remove]
  if(length(x)>0){hits.affy[[Probe]]<-x}
```

```

}

affyUniqueTx<-data.frame(AffyProbe=character(0),RefSeq=character(0))
affyMultiTx<-list()

for (Probe in names(hits.affy)) {
  x<-hits.affy[[Probe]]
  if(length(x)==1){
    affyUniqueTx<-rbind(affyUniqueTx,c(Probe,x))
  }else{affyMultiTx[[Probe]]<-x}
}
names(affyUniqueTx)<-c("AffyProbe","RefSeq")

```

Find affy probes matching by common RefSeqIDs

```

agil.RefSeqs0<-lapply(hits.agil, function(x){x[, "RefSeq"]})
agil.RefSeqs<-list()
for (Probe in names(agil.RefSeqs0)) {
  x<-agil.RefSeqs0[[Probe]]
  Remove<-x[grepl("NR",x)]
  x<-x[!x%in%Remove]
  if(length(x)>0){agil.RefSeqs[[Probe]]<-x}
}

AgilUniqueTx<-data.frame(AgilProbe=character(0),RefSeq=character(0))
AgilMultiTx<-list()

for (Probe in names(agil.RefSeqs)) {
  x<-agil.RefSeqs[[Probe]]
  if(length(x)==1){
    AgilUniqueTx<-rbind(AgilUniqueTx,c(Probe,x))
  }else{AgilMultiTx[[Probe]]<-x}
}
names(AgilUniqueTx)<-c("AgilProbe","RefSeq")

CommonUniqueTx<-intersect(affyUniqueTx$RefSeq,AgilUniqueTx) #0

agil.matches<-lapply(agil.RefSeqs, function(x){
  y<-list()
  for (affyProbe in names(hits.affy)) {
    if(length(intersect(x,hits.affy[[affyProbe]]))>0){
      y[[affyProbe]]<-intersect(x,hits.affy[[affyProbe]])
    }
  }
  return(y)
})

```

Subset by match status

```

noMatch<-list(); SingleMatch<-list(); MultiMatch<-list()
for (Probe in names(agil.matches)) {
  if(length(agil.matches[[Probe]])==0){
    noMatch[[Probe]]<-agil.matches[[Probe]]
  }else if(length(agil.matches[[Probe]])==1){
    SingleMatch[[Probe]]<-agil.matches[[Probe]]
  }else {MultiMatch[[Probe]]<-agil.matches[[Probe]]}
}
length(noMatch); length(SingleMatch); length(MultiMatch)

```

NoMatch 331 SingleMatch 452 MultiMatch 472

Start Matched table

```

AffyProbes<-unlist(lapply(SingleMatch, function(x){names(x)}))
RefSeqs<-unlist(lapply(SingleMatch, function(x){paste(unlist(x),collapse = "/" )}))
Matched<-data.frame(AgilProbe=names(SingleMatch),
                    AffyProbe=AffyProbes,
                    Match=RefSeqs)

```

Remove matched from multimatch—no difference

```

MultiMatch2<-lapply(MultiMatch, function(x){
  x<-x[!names(x)%in%AffyProbes]}

```

Tabulate matches

```
MultiMatch2<-lapply(MultiMatch2, function(x){
  x<-lapply(x, function(y){y<-paste(y,collapse = "/" )})
})

for (Probe in names(MultiMatch2)) {
  df<-data.frame(AgilProbe=Probe,
                 AffyProbe=names(MultiMatch2[[Probe]]),
                 Match=unlist(MultiMatch2[[Probe]]))
  MultiMatch2[[Probe]]<-df
}

MultiMatchTable<-do.call(rbind,MultiMatch2)
```

Unmatched by RefSeq - see if match by gene

```
UsedAffy<-c(MultiMatchTable$AffyProbe,Matched$AffyProbe)
UnusedAffyHits<-hits.affy[!names(hits.affy)%in%UsedAffy]
NoMatchAgil<-names(noMatch)
NoMatchAgilHits<-hits.agil[names(hits.agil)%in%NoMatchAgil]

agil.Genes<-lapply(NoMatchAgilHits, function(x){x[, "Symb"]})

genehits.affy0<-strsplit(affy$Gene.Symbol, " /// ")
names(genehits.affy0)<-affy$Probe.Set.ID
genehits.affy<-genehits.affy0[names(genehits.affy0)%in%names(UnusedAffyHits)]

agil.matches.gene<-lapply(agil.Genes, function(x){
  y<-list()
  for (affyProbe in names(genehits.affy)) {
    if(length(intersect(x,genehits.affy[[affyProbe]]))>0){
      y[[affyProbe]]<-intersect(x,genehits.affy[[affyProbe]])
    }
  }
  return(y)
})

noMatchGene<-list()
SingleMatchGene<-list()
MultiMatchGene<-list()

for (Probe in names(agil.matches.gene)) {
  if(length(agil.matches.gene[[Probe]])==0){
    noMatchGene[[Probe]]<-agil.matches.gene[[Probe]]
  }else if(length(agil.matches.gene[[Probe]])==1){
    SingleMatchGene[[Probe]]<-agil.matches.gene[[Probe]]
  }else {MultiMatchGene[[Probe]]<-agil.matches.gene[[Probe]]}
}
length(noMatchGene);length(SingleMatchGene);length(MultiMatchGene)

AffyProbesGeneMatch<-unlist(lapply(SingleMatchGene, function(x){names(x)}))
Genes<-unlist(lapply(SingleMatchGene, function(x){paste(unlist(x),collapse = "/" )}))
MatchedGene<-data.frame(AgilProbe=names(SingleMatchGene),
                        AffyProbe=AffyProbesGeneMatch,
                        Match=Genes)
```

No match 330 Single 1 Multi 0

Combined Key

```
MatchKey<-rbind(Matched,MultiMatchTable,MatchedGene)
rownames(MatchKey)<-NULL
matches.unique<-unique(MatchKey[,c("AgilProbe", "AffyProbe")])
MatchKey<-MatchKey[rownames(matches.unique),]

# remove the UC probes excluded by max variance filter
MatchKey<-MatchKey[MatchKey$AgilProbe%in%rownames(UC.filteredVar_noNC),]

NonUniqueMatches<-MatchKey[MatchKey$AffyProbe%in%
  MatchKey[duplicated(MatchKey$AffyProbe), "AffyProbe"],]
```

```
#add full list of matched RefSeqs and reannotated UC probes with isoform specificity
```

```
for (i in 1:nrow(NonUniqueMatches)) {
  AgilProbe<-NonUniqueMatches$AgilProbe[i]
  AffyProbe<-NonUniqueMatches$AffyProbe[i]
  probeRS<-hits.affy[[AffyProbe]]
  probeSymb<-genehits.affy0[[AffyProbe]]
  NonUniqueMatches$Symb[i]<-paste(unique(probeSymb),collapse = "/")
  NonUniqueMatches$RefSeqAll[i]<-paste(unique(probeRS),collapse = "/")
  NonUniqueMatches[NonUniqueMatches$AgilProbe==AgilProbe,"AgilSymb"]<-
    UC.maxMean_noNC[AgilProbe,"Symb_ReAnno"]
  NonUniqueMatches[NonUniqueMatches$AgilProbe==AgilProbe,"AgilRS"]<-
    UC.maxMean_noNC[AgilProbe,"RefSeq_ReAnno"]
}

length(unique(NonUniqueMatches$AffyProbe))
#These resulted from 90 non-specific affy probes matching to two agil probes.
# In order to choose the correct match, collapse agil probes by max mean in UC data.
ConfToCollapse<-read.csv("NonUniqueMatchesResolved.csv")
```

```
# subset expression data
datExpr<-exprs(UCmRNAeset)[rownames(exprs(UCmRNAeset))%in%ConfToCollapse$AgilProbe,]
datExpr2<-collapseRows(datExpr,rowGroup = ConfToCollapse$Symb,
  rowID = ConfToCollapse$AgilProbe, method = "maxRowVariance")
ExcludeProbes<-names(datExpr2$selectedRow[datExpr2$selectedRow==F])

UCtoMatch<-UC.filteredVar_noNC[!rownames(UC.filteredVar_noNC)%in%ExcludeProbes,]
MatchKey2<-MatchKey[!MatchKey$AgilProbe%in%ExcludeProbes,]
```

Collapse External data by highest variance per matched probe among probes with multiple matches

```
# Mayo
datExpr<-exprs(MayoNorm)
MatchTab<-MayoCvHCdupCorAll
MatchName<-"M"
datExpr2<-datExpr[rownames(datExpr)%in%MatchKey2$AffyProbe,]
datExpr2<-collapseRows(datExpr,rowGroup = MatchKey2$AgilProbe,
  rowID = MatchKey2$AffyProbe, method = "maxRowVariance")
datExpr3<-datExpr2$datETcollapsed
rownames(datExpr3)<-names(datExpr2$selectedRow[datExpr2$selectedRow==T])
matchResult<-MatchKey2[MatchKey2$AffyProbe%in%
  names(datExpr2$selectedRow[datExpr2$selectedRow==T]),]

for (i in 1:nrow(matchResult)) {
  AgilProbe<-matchResult$AgilProbe[i]
  AffyProbe<-matchResult$AffyProbe[i]
  matchResult[i,"Symb"]<-UCtoMatch[AgilProbe,"Symb_ReAnno"]
  matchResult[i,"logFC_UC"]<-UCtoMatch[AgilProbe,"logFC"]
  matchResult[i,"FC_UC"]<-UCtoMatch[AgilProbe,"FC"]
  matchResult[i,"p_UC"]<-UCtoMatch[AgilProbe,"P.Value"]
  matchResult[i,"FDR_UC"]<-UCtoMatch[AgilProbe,"adj.P.Val"]
  matchResult[i,"AveExpr_UC"]<-UCtoMatch[AgilProbe,"AveExpr"]
  matchResult[i,"mean1_UC"]<-UCtoMatch[AgilProbe,"mean1"]
  matchResult[i,"mean2_UC"]<-UCtoMatch[AgilProbe,"mean2"]
  matchResult[i,"Dir_UC"]<-UCtoMatch[AgilProbe,"Direction"]
  matchResult[i,paste("logFC",MatchName,sep = "_")]<-MatchTab[AffyProbe,"logFC"]
  matchResult[i,paste("FC",MatchName,sep = "_")]<-MatchTab[AffyProbe,"FC"]
  matchResult[i,paste("p",MatchName,sep = "_")]<-MatchTab[AffyProbe,"P.Value"]
  matchResult[i,paste("FDR",MatchName,sep = "_")]<-MatchTab[AffyProbe,"adj.P.Val"]
  matchResult[i,paste("AveExpr",MatchName,sep = "_")]<-MatchTab[AffyProbe,"AveExpr"]
  matchResult[i,paste("mean1",MatchName,sep = "_")]<-MatchTab[AffyProbe,"mean1"]
  matchResult[i,paste("mean2",MatchName,sep = "_")]<-MatchTab[AffyProbe,"mean2"]
  matchResult[i,paste("Dir",MatchName,sep = "_")]<-MatchTab[AffyProbe,"Direction"]
  matchResult[i,"Name"]<-UCtoMatch[AgilProbe,"Name"]
}
rownames(matchResult)<-matchResult$AgilProbe
dir.filter<-matchResult[, "Dir_UC"]==matchResult[,paste("Dir",MatchName,sep = "_")]
matchedDir<-matchResult[dir.filter,]
```

```

matchedDirSig<-matchedDir[matchedDir[,paste("p",MatchName,sep = "_")]<0.05,]

M.datExpr.matched<-datExpr3
M.CvHC.matchedDir<-matchedDir
M.CvHC.matchedDirSig<-matchedDirSig

# Repeat for Nottingham
N.datExpr.matched<-datExpr3
N.CvHC.matchedDir<-matchedDir
N.CvHC.matchedDirSig<-matchedDirSig

# ALL 3
commonSymb<-intersect(M.CvHC.matchedDirSig$Symb,N.CvHC.matchedDirSig$Symb)
commonTable<-M.CvHC.matchedDirSig[M.CvHC.matchedDirSig$Symb%in%commonSymb,]
nottCommon<-N.CvHC.matchedDirSig[rownames(commonTable),]
notCols<-names(nottCommon)[grep("_N",names(nottCommon))]
commonTable<-cbind(commonTable,nottCommon[,names(nottCommon)%in%notCols])

```

GO enrichment

```

library(GOstats)
library(hgu95av2.db)
view.GOstats<-function(toptable,universe,dir){
  if(dir=="up"){toptable<-toptable[toptable$logFC>0,]}
  else{toptable<-toptable[toptable$logFC<0,]}
  selectedEntrezIds<-as.character(toptable$EntrezID)
  params <- new("GOHyperGParams", geneIds = selectedEntrezIds,
    universeGeneIds = universe, annotation = "hgu95av2.db",
    ontology = "BP", pvalueCutoff = 0.05, conditional = FALSE,
    testDirection = "over")
  hgOver <- hyperGTest(params)
  table <- summary(hgOver)
  return(table)
}

```

Common genes

```

commonTable<-UCtoMatch[rownames(UCtoMatch)%in%rownames(M.CvHC.matchedDirSig),]
universeUC<-as.character(fitUC$genes$EntrezID)
universeMayo<-as.character(fitMayo$genes$EntrezID)
universeCommon<-union(universeUC,universeMayo) #same chip/same universe for Roch and Nott
common.up<-view.GOstats(commonTable,universeCommon,"up")
common.down<-view.GOstats(commonTable,universeCommon,"down")

```

WGCNA modules done in same way, universe was EntrezIDs in WGCNA analysis.

Nanostring nCounter validation

quality control initially identified 4 outliers. The fourth is only z score 3.03 so will retain this one

```
library(NanoStringNorm)
```

Load data (outliers removed)

```

load("NSraw0717NoOutliers.rda")
load("NSpheno0717NoOutliers.rda")

```

Normalization

See NanoStringNorm paper: <https://doi.org/10.1093/bioinformatics/bts188>

Code-count normalization: accounts for lane-by-lane variation of the nCounter platform. Normalization occurs by summarizing (i.e. mean, median, sum or geometric mean) the positive control counts and adjusting samples by a factor relative to other samples.

Background correction: negative controls

Sample content normalization: normalization genes summarized (e.g. geometric mean, etc.) and a sample adjustment factor is calculated.

convenience function for normalization

```
myNSnorm<- function(raw,method,logged){
  library(NanoStringNorm)
  norm <- NanoStringNorm(
    x = raw, CodeCount = method[1],
    Background = method[2], SampleContent = method[3],
    OtherNorm = method[4], round.values = TRUE,
    take.log = logged, return.matrix.of.endogenous.probes = TRUE)
  return(norm)
}
```

convenience function to save counts

```
savecounts<-function(countmatrix,namestart){
  date <- gsub("-", "", gsub("2016-", "", Sys.Date()))
  methName <- gsub("geomean", "gm",
    gsub("\\.", "",
      gsub("none", "na",
        gsub("housekeeping", "hk", paste(method, collapse = "_")))))
  filename <- paste(namestart, "_", methName, date, ".csv", sep = "")
  write.csv(countmatrix, file = filename)
  print(paste("Filename:", filename))
  method.sum<-methName
}
```

negative controls were extremely low and the NanoString rep advised against using a background control factor. for content norm, used low.cv.geo.mean which chooses good reference genes (low variation) since pre-selected reference genes (GAPDH, beta actin, PPIA had higher CV)

Since will use voom, do not log transform

```
method <- c("geo.mean", "none", "low.cv.geo.mean", "none")
NSnormNolog<-myNSnorm(raw = NSraw0717NoOutliers, method = method, logged = FALSE)
savecounts(countmatrix = NSnormNolog, namestart = "NSnolog45")
```

```
#there were 6 in the set that were not part of this analysis
load("NSvalidationMAprobes.rda") #maps symbols to microarray probes
NSnormNolog<-NSnormNolog[as.character(NSvalidationMAprobes$GeneSymbol),]
dim(NSnormNolog) #[1] 67 45
```

Differential Expression

Use voom to transform for linear modeling

```
library(limma)
f <- as.factor(NSpheno0717NoOutliers$Group)
design <- model.matrix(~ 0 + f)
levels(f)
## [1] "C" "D" "HC"
colnames(design)=levels(f)
v1<-voom(NSnormNolog,design)
```

Linear modeling

```
fit <- lmFit(v1, design)
contrasts <- makeContrasts((D+C)/2-HC, D-HC, C-HC, levels=design)
fitcon <- contrasts.fit(fit, contrasts)
eb1 <- eBayes(fitcon)
coefficients<-colnames(eb1$coefficients)
```

Function to save results

```
mytop<-function(object,names){
  # returns list, saves csv of top tables
  # object: eg eb
  # names vector of names of top tables
  toplist<-list()
  top1<-topTable(object, coef = coefficients[1], adjust.method = "BH", number = Inf)
  top2<-topTable(object, coef = coefficients[2], adjust.method = "BH", number = Inf)
  top3<-topTable(object, coef = coefficients[3], adjust.method = "BH", number = Inf)
  write.csv(top1,file = paste(names[1],"csv",sep = "."))
  write.csv(top2,file = paste(names[2],"csv",sep = "."))
  write.csv(top3,file = paste(names[3],"csv",sep = "."))
  toplist[["IBSVHC"]] <-top1
```



```

    toplist[["DvHC"]] <-top2
    toplist[["CvHC"]] <-top3
    return(toplist)
}
tops_all<-mytop(eb1,names = c("IBSVHCall","DvHCall","CvHCall"))

```

Determine results in microarray cohort and validation cohort as above

Determine validation

```

load("UC_CvHCtoptables.rda")
MA.C<-UC.CvHCall
MA.C<-MA.C[rownames(NSvalidationMAprobes),]
names(MA.C)<-paste(names(MA.C),"MA",sep = "_")
rownames(MA.C)<-MA.C$Symbol_MA

NS.C.all<-tops_all$CvHC
NS.C.MA<-NStopsMA$CvHC #nanosttring results in microarray cohort
NS.C.val<-NStopsVal$CvHC #nanosttring results in validation cohort

names(NS.C.all)<-paste(names(NS.C.all),"NS",sep = "_")
names(NS.C.MA)<-paste(names(NS.C.MA),"NS",sep = "_")
names(NS.C.val)<-paste(names(NS.C.val),"NS",sep = "_")

length(setdiff(rownames(MA.C),rownames(NS.C.all))) #0

MA.C<-MA.C[order(rownames(MA.C)),]
NS.C.all<-NS.C.all[order(rownames(NS.C.all)),]
NS.C.MA<-NS.C.MA[order(rownames(NS.C.MA)),]
NS.C.val<-NS.C.val[order(rownames(NS.C.val)),]

NSMA.C.all<-cbind.data.frame(NS.C.all,MA.C)
NSMA.C.MA<-cbind.data.frame(NS.C.MA,MA.C)
NSMA.C.val<-cbind.data.frame(NS.C.val,MA.C)

matchDirection<-function(data){
  data<-data[data$P.Value_NS<=0.05,]
  same<- (data$logFC_MA>0&data$logFC_NS>0)|(data$logFC_MA<0&data$logFC_NS<0)
  data<-data[same==TRUE,]
}

NS.C.validated.all<-matchDirection(NSMA.C.all)
NS.C.validated.ma<-matchDirection(NSMA.C.MA)
NS.C.validated.val<-matchDirection(NSMA.C.val)

NS.C.validated.all.FDR<-NS.C.validated.all[NS.C.validated.all$adj.P.Val_NS<0.051,]
NS.C.validated.ma.FDR<-NS.C.validated.ma[NS.C.validated.ma$adj.P.Val_NS<0.051,]
NS.C.validated.val.FDR<-NS.C.validated.val[NS.C.validated.val$adj.P.Val_NS<0.051,]

NSMAvalC<-list()
NSMAvalC$allP<-NS.C.validated.all
NSMAvalC$maP<-NS.C.validated.ma
NSMAvalC$valP<-NS.C.validated.val
NSMAvalC$allFDR<-NS.C.validated.all.FDR
NSMAvalC$maFDR<-NS.C.validated.ma.FDR

for(df in names(NSMAvalC)){
  f<-paste("NSMAvalC",df,".csv",sep = "")
  write.csv(NSMAvalC[[df]],f)
}

```

Association with symptoms

```

load("NSTraits.rda") #Loads pheno data
load("NSnormNolog67x45.rda")
NSvroom<-voom(NSnormNolog)
traits<-traits[traits$IBS=="IBS",]
NS.norm.values<-as.data.frame(t(NSvroom$E))
geneData<-NS.norm.values[rownames(traits),]
genes<-names(geneData)

```


Overall symptoms

```
OverallSx<-data.frame(
  row.names = genes, F.pvalue=numeric(length(genes)),
  B=numeric(length(genes)), Sx.pvalue=numeric(length(genes)))

for (i in 1:nrow(OverallSx)) {
  df1<-cbind(geneData[,row.names(OverallSx)[i]],traits)
  names(df1)[1]<-"gene"
  fit<-lm(BSQ_OverallSx~factor(Sex)+gene,data = df1, na.action = na.exclude)
  sumfit<-summary(fit)
  OverallSx[i,"F.pvalue"]<- pf(sumfit$fstatistic[1], sumfit$fstatistic[2],
    sumfit$fstatistic[3],lower.tail = FALSE)
  OverallSx[i,"B"]<-sumfit$coefficients["gene","Estimate"]
  OverallSx[i,"Sx.pvalue"]<-sumfit$coefficients["gene","Pr(>|t|)"]
}

OverallSx$FDR<-p.adjust(OverallSx$Sx.pvalue,method = "BH")
OverallSx<-OverallSx[order(OverallSx$Sx.pvalue),]
```

Abdominal Pain, Bloating as above

Chapter 3

Code for heatmaps and WGCNA mirrors Chapter 2 code using lncRNA expression data.

Chapter 4

Meta-analysis of TEER experiments

Function to tabulate summary data and statistics from individual experiments

```
getRes<-function(StudyName,StartDays,data,time,xvar,comp){
  dat2<-data[data$Time==time,]
  varname<-xvar[1]
  grp1<-dat2$Percent[dat2[,xvar[1]]==xvar[2]]
  grp2<-dat2$Percent[dat2[,xvar[1]]==xvar[3]]
  tres<-t.test(grp1,grp2)
  se1<-sd(grp1,na.rm = T)/sqrt(length(grp1[!is.na(grp1)]))
  se2<-sd(grp2,na.rm = T)/sqrt(length(grp2[!is.na(grp2)]))
  baseTEERs<-lapply(c(xvar[2],xvar[3]), function(x){
    dat0<-data[data$Time==0,]
    m<-format(mean(dat0$MeanR[dat0[,xvar[1]]==x]/1000),digits = 3)
    sd<-format(sd(dat0$MeanR[dat0[,xvar[1]]==x]/1000),digits = 2)
    msd<-paste(m," (", sd,")",sep = "")
    return(msd)
  })
  df<-data.frame(Study=StudyName,
    DoxStartDay=StartDays[1],InfStartDay=StartDays[2],
    TEER0=StartDays[3],Comparison=comp,Time=time,
    Grp1=xvar[2],Grp2=xvar[3],
    n1=length(grp1[!is.na(grp1)]),n2=length(grp2[!is.na(grp2)]),
    BaseTEER1=baseTEERs[[1]],BaseTEER2=baseTEERs[[2]],
    Mean1=tres$estimate[1],Mean2=tres$estimate[2],
    sd1=sd(grp1,na.rm = T),sd2=sd(grp2,na.rm = T),se1=se1,se2=se2,
    Estimate=tres$estimate[1]- tres$estimate[2],
    EstSE=(tres$conf.int[2]-tres$conf.int[1])/3.92,
    CIlow=tres$conf.int[1],CIhigh=tres$conf.int[2],
    pvalue=format(tres$p.value,digits = 3,scientific = F)
  )
  return(df)
}
```

Example for one experiment (repeated for all with results iteratively added (rbind) to doxResp and infResp)

```
load("Data/TEER_081417.rda")
doxyrespNoOuts.081417$Time<-doxyrespNoOuts.081417$Time-min(doxyrespNoOuts.081417$Time)
infgResp.081417$Time<-infgResp.081417$Time-min(infgResp.081417$Time)

xvar<-c("Doxy","Doxy","Control")
comp = "Dox vs No Dox AFDN-AS1"
StudyName="Aug2017_1"
StartDays=c(6,NA,5)

res_doxyResp.081417<-lapply(c(1,2,3), function(x){
  return(getRes(
    StudyName = StudyName,
    StartDays=StartDays,
    data = doxyrespNoOuts.081417, comp = comp,
    time = x, xvar = xvar))
})
res_doxyResp.081417<-do.call(rbind, res_doxyResp.081417)
doxResp<-res_doxyResp.081417

StartDays=c(6,9,5)

res_infgResp.081417<-lapply(c(1,2,3), function(x){
  return(getRes(StudyName = StudyName, StartDays = StartDays,
    data = infgResp.081417, comp = comp,
    time = x, xvar = xvar))
})
res_infgResp.081417<-do.call(rbind, res_infgResp.081417)
infResp<-res_infgResp.081417

resList<-list()
resList$Dox$Aug.14.17<-res_doxyResp.081417
resList$Infg$Aug.14.17<-res_infgResp.081417
```

Create list with result data

```
metaList<-list()
metaList$Dox$`AFDN-DT`<-doxResp[doxResp$Comparison=="Dox vs No Dox AFDN-AS1",]
metaList$Dox$EV<-doxRespEV[doxRespEV$Comparison=="Dox vs No Dox EV",]
metaList$Dox$AFvEVdox<-doxResp[doxResp$Comparison=="AFDN-AS1 vs EV Dox",]
metaList$Dox$AFvEVnodox<-doxResp[doxResp$Comparison=="AFDN-AS1 vs EV No Dox",]

metaList$Infg$`AFDN-DT`<-infResp[infResp$Comparison=="Dox vs No Dox AFDN-AS1",]
metaList$Infg$EV<-infRespEV[infRespEV$Comparison=="Dox vs No Dox EV",]
metaList$Infg$AFvEVdox<-infResp[infResp$Comparison=="AFDN-AS1 vs EV Dox",]
metaList$Infg$AFvEVnodox<-infResp[infResp$Comparison=="AFDN-AS1 vs EV No Dox",]
```

Forest plots

```
library(meta)
library(forestplot)
```

Functions to calculate total wells, base TEER column and meta-analysis results

```
makeWells<-function(){
  wells <- paste(tableDat$n1 + tableDat$n2, " (",tableDat$n1,
    "/", tableDat$n2, ")", sep = "")
  sumWells <- paste(sum(tableDat$n1 + tableDat$n2), " (",
    sum(tableDat$n1), "/", sum(tableDat$n2), ")", sep = "")
  return(c(wells, sumWells))
}

makeBaseTEER<-function(){
  baseTEERs<-paste(tableDat$BaseTEER1,tableDat$BaseTEER2,sep = ", ")
  return(c(baseTEERs,NA))
}

calcSumRes<-function(){
  sumRes <- metacont(
    n.e = tableDat$n1,
```

```

    mean.e = tableDat$Mean1,
    sd.e = tableDat$sd1,
    n.c = tableDat$n2,
    mean.c = tableDat$Mean2,
    sd.c = tableDat$sd2,
    sm = "MD"
  )
  return(sumRes)
}

```

Function to print plot

```

plotMetaDoxAFday<-function(Time){
  tableDat <- doxResp[doxResp$Time==Time&doxResp$Comparison=="Dox vs No Dox AFDN-AS1",]
  plotTitle<-paste("Response to Doxycycline: T84-TR-AFDN-DT with vs. without doxycycline\n% change in TE
ER ",Time," days post doxycycline", sep = "")
  sumRes<-calcSumRes()
  BaseTeer<-makeBaseTEER()
  Wells<-makeWells()
  tabletext <-
    cbind(
      c("Study",NA, as.character(tableDat$Study), "Summary"),
      c("Dox Start", "(days post-seeding)", as.character(tableDat$DoxStartDay),NA),
      c("n wells","Tot(Dox/No Dox)",Wells),
      c("Baseline TEER","mean (sd), (Dox/No Dox)",BaseTeer),
      c("Estimate","(difference)", round(c(
        sumRes$TE, sumRes$TE.random
      ), 2)),
      c("p value",NA, format(
        c(sumRes$pval, sumRes$pval.random), digits = 3
      ))
    )
  #specify formatting
  labellist<-list()
  labellist$Row1$Col1<-gpar(cex=0.4);labellist$Row1$Col2<-gpar(cex=0.4)
  labellist$Row1$Col3<-gpar(cex=0.4);labellist$Row1$Col4<-gpar(cex=0.4)
  labellist$Row1$Col5<-gpar(cex=0.4);labellist$Row1$Col6<-gpar(cex=0.4)
  labellist$Row2$Col1<-gpar(cex=0.6);labellist$Row2$Col2<-gpar(cex=0.6)
  labellist$Row2$Col3<-gpar(cex=0.6);labellist$Row2$Col4<-gpar(cex=0.6)
  labellist$Row2$Col5<-gpar(cex=0.6);labellist$Row2$Col6<-gpar(cex=0.6)
  labellist$Row3$Col1<-gpar(cex=0.6);labellist$Row3$Col2<-gpar(cex=0.6)
  labellist$Row3$Col3<-gpar(cex=0.6);labellist$Row3$Col4<-gpar(cex=0.6)
  labellist$Row3$Col5<-gpar(cex=0.6);labellist$Row3$Col6<-gpar(cex=0.6)
  labellist$Row4$Col1<-gpar(cex=0.6);labellist$Row4$Col2<-gpar(cex=0.6)
  labellist$Row4$Col3<-gpar(cex=0.6);labellist$Row4$Col4<-gpar(cex=0.6)
  labellist$Row4$Col5<-gpar(cex=0.6);labellist$Row4$Col6<-gpar(cex=0.6)
  print(
    forestplot(
      labeltext = tabletext,graph.pos = 6,
      mean = c(NA,NA, sumRes$TE, sumRes$TE.random),
      lower = c(NA,NA, sumRes$lower, sumRes$lower.random),
      upper = c(NA, NA,sumRes$upper, sumRes$upper.fixed),
      ci.vertices = T,title = plotTitle,line.margin = 0,
      colgap=unit(2,"mm"),is.summary = c(T,rep(F,4), T),
      xlab = "Difference in % Baseline TEER",mar = unit(c(3,3,3,3),"mm"),
      txt_gp = fpTxtGp(label=labellist,title = gpar(cex=0.7),
        summary = gpar(cex=0.6),xlab = gpar(cex=0.5),
        ticks = gpar(cex=0.5)
      ))
    )
  }
}

```

Print plot

```

png(filename = "testMeta.png", width = 6, height = 2, units = "in", res = 300)
plotMetaDoxAFday(3)
dev.off()

```

RNA sequencing

Set-up

```
library(DESeq2)
library(biomaRt)
library(edgeR)
```

DeSeq data set

contains expression data (exprColDat), Sample data (ColDat) and annotation (RowDat)

Expression data

Count data were originally in one file per sample names SampleName.txt

```
files1 <- dir(path = "~/read_counts/")
df1 <- readDGE(files1, path = "~/read_counts/", header = F )
allDat <- as.data.frame(df1$counts)
exprColDat<-allDat
```

Sample data table

```
load("pheno.rda")
rundata<-read.csv("ReadInfo.csv", row.names = 1)
cells<-intersect(row.names(rundata), row.names(pheno))
pheno<-cbind(pheno[cells,], rundata[cells,])
colDat<-pheno[names(allDat),]
colDat$Group <- factor(colDat$Group, levels = c("EV", "AFDN_DT"))
colDat$Dox<- factor(colDat$Dox, levels=c("No_Doxycycline", "Doxycycline"))
colDat$GroupDox[colDat$Group=="AFDN_DT"&colDat$Dox=="Doxycycline"]<-"AF_Dox"
colDat$GroupDox[colDat$Group=="AFDN_DT"&colDat$Dox=="No_Doxycycline"]<-"AF"
colDat$GroupDox[colDat$Group=="EV"&colDat$Dox=="Doxycycline"]<-"EV_Dox"
colDat$GroupDox[colDat$Group=="EV"&colDat$Dox=="No_Doxycycline"]<-"EV"
colDat$GroupDox <- as.factor(colDat$GroupDox)
```

Annotation

```
notgene<-
exprColDat[!row.names(exprColDat)%in%grep("ENSG", row.names(exprColDat), value = T),]

exprColDat<-
exprColDat[row.names(exprColDat)%in%grep("ENSG", row.names(exprColDat), value = T),]

mart.hs <- useMart("ensembl", "hsapiens_gene_ensembl")
an <- getBM(attributes = c("ensembl_gene_id", "hgnc_symbol", "entrezgene", "description", "chromosome_name"),
  filters = "ensembl_gene_id", values = row.names(exprColDat), mart = mart.hs)

an<-an[!duplicated(an$ensembl_gene_id),]
notinAn<-exprColDat[!row.names(exprColDat)%in%an$ensembl_gene_id,]
# all deprecated - most no expression

row.names(an)<-an$ensembl_gene_id
save(an, file="ENSGhsAn.rda")
library(dplyr)
rowDat <- rownames_to_column(exprColDat, "ensembl_gene_id")[,1, drop=F] %>%
  left_join(an, "ensembl_gene_id")
save(rowDat, exprColDat, colDat, file = "CountDataAll.rda")
```

Clustering identifies W16 as outlier

For clustering, counts are transformed for linear modeling following exclusion of genes with low counts (these do not need to be excluded for differential expression, though genes with very low counts are removed to reduce the memory needed)

```
dds <- DESeqDataSetFromMatrix (countData = exprColDat,
  colData = colDat,
  design = ~Group+Dox+Group:Dox)

# remove very low count genes
keep <- rowSums(counts(dds)) >= 10
dds <- dds[keep,]

# remove low count genes
countsGr10in30perfilt<-function(dds){
```

```

n30percent<-.3*dim(dds)[2]
countsGr10in30per<-apply(counts(dds),1,function(y){
  y2<-ifelse(y>10,1,0)
  y3<-ifelse(sum(y2)>n30percent,TRUE,FALSE)
  return(y3)
})
ddsfilt<-dds[countsGr10in30per,]
return(ddsfilt)
}
dds<-countsGr10in30perfilt(dds)
vs<-vst(dds,blind = FALSE)
edata<-as.data.frame(assay(vs))
edata<-normalizeQuantiles(as.matrix(edata))
datExpr<-as.data.frame(t(edata))
rownames(datExpr)<-colnames(edata)
library(WGCNA)
sampleTree = hclust(dist(datExpr), method = "average")
pdf(file = "SigSampClust042419.pdf")
par(cex = 0.6);
par(mar = c(0,4,2,0))
plot(sampleTree,
      main = "Sample clustering to detect outliers",
      sub="", xlab="", cex.lab = 1.5,
      cex.axis = 1.5, cex.main = 2)
dev.off()

```

Remove W16

```

colDat<-colDat[rownames(colDat)!="W16",]
exprColDat<-exprColDat[,rownames(colDat)]
save(exprColDat,colDat,rowDat,file = "CountDataAllnoW16.rda")

```

Principal components plot

```

dds<-dds[,dds$Sample!="W16"]
vsd <- vst(dds)
data <- plotPCA(vsd, intgroup = c("Group","Dox","GroupDox2","AFDTpcr"), returnData=TRUE)
percentVar <- round(100 * attr(data, "percentVar"))

library(ggplot2)
load("/myColors.rda")

ggplot(data, aes(PC1PC2, fill = GroupDox, alpha = GroupDox, size = AFDTpcr)) +
  geom_point(shape=21) +
  xlab(paste0("PC1: ", percentVar[1], "% variance")) +
  ylab(paste0("PC2: ", percentVar[2], "% variance")) +
  scale_alpha_manual(name = "Group", values = c(0.5, 1, 0.5, 1)) +
  scale_fill_manual(name = "Group", values = c(
    myColors$main$yellow1, myColors$main$yellow1,
    myColors$main$blue1, myColors$main$blue1
  ))
) +
  scale_size_continuous(name = "AFDN-DT PCR Expr.",range = c(1,10)) +
  theme_bw() +
  guides(
    fill = guide_legend(order = 1,override.aes = list(size=5)),
    alpha = guide_legend(order = 1), size = guide_legend(order = 2)
  )
)

```

Differential expression

```

library(tidyverse)
load("CountDataAllnoW16.rda")

```

Main effects of doxycycline and cell line

```

dds <- DESeqDataSetFromMatrix (countData = exprColDat,
                              colData = colDat,
                              rowData = rowDat,

```

```

                                design = ~Dox+Group)
keep <- rowSums(counts(dds)) >= 10
dds <- dds[keep,]
dds <- DESeq(dds)
resultsNames(dds)
# [1] "Intercept"
# [2] "Dox_Doxycycline_vs_No_Doxycycline"
# [3] "Group_AFDN_DT_vs_EV"

# Dox
res<-results(dds, name = "Dox_Doxycycline_vs_No_Doxycycline")
summary(res)
# out of 14847 with nonzero total read count
# adjusted p-value < 0.1
# LFC > 0 (up)      : 3, 0.02%
# LFC < 0 (down)    : 2, 0.013%
# outliers [1]      : 16, 0.11%
# low counts [2]     : 0, 0%
# (mean count < 1)

ResList<-list()
ResList$DoxMain<-res

# Group
res<-results(dds, name = "Group_AFDN_DT_vs_EV")
summary(res)
# out of 14847 with nonzero total read count
# adjusted p-value < 0.1
# LFC > 0 (up)      : 157, 1.1%
# LFC < 0 (down)    : 134, 0.9%
# outliers [1]      : 16, 0.11%
# low counts [2]     : 9771, 66%
# (mean count < 32)

ResList$GroupMain<-res

```

Interaction term for dox effect in AF vs EV

```

dds <- DESeqDataSetFromMatrix (countData = exprColDat,
                                colData = colDat,
                                rowData = rowDat,
                                design = ~Dox+Group+Dox:Group)

keep <- rowSums(counts(dds)) >= 10
dds <- dds[keep,]
dds <- DESeq(dds)
resultsNames(dds)
# [1] "Intercept"
# [2] "Dox_Doxycycline_vs_No_Doxycycline"
# [3] "Group_AFDN_DT_vs_EV"
# [4] "DoxDoxycycline.GroupAFDN_DT"

res<-results(dds, name="DoxDoxycycline.GroupAFDN_DT")
summary(res)

# out of 14847 with nonzero total read count
# adjusted p-value < 0.1
# LFC > 0 (up)      : 0, 0%
# LFC < 0 (down)    : 0, 0%
# outliers [1]      : 0, 0%
# low counts [2]     : 0, 0%
# (mean count < 0)

ResList$GroupDoxInt<-res

```

Dox effect within groups, group effect within Dox

```
dds <- DESeqDataSetFromMatrix (countData = exprColDat,  
                              colData = colDat,  
                              rowData = rowDat,  
                              design = ~0+GroupDox)  
  
keep <- rowSums(counts(dds)) >= 10  
dds <- dds[keep,]  
dds <- DESeq(dds)  
resultsNames(dds)  
# [1] "GroupDoxAF"      "GroupDoxAF_Dox" "GroupDoxEV"      "GroupDoxEV_Dox"  
  
# AF+dox vs AF  
res<-results(dds, contrast = c("GroupDox", "AF_Dox", "AF"))  
summary(res)  
# out of 14847 with nonzero total read count  
# adjusted p-value < 0.1  
# LFC > 0 (up)      : 7, 0.047%  
# LFC < 0 (down)    : 9, 0.06%  
# outliers [1]      : 0, 0%  
# low counts [2]     : 13529, 91%  
# (mean count < 185)  
  
ResList$GroupDox_AF_Dox_vs_AF<-res  
  
# EV+dox vs EV  
res<-results(dds, contrast = c("GroupDox", "EV_Dox", "EV"))  
summary(res)  
# out of 14847 with nonzero total read count  
# adjusted p-value < 0.1  
# LFC > 0 (up)      : 0, 0%  
# LFC < 0 (down)    : 0, 0%  
# outliers [1]      : 0, 0%  
# low counts [2]     : 0, 0%  
# (mean count < 1)  
  
ResList$GroupDox_EV_Dox_vs_EV<-res  
  
#AF vs EV  
res<-results(dds, contrast = c("GroupDox", "AF", "EV"))  
summary(res)  
  
# out of 14847 with nonzero total read count  
# adjusted p-value < 0.1  
# LFC > 0 (up)      : 91, 0.61%  
# LFC < 0 (down)    : 57, 0.38%  
# outliers [1]      : 0, 0%  
# low counts [2]     : 12089, 81%  
# (mean count < 80)  
  
ResList$GroupDox_AF_vs_EV<-res  
  
ResAnList<-lapply(ResList, function(x){  
  xAn<-cbind(as.data.frame(x),rowData(dds)[,2:4])  
  xAn<-xAn[!is.na(xAn$padj),]  
  xAn<-xAn[order(xAn$pvalue),]  
  xAn$FC<-2^xAn$log2FoldChange  
  return(as.data.frame(xAn))  
})  
  
SigCountsTally2<-function(x){  
  xUp=x[x$log2FoldChange>0,]  
  xDown=x[x$log2FoldChange<0,]  
  FDR10up=nrow(xUp[xUp$padj<=0.1,])  
  FDR10down=nrow(xDown[xDown$padj<=0.1,])  
  FDR05up=nrow(xUp[xUp$padj<=0.05,])  
  FDR05down=nrow(xDown[xDown$padj<=0.05,])
```

```

P05up=nrow(xUp[xUp$pvalue<=0.05,])
P05down=nrow(xDown[xDown$pvalue<=0.05,])
return(c(FDR10=paste(FDR10up+FDR10down, " (", FDR10up, "/", FDR10down, ")", sep = ""),
          FDR05=paste(FDR05up+FDR05down, " (", FDR05up, "/", FDR05down, ")", sep = ""),
          P05=paste(P05up+P05down, " (", P05up, "/", P05down, ")", sep = "")))
}

makeSigCountsTable2<-function(y){
  yFC2<-y[abs(y$log2FoldChange)>=1,]
  return(data.frame(All=SigCountsTally2(y),
                    FC2=SigCountsTally2(yFC2)))
}

write.csv(do.call(cbind,lapply(ResAnList, makeSigCountsTable2)),
          "SigCountsTable.csv")

```

WGCNA

Single network

```

library(WGCNA)
options(stringsAsFactors = F)

dds <- DESeqDataSetFromMatrix (countData = exprColDat,
                              colData = colDat,
                              design = ~Group+Dox+Group:Dox)

keep <- rowSums(counts(dds)) >= 10
dds <- dds[keep,]
dds <- DESeq(dds)
dds<-countsGr10in30perfilt(dds)
dim(dds) #[1] 9005 15
vs<-vst(dds,blind = FALSE)
edata<-as.data.frame(assay(vs))
edata<-normalizeQuantiles(as.matrix(edata))
datExpr.all<-as.data.frame(t(edata))
rownames(datExpr.all)<-colnames(edata)

softPower=18
netType="signed"
deepSplit = 1
MEDissThres = 0.25
minModuleSize = 30
fname.root<-"18_30_1_25_all"

adjacency = adjacency(datExpr, power = softPower, type = netType);
TOM = TOMsimilarity(adjacency);
dissTOM = 1-TOM
rm(adjacency)
geneTree = hclust(as.dist(dissTOM), method = "average");
dynamicMods = cutreeDynamic(dendro = geneTree, distM = dissTOM,
                           deepSplit = deepSplit, pamRespectsDendro = FALSE,
                           minClusterSize = minModuleSize);
dynamicColors = labels2colors(dynamicMods)
MEList = moduleEigengenes(datExpr, colors = dynamicColors,excludeGrey = T)
MEs = MEList$eigengenes
MEDiss = 1-abs(cor(MEs));
METree = hclust(as.dist(MEDiss), method = "average");
merge = mergeCloseModules(datExpr, dynamicColors,
                          cutHeight = MEDissThres,
                          verbose = 0, getNewUnassdME = FALSE)

mergedColors = merge$colors;
mergedMEs = merge$newMEs
moduleColors = mergedColors
colorOrder = standardColors(150);
moduleLabels = match(moduleColors, colorOrder);
MEs = mergedMEs
MEs0 = moduleEigengenes(datExpr, moduleColors,excludeGrey = T)$eigengenes
MEs = orderMEs(MEs0)

```



```
pdf(paste("EigenNet_", fname.root, ".pdf", sep = ""), width = 7, height = 6.5)
plotEigengeneNetworks(MEs, setLabels = NULL, plotDendrograms = F,
                      signed = T, setMargins = T)
dev.off()
```

Correlate modules with group/treatment

```
load("datTraitsAll.rda")
datTraits<-datTraits[,c(2,3,5,6,4)]

datTraitsAFvEV[datTraitsGroup==1&datTraitsDox==0]<-1
datTraitsAFvEV[datTraitsGroup==0&datTraitsDox==0]<-0

datTraits<-datTraits[,c(1,6,2:5)]

ylabs<-paste(labelslabel, " (", labelssize, ")", sep = "")

moduleTraitCor = cor(MEs.all,
                     datTraits,
                     use = "p")
moduleTraitPvalue = corPvalueStudent(moduleTraitCor,
                                     nrow(datExpr.all))
textMatrix = paste(signif(moduleTraitCor, 2),
                   " (", signif(moduleTraitPvalue, 1), ")",
                   sep = "")
dim(textMatrix) = dim(moduleTraitCor)

textMatrix2<-textMatrix

for(i in 1:dim(textMatrix)[1]){
  for (j in 1:dim(textMatrix)[2]) {
    if (moduleTraitPvalue[i,j]<0.05) {
      textMatrix2[i,j]<-paste(textMatrix2[i,j], "*", sep = "")
    }
  }
}

xlabs<-c("Group\n(all)",
         "Group\n(no Dox)",
         "Dox\n(all)",
         "Dox\n(AFDN-DT)",
         "Dox\n(EV)",
         "AFDN-DT\nPCR Expr.")

png(file = "SingleNetTraitCorCombo.png", width = 6.5,
    height = 3.3, units = "in", res = 300);par(
  mar = c(2, 10, 0.5, 0.1)
);labeledHeatmap(
  Matrix = moduleTraitCor,
  textMatrix = textMatrix2,
  xLabels = xlabs,
  yLabels = names(MEs.all),
  ySymbols = ylabs,
  yColorLabels = TRUE,
  colors = blueWhiteRed(50), setStdMargins = FALSE,
  cex.text = 0.6, xLabelsAngle = 0,
  cex.lab.y = 0.6, cex.legend = 0.5,
  cex.lab.x = 0.6,
  xLabelsAdj = c(0.5, 0.5));dev.off()
```

Gene info table

```
notgrey<-!moduleColors=="grey"

load("ENSGHsAn.rda")
an<-an[colnames(datExpr[,notgrey]),]
makeGeneInfo<-function(geneData, modules, geneModMemb){
  geneInfo0 = data.frame(ENSG = geneData$ensembl_gene_id,
                        geneSymbol = geneData$hgnc_symbol,
                        EntrezID = geneData$entrez,
                        Name=geneData$description,
                        moduleColor = modules)
```

```

Modmemb<-numeric()
for(i in 1:nrow(geneInfo0)){
  modColor<-modules[i]
  Modmemb[i]<-geneModMemb[i,modColor]
}
geneInfo0$MM<-Modmemb
geneOrder = order(
  factor(geneInfo0$moduleColor,levels = names(geneModMemb)), ~geneInfo0$MM)
geneInfo<-geneInfo0[geneOrder,]
geneInfo<-geneInfo[geneInfo$geneSymbol!="",]
return(geneInfo)
}

geneModuleMembership = as.data.frame(cor(datExpr[,notgrey], MEs, use = "p"))
names(geneModuleMembership) = substring(names(MEs),3)
geneInfo.all<-makeGeneInfo(an,moduleColors[notgrey],geneModuleMembership)
write.csv(geneInfo.all, paste(fname.root,"_INFO.csv",sep = ""))

```

Consensus network

```

dds <- DESeqDataSetFromMatrix (countData = exprColDat,
                              colData = colDat,
                              rowData = rowDat,
                              design = ~0+GroupDox)
ddsList<-list(AF=dds[,dds$Group=="AFDN_DT"],EV=dds[,dds$Group=="EV"])
ddsList<-lapply(ddsList, function(x){
  keep<-rowSums(counts(x))>=5; return(x[keep,])
})

ddsDeSeqList<-lapply(ddsList, DESeq)

lapply(ddsDeSeqList, dim)
# $AF
# [1] 15034      7
# $EV
# [1] 15174      8

ddsDeSeqListfilt<-lapply(ddsDeSeqList, countsGr10in30perfilt)

lapply(ddsDeSeqListfilt, dim)
# $AF
# [1] 8786      7
#
# $EV
# [1] 8943      8

vsList<-lapply(ddsDeSeqListfilt, function(x){
  y<-vst(x,blind = FALSE)
})

edataList<-lapply(vsList, function(x){
  x2<-as.data.frame(assay(x))
  return(x2)
})

colramp = colorRampPalette(c(3,"white",2))(20)

for (i in names(edataList)) {
  df<-edataList[[i]]
  pdf(file = paste(i,"densplot.pdf",sep = "_"))
  plot(density(df[,1]),col=colramp[1],lwd=3,ylim=c(0,.30))
  for(j in 2:8){lines(density(df[,j]),lwd=3,col=colramp[j])}
  dev.off()
}

library(limma)
norm_edataList<-lapply(edataList,function(x){
  x2<-normalizeQuantiles(as.matrix(x))
  return(x2)
})

```

```

}))

for (i in names(norm_edataList)) {
  df<-norm_edataList[[i]]
  pdf(file = paste(i,"densplotNorm.pdf",sep = "_"))
  plot(density(df[,1]),col=colramp[1],lwd=3,ylim=c(0,.30))
  for(j in 2:8){lines(density(df[,j]),lwd=3,col=colramp[j])}
  dev.off()
}

datExprList<-lapply(norm_edataList, function(x){
  x2<-as.data.frame(t(x))
  rownames(x2)<-colnames(x)
  return(x2)
})

genes<-intersect(names(datExprList$AF), names(datExprList$EV)) #8406

multiExpr<-multiData(AF=datExprList$AF[rownames(datExprList$AF)!="W16",genes], EV=datExprList$EV[,genes])

# use recommended power of 18 for <20 samples
sp = 18; ms = 30; ds = 1; mch = 0.15; Fname<-"18_30_1_15"

net = blockwiseConsensusModules(
  multiExpr, power = sp, minModuleSize = ms, deepSplit = ds,pamRespectsDendro = FALSE,
  networkType = "signed", mergeCutHeight = mch, numericLabels = TRUE,
  minKMEtoStay = 0.2, verbose = 5,maxBlockSize = 20000)

setLabels = c("AFDN-DT", "EV"); shortLabels = c("AF", "EV")
consMEs = net$multiMEs
consModuleLabels = net$colors;
consModuleColors = labels2colors(consModuleLabels)
consMEsC = multiSetMEs(multiExpr, universalColors = consModuleColors)
mods<-substring(names(consMEsC$AF$data),3)
mods<-mods[!mods%in%"grey"]
consMEs.ConsOrder<-consensusOrderMEs(consMEsC)

MyPlotEigenNets<-function(multiME){
  p<-plotEigengeneNetworks(
    multiME,
    setLabels,
    letterSubPlots = FALSE, Letters = NULL,
    excludeGrey = TRUE, greyLabel = "grey",
    plotDendrograms = FALSE, plotHeatmaps = TRUE,
    setMargins = TRUE, marDendro = NULL,
    colorLabels = TRUE, signed = TRUE,
    heatmapColors = NULL,
    plotAdjacency = FALSE,
    printAdjacency = FALSE, cex.adjacency = 0.9,
    coloredBarplot = TRUE, barplotMeans = FALSE, barplotErrors = FALSE,
    plotPreservation = "standard",
    printPreservation = FALSE, cex.preservation = 0.9,
    marHeatmap = c(3,3,2,1),
    xlimPreservation = c(0, 1),
    xLabelsAngle = 90)
  print(p)
}

pdf(paste("EigNet",Fname,".pdf",sep = ""))
MyPlotEigenNets(consMEs.ConsOrder)
dev.off()

```

Correlation with doxycycline

```

load("datTraitsAll.rda")
datTraits<-datTraits[,3:4]
labels<-read.csv("Cons18_30_1_15_2_labels.csv") #module colors converted to names
ylabs<-paste(labels$Label, " (",labels$Size,")",sep = "")

moduleTraitCorAF = cor(consMEs.ConsOrderAFdata,

```

```

        datTraits[rownames(multiExprAFdata),],
        use = "p")
moduleTraitPvalueAF = corPvalueStudent(moduleTraitCorAF,
                                       nrow(multiExprAFdata))
textMatrixAF = paste(signif(moduleTraitCorAF, 2),
                     " (", signif(moduleTraitPvalueAF, 1), ")",
                     sep = "")
dim(textMatrixAF) = dim(moduleTraitCorAF)

for(i in 1:dim(textMatrixAF)[1]){ #add * to significant results
  for (j in 1:dim(textMatrixAF)[2]) {
    if (moduleTraitPvalueAF[i,j]<0.05) {
      textMatrixAF[i,j]<-paste(textMatrixAF[i,j],"*",sep = "")
    }
  }
}

moduleTraitCorEV = cor(consMEs.ConsOrderEVdata,
                      datTraits[rownames(multiExprEVdata),],
                      use = "p")
moduleTraitPvalueEV = corPvalueStudent(moduleTraitCorEV,
                                       nrow(multiExprEVdata))
textMatrixEV = paste(signif(moduleTraitCorEV, 2),
                     " (", signif(moduleTraitPvalueEV, 1), ")",
                     sep = "")
dim(textMatrixEV) = dim(moduleTraitCorEV)

for(i in 1:dim(textMatrixEV)[1]){
  for (j in 1:dim(textMatrixEV)[2]) {
    if (moduleTraitPvalueEV[i,j]<0.05) {
      textMatrixEV[i,j]<-paste(textMatrixEV[i,j],"*",sep = "")
    }
  }
}

moduleTraitCor<-cbind(moduleTraitCorAF[1:23,],moduleTraitCorEV[1:23,])
textmatrix<-cbind(textMatrixAF[1:23,],textMatrixEV[1:23,])

xlabs<-c("Doxycycline\n(AFDN-DT)",
         "AFDN-DT\nPCR Expr.\n(AFDN-DT)",
         "Doxycycline\n(EV)",
         "AFDN-DT\nPCR Expr.\n(EV)")

png(file = "ConsTraitCorComb.png", width = 6.5,
    height = 3.5,units = "in",res = 300);par(
  mar = c(1.6, 11, .1, 0.1)
);labeledHeatmap(
  Matrix = moduleTraitCor,
  textMatrix = textmatrix,
  xLabels = xlabs,
  yLabels = names(consMEs.ConsOrderAFdata)[1:23],
  ySymbols = ylabs,
  yColorLabels = TRUE,
  colors = blueWhiteRed(50), setStdMargins = FALSE,
  cex.text = 0.6, xLabelsAngle = 0,
  cex.lab.y = 0.6,cex.legend = 0.5,
  cex.lab.x = 0.55,
  xLabelsAdj = c(0.5,0.5));dev.off()

```

GO enrichment

See Chapter 2 code, page 200.

Chapter 5

Basic process for RNA seq data

- counts downloaded from recount project (<https://jhubiostatistics.shinyapps.io/recount/>)
- scale counts
- DESeq
- limit to genes with counts > 10 in at least 30%
- vst (DESeq2)
- normalizeQuantiles (limma)

Creation of expression matrices

```
#source('http://bioconductor.org/biocLite.R')
#biocLite('recount')
library('recount')
library('DESeq2')
library(limma)
```

Colon

```
load("rse_geneSig34.rda")
rse<-rse_geneSig34
rseSig <- scale_counts(rse_geneSig34)
```

Scale counts with recount package

```
rse <- scale_counts(rse)
```

Run DESeq

control for RIN and batch

```
dds <- DESeqDataSet(rse, design = ~as.numeric(smrin)+smgebtch)
dds <- DESeq(dds)
```

Filter to counts>10 in 30%

```
n30percent<-.3*dim(rse)[2]
countsGr10in30per<-apply(counts(dds),1,function(x){
  x2<-ifelse(x>10,1,0)
  x3<-ifelse(sum(x2)>n30percent,TRUE,FALSE)
  return(x3)
})
dds <- dds[ countsGr10in30per, ]
```

Variance stabilization

```
vsd <- vst(dds, blind = FALSE)
edata=as.data.frame(assay(vsd))
```

Scatter plot and normalization

```
colramp = colorRampPalette(c(3,"white",2))(20)
plot(density(edataSig[,1]),col=colramp[1],lwd=3,ylim=c(0,.30))
for(i in 2:73){lines(density(edataSig[,i]),lwd=3,col=colramp[i])}

norm_edataSig = normalizeQuantiles(as.matrix(edataSig))
plot(density(norm_edataSig[,1]),col=colramp[1],lwd=3,ylim=c(0,.30))
for(i in 2:73){lines(density(norm_edataSig[,i]),lwd=3,col=colramp[i])}

datExprSig<-as.data.frame(t(norm_edataSig))
rownames(datExprSig)<-rse_geneSig34$Person
```

Repeat for Brain

```
cortFilt<-rse_geneCortex$Person%in%samps # limit to samples with sigmoid data
rse_geneCort34<-rse_geneCortex[,cortFilt]
dim(rse_geneCort34) #[1] 58037 34
```

```

rseCort34<-scale_counts(rse_geneCort34)
ddsCort<-DESeqDataSet(rseCort34,~as.numeric(smrin))
ddsCort<-DESeq(ddsCort)

n30percent<-0.3*dim(rseCort34)[2]
countsGr10in30per<-apply(counts(ddsCort),1,function(x){
  x2<-ifelse(x>10,1,0)
  x3<-ifelse(sum(x2)>n30percent,TRUE,FALSE)
  return(x3)
})

ddsCort <- ddsCort[ countsGr10in30per, ]
nrow(ddsCort) #24048
vsdCort <- vst(ddsCort, blind = FALSE)
edataCort=as.data.frame(assay(vsdCort))
colramp = colorRampPalette(c(3,"white",2))(20)

plot(density(edataCort[,1]),col=colramp[1],lwd=3,ylim=c(0,.30))
for(i in 2:33){lines(density(edataCort[,i]),lwd=3,col=colramp[i])}

norm_edataCort = normalizeQuantiles(as.matrix(edataCort))
plot(density(norm_edataCort[,1]),col=colramp[1],lwd=3,ylim=c(0,.30))
for(i in 2:33){lines(density(norm_edataCort[,i]),lwd=3,col=colramp[i])}

datExprCort<-as.data.frame(t(norm_edataCort))
rownames(datExprCort)<-rseCort34$Person

```

Remove X and Y chromosome genes

```

load("genedataAllcolon.rda") # probes

ensemblIDs <- sapply( strsplit( as.character(genedataAll$Probe), split="\\." ), "[", 1 )

library( "biomaRt" )
ensembl = useMart( "ensembl", dataset = "hsapiens_gene_ensembl" )
genemap <- getBM( attributes = c("ensembl_gene_id", "entrezgene", "hgnc_symbol", "description", "chromosome_name"),
                  filters = "ensembl_gene_id",
                  values = ensemblIDs,
                  mart = ensembl )

idx <- match( ensemblIDs, genemap$ensembl_gene_id )
genedataAll<-data.frame(Probe=genedataAll$Probe, ENS=ensemblIDs,EntrezID=genemap$entrezgene[idx],Symbol=genemap$hgnc_symbol[idx], Name=genemap$description[idx],Chr=genemap$chromosome_name[idx])

noXY<-as.character(genedataAll$Probe[!genedataAll$Chr%in%c("X","Y")])

datExprSigNoXY<-datExprSig[,names(datExprSig)%in%noXY]
datExprSigNoXYconsSamp<-datExprSigNoXY[ConsensusSample,]

datExprCortNoXY<-datExprCort[,names(datExprCort)%in%noXY]
datExprCortNoXYconsSamp<-datExprCortNoXY[ConsensusSample,]
save(datExprCortNoXY,file = "datExprCortNoXY.rda")
save(datExprCortNoXYconsSamp,file = "datExprCortNoXYconsSamp.rda")

Genes<-intersect(names(datExprCortNoXYconsSamp),names(datExprSigNoXYconsSamp))
datExprBrain<-datExprCortNoXYconsSamp[,Genes]
datExprColon<-datExprSigNoXYconsSamp[,Genes]
multiExpr<-multiData(Colon=datExprColon, Brain=datExprBrain)

```

Of the 34 samples with sigmoid colon and cerebral cortex there are 29 that are not outliers in either (outliers based on sample clustering in WGCNA)

Softpower threshold 9 worked for both brain and colon. For signed network, power is doubled.

WGCNA

Tissue-specific and consensus networks

Run WGCNA as in Chapter 4 (Tissue-specific: page 210, Consensus: page 212) with the following parameters (used for both tissue specific and consensus analyses)

```
library(WGCNA)
options(stringsAsFactors = FALSE)
allowWGCNAThreads()

#parameters-----
softPower = 18
netType = "signed"
minModuleSize = 80
deepSplit = 1
MEDissThres = 0.15
```

Unified network

Combine top 10,000 by max variance

```
load("datExprSigNoXYconsSamp.rda")
load("datExprCortNoXYconsSamp.rda")

c<-datExprSigNoXYconsSamp
b<-datExprCortNoXYconsSamp

names(c)<-paste("c",names(c),sep = "_")
names(b)<-paste("b",names(b),sep = "_")

tb<-as.data.frame(t(b))
tb$var<-apply(tb,1,var)
tb<-tb[order(tb$var,decreasing = TRUE),]
topVarB<-rownames(tb)[1:10000]
b<-b[,names(b)%in%topVarB]

tc<-as.data.frame(t(c))
tc$var<-apply(tc,1,var)
tc<-tc[order(tc$var,decreasing = TRUE),]
topVarC<-rownames(tc)[1:10000]
c<-c[,names(c)%in%topVarC]

datExpr<-cbind(b,c)
```

Run WGCNA

```
# parameters
softPower = 12
netType = "signed"
minModuleSize = 80
deepSplit = 1
MEDissThres = 0.15
```

Gene Ontology Enrichment

See Chapter 2 code, page 200.