Functional Genomics at the EHEC-Intestinal Interface: Mechanisms of Pathogenicity and the Host Response

HARVARD UNIVERSITY Graduate School of Arts and Sciences

DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the Division of Medical Sciences in the subject of Microbiology and Immunobiology have examined a dissertation entitled

Functional Genomics at the EHEC-Intestinal Interface: Mechanisms of Pathogenicity and the Host Response

presented by Alyson R. Warr

candidate for the degree of Doctor of Philosophy and hereby certify that it is worthy of acceptance.

Signature: _____

Typed Name: Dr. Simon Dove

Signature: _____

Typed Name: Dr. Sophie Helaine Wayne Lencer

Signature: _____ Wayne Lencer (Dec 3, 2020 12:03 EST)

Typed Name: Dr. Wayne Lencer

Signature: _____ Vanessa Sperandio (Dec 3, 2020 10:21 CST)

Typed Name: Dr. Vanessa Sperandio

Date: November 17, 2020

Functional Genomics at the EHEC-Intestinal Interface: Mechanisms of Pathogenicity and the Host Response

A DISSERTATION PRESENTED

BY

Alyson R. Warr

TO

The Division of Medical Sciences

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

Doctor of Philosophy

IN THE SUBJECT OF

Microbiology and Immunobiology

Harvard University

Cambridge, Massachusetts

November 2020

Dissertation Advisor: Dr. Matthew K. Waldor Alyson R. Warr

Functional Genomics at the EHEC-Intestinal Interface: Mechanisms of Pathogenicity and the Host Response

Abstract

Enterohemorrhagic *Escherichia coli* (EHEC) is a human foodborne pathogen that infects the colon, leading to hemorrhagic colitis and the potentially fatal hemolytic uremic syndrome. This dissertation aims to address the host-pathogen relationship between EHEC and the colonic mucosa *in vivo*. We used functional genomics to examine both requirements for EHEC pathogenicity and the enteric transcriptional response to infection in infant rabbits, a physiologically relevant animal model that closely recapitulates human EHEC intestinal disease.

We used transposon-insertion sequencing screens to define EHEC genes required for both growth *in vitro* and colonic colonization. A transposon mutant library of over 100,000 unique EHEC mutants was analyzed to identify mutants whose frequency was reduced in different conditions. Underrepresented mutants provided knowledge of the pathogen's essential gene set; moreover, passage of the transposon library through the infant rabbit colon yielded insight into the EHEC loci required for optimal intestinal colonization (Chapter 2). We identified many new colonization-promoting genes, including the widely conserved inner membrane protein CvpA (Chapter 2). Functional characterization of CvpA revealed that it is a member of the EHEC bile resistance repertoire and is linked genetically to the σ^E extracytoplasmic stress response (Chapter 3).

We also assessed how EHEC's potent toxin, Shiga toxin (Stx), modifies the histopathologic and transcriptional response to infection in the colon. We found that Stx is required for severe, hemorrhage and extensive apoptosis in the colon. Cellular-compartment-specific transcriptomics revealed that Stx dramatically remodels the host transcriptome during infection, particularly in

epithelial cells. Genes related to coagulation and immune signaling pathways were differentially expressed in the presence of Stx, revealing the importance of this virulence factor in shaping the host response to infection (Chapter 4.)

Collectively, these findings deepen understanding of the biology at the EHEC-intestinal interface. As no treatment exists for EHEC infection aside from rehydration therapy, these data will be useful to inform future work on development of new therapeutics.

Contents

Listing of Figures

Listing of Tables

This thesis is dedicated to my grandmother Claire,
who dreamed as a little girl of going to Harvard.

Acknowledgements

I am indebted to many people who have helped me during this long and arduous journey.

Firstly, I express gratitude to my PhD advisor, Dr. Matthew Waldor. Matt has an *infectious* enthusiasm for science. Few professors will sit by your side editing your first manuscript or dissect a mouse with you – and I benefited enormously from that mentorship. Thank you, Matt, for accepting me into your lab and giving me the freedom to explore my own scientific path.

I owe so much to the scientists in the Waldor lab who were great colleagues and friends. It's easier to clean up after a failed experiment side-by-side with a friend offering words of wisdom, a reason to laugh, and the hope to try again tomorrow. I am especially grateful to my first mentors Carole and Troy, who trained me with patience when I was just starting. I also give special thanks and express enormous gratitude to my lab mate, classmate, and friend, Brandon, who is a diligent scientist and truly selfless.

I am grateful to have been part of the Microbiology community at HMS for the last 6 years. I met so many amazing scientists who were also genuinely kind and helpful. I am extremely grateful to the Microbiology curriculum fellow Deepali, who is a never-ending source of support and advocacy. There are so many little things I will miss, like beer hour and caramel brownies at Tuesday talks.

I give thanks to my dissertation advisory committee (Drs. Eric Rubin, Sarah Fortune and Simon Dove) for excellent feedback over the years and to my dissertation defense committee (Drs. Simon Dove, Sophie Helaine, Wayne Lencer and Vanessa Sperandio) for serving as my examiners.
I also thank the great mentorship of my undergraduate research advisor, Dr. Steven Sandler. In his laboratory I made my first discoveries and became fascinated with *E. coli*. Steve encouraged me to be very independent as an undergraduate researcher, which is why I had the nerve to pursue a PhD.

Thank you to the BBS office – Kate, Anne, Danny – who helped me on so many occasions, especially when I was a lost first year. I am extremely grateful to my classmates and friends who provided a much-needed network of support over the last 6 years – thank you for sticking with me to the very end. In particular: Kate, Olivia, Sam, Brandon, Pauline, and Greg – you are all fantastic scientists – thank you for challenging me, inspiring me, and helping me. Thank you to Greg's van for surviving 5 long drives to Provincetown. Our G2 skit was legendary, and the real lesson was the friends we made along the way. I was grateful to find other communities outside of the lab at Harvard in both Science in the News and the mental health initiative. I am so happy to have worked with so many amazing individuals who sought to leave the graduate school experience a better one than the way we found it. Thank you also to the entire UMass crew and especially to Nasim, Kelsie, Liz, Sarah, Maggie, Daryl, Lia, Heika, Brit, and Catie who have encouraged me for years.

I thank my supportive family: my parents and my brother Matt. My parents fostered my love for science from an early age. My dad has a tenacity and work ethic I admire deeply, and his forceful determination has served as a reminder to be relentless in the pursuit of my dreams. My mom is an excellent nurse and taught me how to have compassion for the world and put my brain to good use helping others. During grad school I gained a new family, the Manteigas, who have filled my life with new experiences, joys and laughter. Thank you for welcoming me.

Lastly, I express deep gratitude and appreciation to my husband John, who is a world-class scientist and my best friend. His encouragement and love served as my buoy during this most challenging period of my life. We are now embarking on a new road, which is both terrifying and exhilarating. I know the challenges we face ahead will be brighter and easier to carry simply because of the enduring love we have for each other. Thank you for helping me achieve my dream.

1

Introduction

Diarrheal diseases are a significant global public health burden

Globally, diarrheal disease is among the leading causes of human mortality. In 2019, 1.5 million deaths were attributable to diarrheal illness[1] (Figure 1.1A). Diarrheal disease is especially dangerous early in life, accounting for 10% of all deaths in children under the age of five (Figure 1.1B)[1,2]. Even in non-fatal cases in children, bouts of diarrhea greatly diminish quality of life, exacerbating malnutrition, impairing immunity, and leading to delays in cognitive development[3]. Diarrheal diseases are mainly caused by gastrointestinal (GI) infections with viral and bacterial pathogens, which are typically acquired by ingestion of contaminated food or water.

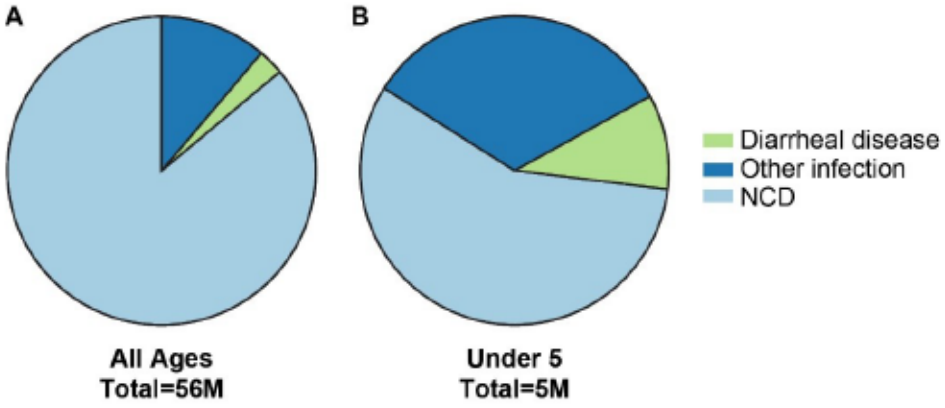


Figure 1.1: The global burden of diarrheal disease.

A) Of 56 million (M) deaths across all age groups in 2019, 3% were attributed to diarrheal diseases (green), 11% to other infections (dark blue), and 86% to non-communicable diseases (NCD, light blue). B) Of the 5 million deaths in children under the age of five years old in 2019, 10% of deaths were attributable to diarrheal disease, 33% other infections, and 57% NCDs. Data from Global Burden of Disease Study 2019 Results.

The World Health Organization (WHO) considers diarrhea to be 'preventable and treatable'. Implementation of preventative 'WASH' (water safety, sanitation, and hygiene) techniques and oral rehydration therapy (ORT) have contributed to the decline in worldwide mortality from diarrheal diseases for children and adults under the age of 70 over the last four decades[4]. Despite progress, this preventable disease remains a major killer of young children, and for adults who are 70 and older the proportion of deaths

attributable to diarrheal diseases is actually on the rise[5]. The WHO seeks to eliminate childhood deaths from diarrhea through a combination of preventative and treatment strategies including WASH, ORT, vitamin A supplementation, and development of affordable and efficacious therapeutics including vaccines[6]. Introduction of a vaccine against the diarrheal pathogen rotavirus, for example, reduced mortality in children around the world[7,8]. There are no licensed vaccines for other diarrheal pathogens aside from *Vibrio cholerae*[9].

The subject of this dissertation is the diarrheal pathogen enterohemorrhagic *Escherichia coli* (EHEC). EHEC has gained widespread attention as a public health risk since its emergence 40 years ago as the causative agent of a severe form of diarrheal illness. Consumption of food contaminated with EHEC, typically ground beef or leafy greens, often leads to a self-limiting hemorrhagic colitis, but can progress to severe complications causing renal or nervous system damage in 5-10% of patients. Hemolytic uremic syndrome (HUS) is a potentially lethal complication of EHEC infection, characterized by anemia, thrombocytopenia, and acute kidney injury. HUS is one of the leading causes of acute kidney injury in children[10]. Young and aged populations are at higher risk of adverse outcomes; children are particularly susceptible to severe complications[11], and individuals over the age of 60 are the most likely to die from EHEC infection[11]. Supportive rehydration therapy is the primary treatment for EHEC infection. Antibiotics are associated with a higher frequency of complications and therefore contraindicated[12,13], and there is currently no vaccine or therapeutic available[14]. Tightening of food safety regulations after EHEC outbreaks in the 1990s limited the spread of foodborne disease in the following decades, but 30 years later this pathogen remains a threat to public health[15]. Global analyses estimate EHEC causes up to 2.8 million cases annually, with the highest incidence in the WHO-defined Regions of the Americas, Eastern Mediterranean Region, and South-East Asia Region[16].

A deep understanding of the pathogen-host interaction can greatly accelerate rational design of beneficial therapeutic interventions for EHEC. Historically, this type of analysis has been difficult, as *in vivo* studies of EHEC have been hampered due to lack of a suitable animal model. Mice, the typical laboratory model system of choice, do not exhibit diarrhea or colonic pathology after oral inoculation with EHEC[17,18]. In contrast, infant rabbits have proven to be a physiologically relevant model of EHEC infection; infected animals

develop diarrhea and intestinal pathology that closely resembles that observed in human EHEC disease[19–23]. With this model in hand, and with the advent of high-throughput sequencing and high-resolution comprehensive functional genomic approaches, we were poised to amass a wealth of knowledge about the host-pathogen interactions that occur during EHEC infection. We first conducted an in vivo bacterial transposon-insertion mutagenesis screen and identified the set of genes required for EHEC to grow in vitro and colonize the colon[24]. These colonization promoting-genes included those studied previously, such as components of the Type 3 secretion system, and many new ones (Chapter 2). We further investigated the function of one of these genes, *cpxA*, and discovered that it is part of the EHEC bile resistance repertoire and encodes a widely conserved inner membrane protein linked to the σ^E-extracytoplasmic stress response[25] (Chapter 3). We also used cellular compartment-specific transcriptomics to characterize the host response to EHEC and its potent exotoxin, Shiga toxin (Stx), and discovered that over 500 host genes are transcribed specifically in response to Stx[26]. Many of these genes encode proteins that mediate coagulation and immune signaling pathways, highlighting the ability of Stx to powerfully remodel the host response to infection (Chapter 4).

Escherichia coli as a commensal and a pathogen

The bacterial species *Escherichia coli* is a Gram-negative, facultative anaerobe and a member of the class Gammaproteobacteria of the family Enterobacteriaceae. The species was first isolated from the stool of infants in 1885 by the German scientist Dr. Theodor von Escherich[27]. As a fast-growing, metabolically flexible bacteria, *E. coli* quickly rose to prominence after its discovery to become one of the most widely used model organisms. Studies using *E. coli* led to some of the most fundamental discoveries of modern biological science, such as the semi-conservative nature of DNA replication[28].

Today, *E. coli* is recognized as a member of the human GI microbiota, the consortium of commensal microbes which live symbiotically in the GI tract. *E. coli* finds its niche in the colon, where microbial density reaches a staggering 10^{14} cells[29,30]. Enterobacteria with the ability to consume oxygen, like *E. coli*, colonize the infant human colon within days of birth[31–35]. The consumption of oxygen in the infant colon by these pioneer

species creates a suitable environment for the strictly anaerobic members of the consortium such as *Bacteroidetes sp.* that eventually dominate. In adults, *E. coli* has variable abundance in individuals[36], but is generally thought to compose less than 0.1% of the total microbiota[37]. Although it is not a major member of the microbiota, *E. coli* provides important functions for the host. *E. coli* produces essential vitamins for host usage such as K_2[38], and as a member of the microbiota, provides at least two mechanisms of defense against invading pathogens. The first is through colonization resistance, a phenomenon where pathogens are unable to access the intestinal niche due to prior occupancy by commensals[39,40]. The other is immunological training, where the presence of diverse commensal microbes in the intestine early in life guides the development of an appropriately responsive immune system[41–43].

Not all *E. coli* strains are commensals; several have evolved virulence attributes. These pathogenic strains are historically categorized into 'pathotypes' based on possession of certain virulence factors and a stereotypical set of symptoms observed during infection. There are 7 major diarrheagenic pathotypes of *E. coli*: Shiga toxin-producing (STEC), enteropathogenic (EPEC), enterotoxigenic (ETEC), enteroaggregative (EAEC), enteroinvasive (EIEC, which includes *Shigella sp.*), diffusely adherent (DAEC), and adherent-invasive *E. coli* (AIEC)[44–48]. There are also several non-diarrheagenic pathotypes of *E. coli* that colonize organs other than the intestine, collectively referred to as ExPEC, or extra-intestinal *E. coli*. ExPEC pathotypes include strains which cause urinary tract infections (UPEC), sepsis (SEPEC), and neonatal meningitis (NMEC)[46,49]. Some isolates of pathogenic *E. coli* do not fit neatly into these categories, necessitating the descriptor 'hybrid' or 'atypical' to describe strains with a mix of virulence attributes from different pathotypes, such as Shiga toxin-producing enteroaggregative *E. coli* [45], or atypical EPEC[50].

Identification of hybrid and atypical isolates led to attempts to classify *E. coli* strains by genomic similarity instead of phenotype or possession of virulence factors. To date, there are over 100,000 publicly available sequenced *E. coli* genomes, and comparisons have revealed a remarkable genetic diversity across *E. coli* isolates[51]. The species as a whole has a large pan-genome of about 20,000 genes, with each individual genome containing about 5,000 genes. Less than 2,000 of these genes are conserved across all isolates, constituting the core genome[51–57]. The additional 3,000 accessory genes unique to each isolate are typically acquired

horizontally, either from integration of phage elements, acquisition of plasmids, or other recombination events[51–57]. This genetic plasticity explains the diversification of phenotypically distinct strains and the adeptness of *E. coli* in carving out new niches[58–60]. Creating phylogeny from these *E. coli* genomes has led to a 'phylogroup' classification system, where as many as 14 distinct phylogroups are recognized[51]. Interestingly, many phylogroups contain both commensal and pathogenic strains, arguing for acquisition of virulence several independent times.

STEC is considered the most dangerous and notorious type of pathogenic *E. coli*. Enterohemorrhagic *E. coli* (EHEC), the focus of this dissertation, is characterized as a subset of STEC that produce and release one or more Shiga toxins (Stx) and contain the Locus of Enterocyte Effacement (LEE), a gene locus enabling intimate attachment to the colonic epithelium. There are seven major serotypes EHEC, which are classified by their unique O (LPS) and H (flagella) antigen structure[61]. Serotype O157:H7 is the most prominent worldwide, but non-O157 serotypes are gaining in prevalence[62]. The strain EDL933, the prototypical EHEC O157:H7 strain, was isolated from contaminated ground beef from the first recognized outbreak in 1982[63], and is used in all experiments presented in Chapters 2-4.

Emergence of a pathogen: evolution of EHEC

EHEC was first identified in 1982 as the causative agent of 'hamburger disease' after a mysterious outbreak linked to ground beef served at McDonald's sickened Americans in Oregon and Michigan[64–68]. The pathogen gained widespread recognition and notoriety after an outbreak linked to contaminated beef patties served at the fast-food restaurant Jack in the Box sickened more than 700 Americans across several states in 1993[69–71]. The majority of the victims were children, several of whom required extensive kidney dialysis. In four children, the infection was fatal[71].

The sudden appearance of this new virulent serotype of *E. coli* was initially surprising. Retrospective analyses revealed that illness similar to that caused by EHEC (HUS preceded by a prodromal diarrhea phase) had occasionally been observed in human patients and animals before the 1980s[72–78], but overall the incidence was extremely low. Within years it had progressed from a 'rare serotype'[67] to a widespread human pathogen.

Changes in food processing systems in the late 20th century are likely to blame for the emergence of EHEC. In the 1970s, the massive industrialization of beef and poultry production, often with non-hygienic conditions, led to increased incidence of several bacterial foodborne pathogens, including *Salmonella,* and the emergence of new pathogens like *Campylobacter* and EHEC which have natural reservoirs in food animals[79]. Cattle represent a significant reservoir of EHEC[80]. EHEC colonizes the GI tract of up to 70% of these ruminants asymptomatically[81], and cattle can shed up to 10^4 CFU of EHEC per gram of feces[82]. 'Super-shedder' animals are likely a significant source of contamination of beef carcasses and leafy greens[83,84]. The rise of foodborne illness during this time, and the 1993 EHEC outbreak in particular, prompted several key changes to food safety systems around the world, including increased scrutiny of animal product inspections and more robust food recall systems[15,85,86]. These changes reduced the incidence of EHEC infection in the United States during the early 2000s by 46%[15], but outbreaks continue to sicken Americans every year[87].

EHEC O157:H7 diverged from an ancestral strain as recently as 400 years ago[88]. The ancestral strain is *E. coli* O55:H7, which is Stx-negative and causes non-bloody diarrhea. Evolution likely occurred in a step-wise fashion that included acquisition of novel virulence attributes through horizontal gene transfer (HGT)[88–92]. For example, Shiga toxin was acquired by EHEC through chromosomal integration of bacteriophage genomes that contained the Stx genes. These phage likely originally acquired this DNA Stx genes in *Shigella dysenteriae*[93–96]. This propensity for HGT makes *E. coli* notable as an extremely versatile species and a seemingly limitless ability to adapt to new niches. More recently, in 2011, a new Shiga-toxin producing *E. coli* strain emerged (O104:H4) and sickened thousands of people in Germany in an outbreak linked to contaminated sprouts[97–101]. This hybrid EAEC/STEC isolate illustrates the power of HGT to produce new variants of *E. coli* pathogens and underscores the importance of therapeutic strategies to combat mobile virulence factors like Stx that can transfer between strains.

Entry into the host: initial innate barriers

Contiguous with the external environment, the human GI tract is constantly exposed to foreign material and microbes. A multitude of barriers exist to prevent microbes from colonizing the human GI tract and causing

disease. The initial barriers a microbe faces are part of the innate immune system, and these physical, chemical, and biological 'firewalls' are typically extremely successful in preventing microbial infiltration. Pathogens such as EHEC, though, are highly evolved to evade and overcome these host barriers.

Upon consumption of contaminated food, EHEC enters the acidic stomach (pH 1.5-3.5), which is an important barrier for pathogen elimination[102,103]. Enteric pathogens typically induce an acid tolerance response (ATR) upon exposure to acidic conditions that can be protective up to pH 3.0[104]. This response includes removing or sequestering acidifying hydrogen ions from the cell cytosol to maintain internal pH homeostasis[104,105]. Invading enteric bacteria that rely on ATRs, such as *Vibrio cholerae*, cannot survive prolonged exposure to gastric juices. Only a large infectious dose ($>10^6$ cells) permits passage of a few cells by chance to the intestine to establish disease. In comparison, EHEC has additional mechanisms that enable extreme acid resistance (XAR), permitting prolonged survival in pH < 2.5. XAR systems include antiporters which pair decarboxylation of amino acids with export of cytoplasmic hydrogen ions[104–108]. XAR mechanisms are likely to contribute to the extremely low infectious dose of EHEC and other similarly acid-resistant organisms like *Shigella*, which can cause disease after consumption of as little as 10-100 cells[70,104,109,110]. Mutations in XAR genes likely lead to lower infectivity[111].

After passage through the stomach, EHEC enters the small intestine (SI). The intestinal environment poses significant challenges to EHEC, including substantial fluctuations in temperature, osmolarity, oxygen concentration, and nutrient availability[112,113]. Mechanical shear force from peristalsis and host-derived antimicrobial peptides (AMPs), kill microbes that move through the central cavity of the intestine, the lumen[112,113]. Enteric pathogens must also survive exposure to bile, an aqueous secretion of the liver that contains bile salts, potent antimicrobial compounds. Bile salts cause a wide range of deleterious effects in bacterial cells, including damage to the cell envelope and DNA, widespread protein misfolding, and redox stress[114,115]. Mechanisms for resisting these insults include efflux pumps to remove bile salts from the bacterial cytosol, chaperones to ameliorate protein folding stress, and global stress response pathways[114,115]. Chapters 2 and 3 detail the identification and characterization of a novel member of the EHEC bile resistance repertoire,

CvpA[24,25]. We discovered that CvpA is linked genetically to the σ^E-extracytoplasmic stress response, and that activation of this response is required for optimal bile resistance[25].

The colonic niche

EHEC preferred niche is the colon. Functionally, the colon absorbs fluid from waste, houses the abundant GI microbiota, and maintains a robust mucosal immune system poised to defend against the constant barrage of foreign material. Anatomically, the colon is composed of four layers of tissue surrounding the lumen (Figure 1.2A). The three outermost layers (serosa, muscularis, and submucosa) contain connective tissue that provides structure, muscles that contract to provide peristalsis motion, a nerve plexus that innervates the tissue, and vessels of both the circulatory and lymphatic system. The innermost layer is the mucosa, which has three major components: the epithelium, lamina propria, and a muscle layer called the muscularis mucosae. As EHEC passes through the colonic lumen it interfaces with the microbiota and colonic mucosa.

In the lumen, EHEC faces competition with the teeming microbiota for nutrients and physical space[116]. Nutritionally, the colon is almost devoid of free glucose, as simple dietary monosaccharides are absorbed almost entirely in the SI[116,117]. Low glucose concentration in the colon triggers changes in EHEC gene expression, activating the expression of virulence and colonization genes[116,118]. The main source of carbohydrates in the colon are complex polysaccharides like starch and fiber that pass through the SI undigested and are fermented in the colon by commensal bacteria[119,120]. Commensals like *Bacteroides thetaiotaomicron* express a suite of metabolic enzymes which digest polysaccharides into simpler sugars which can be taken up by EHEC for its own metabolism and as a signal to activate virulence gene expression[117,121–124]. Commensals also produce short-chain fatty acids (SCFAs), which EHEC can also detect to regulate gene expression. For example, butyrate can induce genes required for colonization and motility in EHEC[125,126]. SCFAs can also be protective to the host during EHEC infection; acetate is thought to limit damage in the colon by fortifying the intestinal barrier[127].

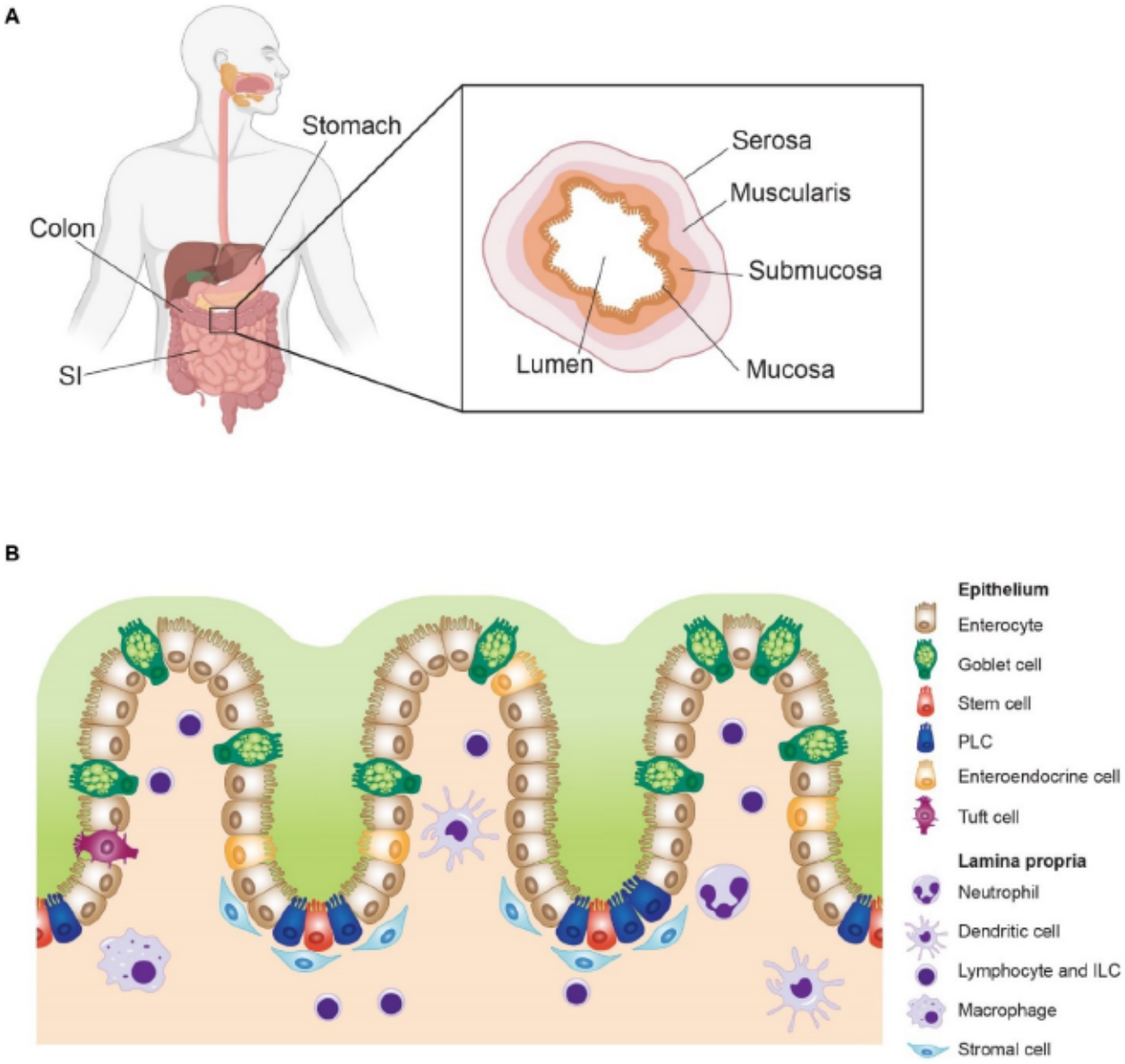


Figure 1.2: Human colon anatomy.

A) The gastrointestinal tract is a tube contiguous with the external environment. Consumed EHEC cells first encounter the stomach, then travel through the small intestine (SI) and find their niche in the colon. Inset on left displays a cross section of the four tissue layers of the colon that surround the inner lumen. Image adapted from BioRender. B) Cellular components of the colonic mucosa. The green layer above epithelium represents the colonic mucus layer. PLC: Paneth-like cell. ILC: Innate-lymphoid cell.

Directly adjacent to the colonic lumen is the mucus layer. Specialized epithelial cells called goblet cells secrete gel-forming mucins to establish this barrier. The mucus barrier is comprised of a looser outer layer containing commensals and an adherent sterile inner layer[128,129]. *B. thetaiotaomicron* can cleave intestinal mucin glycoproteins, liberating the sugar fucose into the intestinal lumen. EHEC senses fucose with a two-component system, FusKR, to activate genes required for colnonization[130,131]. EHEC also secretes a protease, StcE, which degrades mucus and permits access to the underlying epithelium[132].

The epithelium (Figure 1.2B) is a single layer of epithelial cells and constitutes a critical part of the host innate defense system, representing the first cellular barrier between the external world and the internal sterile environment[133]. Epithelial cells are joined tightly together to create an impermeable barrier to microbes. This barrier is formed by three main types of transmembrane junctional protein complexes which provide extracellular linkages between cells[134]. These complexes also interact with intracellular cytoskeletal proteins, which modulate cell structure and dictate cellular permeability[134,135]. Tight junctions are formed by interactions between transmembrane proteins of the occludin and claudin families. Adherens junctions are formed by Ca^{2+}-dependent bonds between cadherin proteins[134]. Desmosomes also mediate Ca^{2+}-dependent inter-cellular linkages[134]. Inflammatory stimuli regulate the tightness of junction barriers and expression of junction proteins to facilitate barrier maintenance[136].

Structurally, the colonic epithelial sheet is flat, lacking the villus protrusions characteristic of the SI. This sheet contains invaginations called the Crypts of Lieberkühn (crypts), where pluripotent stem cells reside[137–142]. These stem cells proliferate into a limited set of epithelial progeny cells, which migrate from the crypt towards the epithelial surface[143]. It is estimated that the lifetime of mature colonic epithelial cells is very short, approximately 5-7 days[144–146]. The rapid cell turnover of the colon maintains barrier integrity, as dead or damaged cells are quickly sloughed and replaced. The most abundant cell type (~80% of epithelial cells) is the enterocyte[142]. Enterocytes are simple columnar epithelial cells that provide an absorptive function to the colon. Their apical surface is covered with protrusions of the cell membrane called microvilli which increase surface area.

The colon also contains several specialized epithelial secretory cell types. The most abundant of these are mucin-secreting goblet cells. Other secretory cells release molecules which damage microbes, stimulate the immune system, and modify the behavior of other cells in the colon. Deep-crypt secretory cells (REG4+ cells similar to Paneth cells in the SI[142,147]) release antimicrobial peptides (AMPs) that damage bacterial membranes[148]. They also release cytokines that regulate stem cell signaling in response to stimuli such as barrier breach, and help maintain the proliferative stem cell niche[142,147]. Enteroendocrine cells produce and release peptide hormones into the bloodstream that regulate appetite, insulin release, motility, and intestinal inflammation[142,149]. Tuft cells secrete cytokines and are important for regulation of immune responses against helminths[150,151]. Transcriptomic analyses at single cell resolution have revealed that intestinal epithelial cells types can also further subspecialize, and have helped characterize their role during infection[152–155].

Beneath the epithelial layer is the lamina propria (LP). The LP is a matrix of loose connective tissue which contains cells of both mesenchymal and hematopoietic lineages that support and respond to epithelial stimuli. Mesenchymal stromal cells form a matrix of support beneath the epithelium, and likely have additional roles in bridging immune signaling between the epithelial cells and the LP immune cells[156–158]. Even in a non-diseased state, a variety of immune cells can be found in the LP. These include innate immune tissue-resident 'sentinel cells' such as tissue-resident macrophages and dendritic cells (DCs), which survey the tissue for damage and pathogens[159–161]. Neutrophils are not abundant in non-inflamed tissue, but increase in number rapidly during infection. Lymphocytic cells are numerous in the LP, and include CD4+ T cells, CD8+ T cells, and plasma cells[159–161]. Plasma cells are also present and actively secreting large amounts of IgA antibody, which aids in mucosal defense[162]. Innate lymphoid cells (ILCs) are also present, and help bridge signaling between the innate and adaptive immune responses[159–161,163]. Macroscopically, structures of lymphoid follicles analogous to SI Peyer's patches are found throughout the colon[159]. Specialized epithelial cells called Microfold (M) cells reside in these regions, and sample antigens from the colonic lumen which are presented to the underlying mature naïve lymphocytes[159,164].

The Type 3 secretion system facilitates secure epithelial attachment

After passage through the mucus layer, EHEC attaches to epithelial cells. Initial attachment is mediated by EHEC adhesin proteins including long polar fimbriae (LPF) and Curli fibers[165]. LPF and Curli bind tightly to extracellular matrix proteins, including fibronectin, laminin and collagen[166–168]. When in close proximity to epithelial cells, EHEC can engage its type three secretion system (T3SS) to mediate secure attachment. This intimate attachment is characterized by attaching and effacing (A/E) lesions, where EHEC binds tightly to the epithelial cell surface on actin pedestals (Figure 1.3A), remodeling host cell actin and destroying the absorptive microvilli[169,170]. The T3SS is a molecular syringe which allows EHEC to translocate effector proteins directly into the host cell cytosol (Figure 1.3D)[171,172]. To form A/E lesions, EHEC injects proteins Tir and EspFu into the host epithelial cell cytosol via the T3SS. Tir inserts into the host cell membrane, and cytosolic EspFu mediates interaction between Tir and host N-WASP, stimulating the actin-nucleating complex Arp2/3 and promoting polymerization of host cell actin beneath EHEC[173–176]. EHEC displays on its surface the adhesin Intimin, which interacts with Tir on the host cell membrane. This interaction creates a strong bond to keep EHEC at the epithelial cell surface (Figure 1.3B)[177–183]. Tir and Intimin are required for EHEC to colonize the colon; mutations in these genes renders the resulting strain unable to reach the same burden in infected animals or cause severe disease[184,171,185,180,186,177,182,21].

The T3SS structural genes are encoded within a pathogenicity island named the Locus of Enterocyte Effacement (LEE) (Figure 1.3C)[187]. Of EHEC's 49 effector proteins, 6 are encoded within the LEE (*tir, espF, espG, espH, espZ,* and *map*) and 43 are encoded elsewhere in the genome and are termed non-LEE encoded effectors (Nle). The contributions of Tir and the EspFu to colonization have been studied in some detail, but the contribution of other effector proteins to EHEC colonization has received less attention. We completed a screen of >100,000 EHEC mutants to identify proteins required for colonization, and a thorough analysis of the contribution of all effector proteins is discussed in detail in Chapter 2[24].

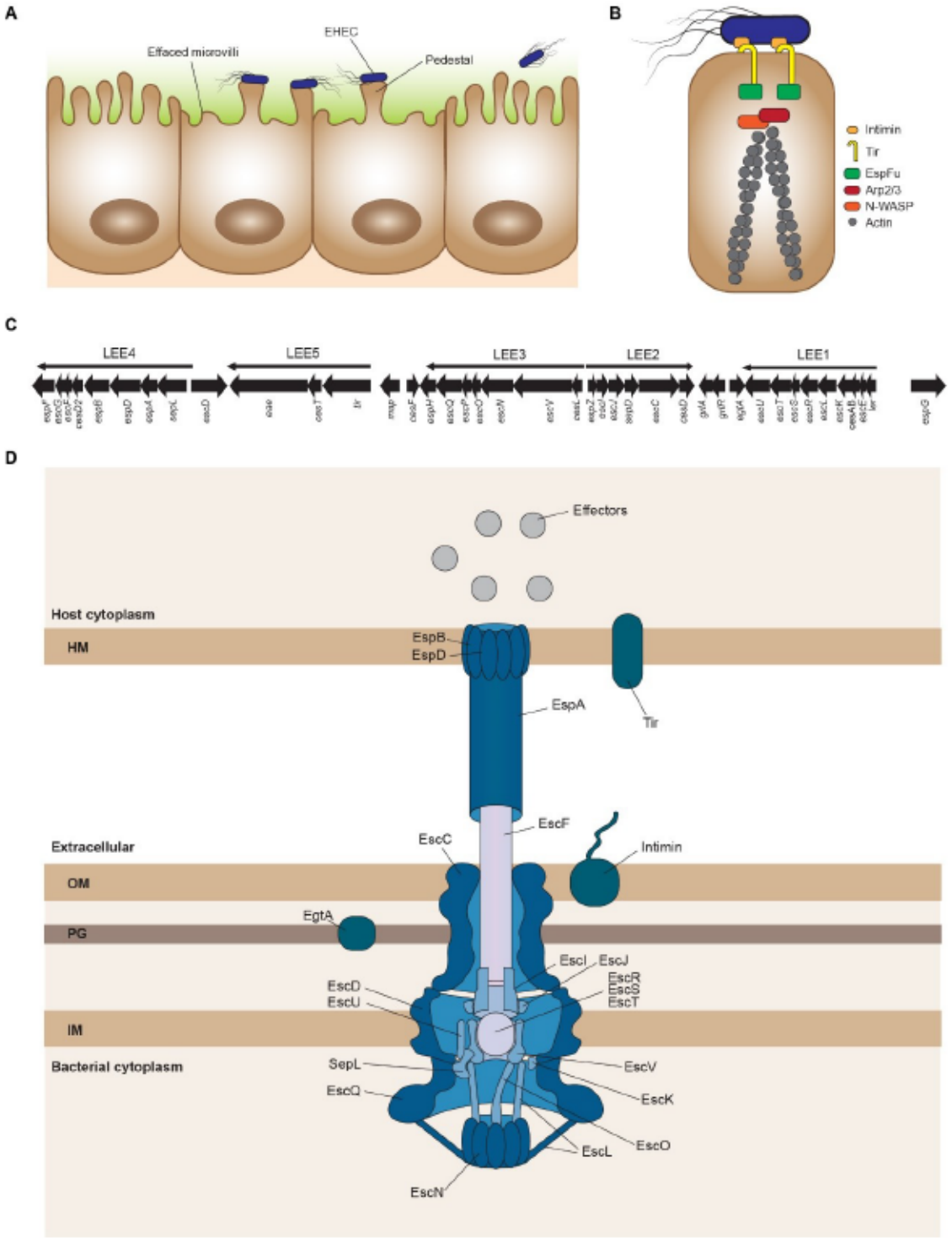


Figure 1.3: EHEC attachment to the colonic epithelium is mediated by the T3SS.

A) EHEC formation of attaching and effacing lesions on colonic enterocytes. B) Protein-protein interactions creating the pedestal. C) Genes in the Locus of Enterocyte Effacement. D) Structure of the type three secretion system crossing the bacterial inner membrane (IM), peptidoglycan (PG), outer membrane (OM) and penetrating the host cell membrane (HM) to deliver effector proteins into the host cytoplasm.

Gene expression in the LEE is activated by the transcriptional regulator protein Ler. An extensive network of signals converge on regulation of Ler activity[188]. A variety of bacterial- and host-derived compounds including autoinducers, bile salts and hormones can act as signals for activation and repression of the T3SS[189–192]. Microbiota-produced metabolites like fucose can induce *ler* gene expression, activating the T3SS[130]. Host hormones can also serve as activation cues. Epinephrine and norepinephrine, recognized by the EHEC adrenergic two-component system QseBC, activate the T3SS[193–196]. QseBC can also activate the T3SS through detection of the EHEC secreted autoinducer AI-3, a marker of high bacterial density[197,198]. Other host hormones, like serotonin and endocannabinoids, can inhibit T3SS[199,200]. Targeting these molecules could provide alternative strategies to disrupt EHEC colonization capacity *in vivo*[201].

Shiga toxin is a ribotoxic protein released into the colonic lumen

During colonic colonization, EHEC releases the exotoxin Shiga toxin (Stx), which is responsible for the most severe manifestations of EHEC disease. Stx is exquisitely potent, with an intravenous median lethal dose of 2 ng/kg, it is one of the most lethal toxins known to humankind, second only to botulinum neurotoxin[202].

Shiga toxin (Stx) was first recognized as a dangerous toxin 100 years before EHEC was discovered. In 1898, the Japanese microbiologist Dr. Kiyoshi Shiga identified the etiological agent of epidemic dysentery, *Shigella dysenteriae*[203,204]. Shiga and others showed experimentally that *S. dysenteriae* produced a toxic factor[205,206], but lack of purification techniques made separation of its effects from bacterial endotoxin difficult[207]. New biochemical techniques in the early 1950s allowed purification of Stx from bacterial cells without significant contaminating LPS[208], allowing for the first time studies of Stx's toxic effects. Administration of Stx to animals and cultured cells provided evidence that Stx exposure led to neurotoxicity in the central nervous system[209,210], severe intestinal disease[211,212], and epithelial cell damage[213]. Later, administration of strains of *Shigella dysenteriae* with intact or disrupted Stx genes to human volunteers and non-human primates provided evidence that Stx was important for severe intestinal disease (dysentery) and inflammation[214–217].

It was discovered in 1977 that some strains of pathogenic *E. coli* produced a potent toxin that had a strong cytotoxic effect on Vero kidney epithelial cells[218]. 'Verotoxin' was recognized to be Shiga toxin after

the EHEC outbreak in 1982, when molecular biologists confirmed that Verotoxin produced by EHEC O157:H7 was nearly identical to Shiga toxin produced by *S. dysenteriae*[219–222]. It is now appreciated that the Shiga toxins produced by EHEC and *S. dysenteriae* are members of a large family of Shiga toxin proteins[223]. There are two major groups of Shiga toxin proteins which are antigenically distinct, Stx1 and Stx2. Stx2 is more potent and associated with severe disease[224,225]. EHEC isolates can harbor several distinct variants of Shiga toxins from either group[226], and Shiga toxin produced by EHEC is directly linked to the ability of this pathogen to cause extraintestinal disease like HUS[66,227,228].

Genes encoding Stx in the EHEC genome are found on a lambdoid prophage, pointing to its horizonal acquisition, potentially from *S. dysenteriae*[93–96]. The Stx genes are controlled by a phage promoter, and prophage induction is required for expression[229,230]. Typically, phage genes and Stx are quiescent, but in response to stress stimuli like DNA damage, phage promoters are activated and *stx* genes transcribed. Lysis of the bacterial cell by phage ultimately releases the toxin[229,230]. It is not immediately clear what benefit this provides to EHEC. This behavior could be a type of 'bacterial altruism,' where the sacrifice of some percentage of the population provides a benefit, such as nutrients, to the remaining intact cells[231]. Stx increases frequency and severity of diarrhea in infant rabbits inoculated with EHEC[21], which may facilitate pathogen transmission and thus be evolutionarily beneficial.

Shiga toxins are AB_5 subunit toxins. The A subunit is composed of two domains linked by a disulfide bond; the catalytic A1 domain, and the A2 domain are tethered noncovalently to the pentameric B subunit[232]. The B subunit binds to the host cell surface glycolipid globotriaosylceramide (Gb3)[233–243]. After binding, Stx is transported to the endoplasmic reticulum (ER) by retrograde transport (Figure 1.4)[244,245]. In the ER, the A and B subunits dissociate, are unfolded by chaperones, and the A1-A2 disulfide bond is cleaved[226,246]. The A1 subunit mimics an unfolded protein and is perceived as a substrate for ER-associated protein degradation (ERAD)[246]. A1 is translocated to the cytosol for degradation by ERAD machinery, but instead subverts protease interaction and refolds[226,246]. In the cytosol, the A1 subunit acts as an RNA N-glycosidase, catalyzing the removal of a specific adenine residue from the 28S rRNA, inactivating protein synthesis[247,248].

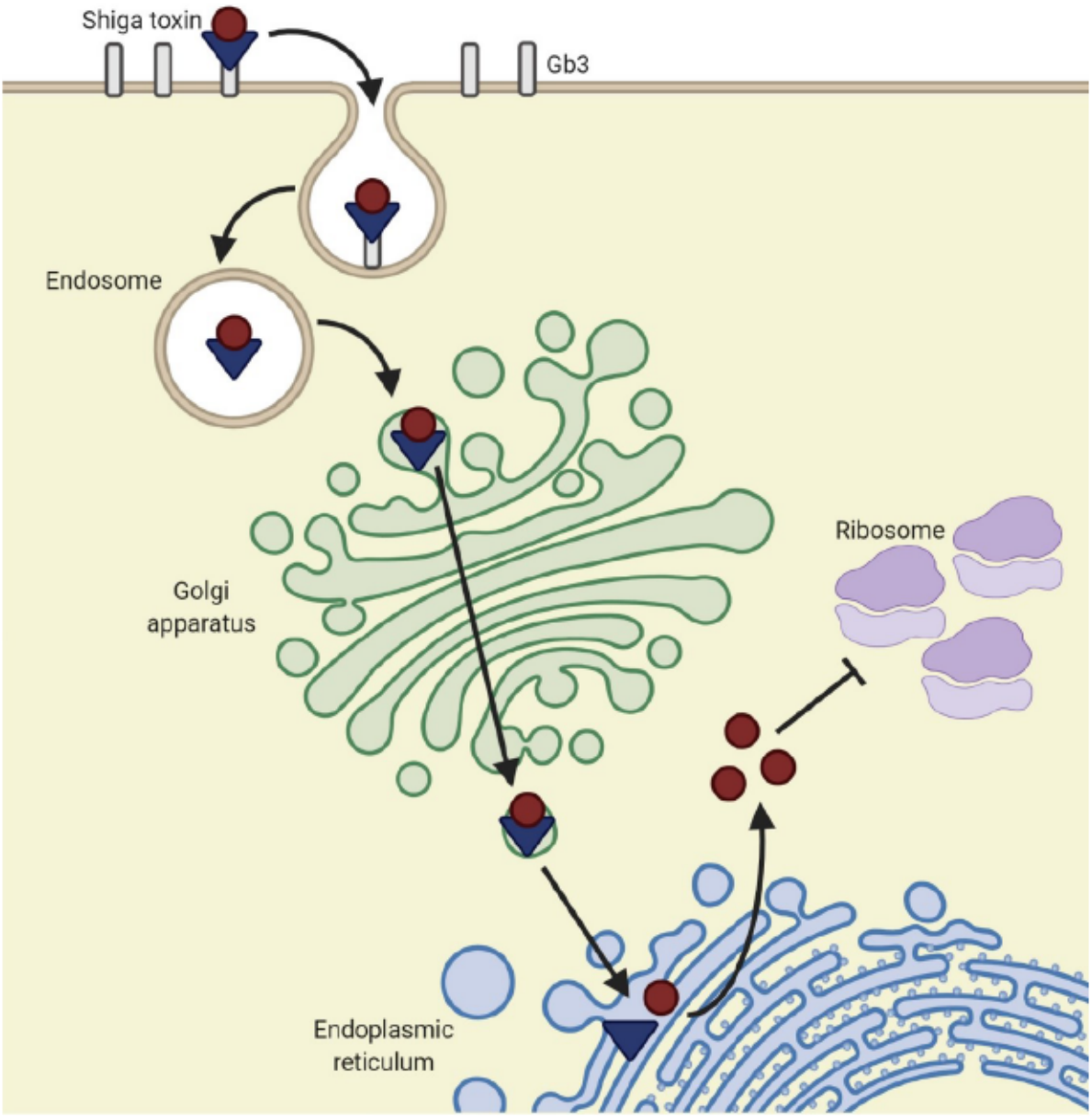


Figure 1.4: Retrograde transport of Shiga toxin.

Shiga toxin B subunit (blue triangle) binds to host surface molecule Gb3, triggering its endocytosis. Shiga toxin travels by retrograde transport through the Golgi apparatus to the endoplasmic reticulum (ER). In the ER, the Shiga toxin B subunit is separated from the toxic A subunit (red circle). The A subunit exists the ER into the host cytoplasm, where it depurinates a specific adenine residue in the 28S ribosomal RNA, inactivating the ribosome. Created with BioRender.

Depurination of the 28s rRNA triggers the 'ribotoxic stress response'[249–252]. Structural changes to the ribosomal peptidyltransferase domain rapidly trigger a protein kinase cascade, activating mitogen-activated protein kinases (MAPK) and stress-activated protein kinases (SAPK/JNK). These kinases activate transcription factors like FOS, JUN, NFKB, ATF2 and p38, which mediate changes in gene expression. mRNA stabilization may also contribute to this phenomenon of robust gene expression of certain transcripts known as 'super-induction[249,253]. In vitro, super-induced genes in intoxicated epithelial and immune cells typically include inflammatory cytokines[249,250,254–263]. Protein biosynthesis in intoxicated cultured cells begins to decline as rapidly as 30 minutes post depurination of the ribosomal rRNA in highly sensitive cells, and several hours in less sensitive cell types[249]. How the opposing effects of increased transcriptional and decreased translational activity manifest at the protein-level depends on the cell type and is not completely understood[249]. Apoptosis, indicated by blebbing nuclei and DNA degradation, begins at least 6 hours post intoxication[264], with loss of membrane integrity and viability after 48 hours[265].

Of note, although Stx has been associated with pathology in the colon[21,23,254,266,267], it is unclear mechanistically how Stx interacts with colonic epithelial cells because the presence of the Gb3 receptor in the colon has been disputed[268–270]. It has been suggested that Stx could intoxicate epithelial cells through interacting with other receptors, like Gb4[270,271]. Another possibility is that Stx does not intoxicate epithelial cells at all, but rather transverses through the epithelial layer by other mechanisms, such as interaction with antigen-sampling Microfold cells, to interact with endothelial and immune cells in the underlying tissue[272–278]. Chapter 4 discusses in detail the effect of Shiga toxin on the pathology and transcriptome of epithelial and immune cells in the colon in a physiologically relevant animal model[26].

Specific mucosal immune response to EHEC and Stx

EHEC and Stx interact closely with epithelial and immune cells in the colon, and these cell types work together to respond to and ultimately clear the infection. Epithelial cells are the first to interact with invading pathogens and have an important task to 'sound the alarm' to the immune system that a threat is present. Epithelial cells do this by recognizing compounds and structures produced by microbes, known as microbe

associated molecular patterns (MAMPs), such as LPS, peptidoglycan (PG) and flagellin. MAMPs are recognized by cell surface receptors known as pattern-recognition receptors (PRRs) which upon binding, trigger transcriptional changes such as expression of cytokines[133]. Toll-like-receptors (TLRs) are one major class of PRRs: TLR4 and TLR5 are key for recognizing EHEC-derived MAMPs LPS and flagellin[279]. In response to EHEC infection, epithelial cells secrete proinflammatory cytokines and chemokines such as CXCL8, TNFα, IL-6, CCL2 and CCL5[252,256,257,262,263,276,280–287]. These cytokines markedly remodel immune cell behavior[160,288].

EHEC also causes significant tissue destruction, leading to breaches in the epithelial barrier. Colonoscopy and histologic evaluation of EHEC-infected human patient biopsies reveal gross hemorrhage, edema, inflammation and focal necrosis[266,289–291]. Epithelial tissue destruction leads to release of damage-associated molecular patterns (DAMPs), molecules which signal to the immune system that cellular integrity has been compromised[292]. These molecules, such as mitochondrial DNA and cytosolic proteins, are typically sequestered inside cells, but can activate immune cells when released after cellular damage[292]. EHEC can also directly interact with immune cells in the LP after tissue damage. Underlying immune cells can directly recognize EHEC MAMPs through processes similar to those in epithelial cells, leading to the release of additional inflammatory cytokines, such as CXCL1, IL1α/β, and G-CSF[160,255,258,259,261,293–299]. Importantly, chemokine production draws migrating leukocytes from the blood to the site of infection. The first to arrive are phagocytic neutrophils and later monocytes, which rapidly differentiate into macrophages. Massive neutrophil infiltration is a hallmark of EHEC infection and has been observed in many clinical specimens[266,289–291]. Neutrophils and monocytes may serve as cellular 'carriers' to transport Stx from the colon to the blood and distal sites[276,300–303]. Elevated white blood cell count predicts a poor prognosis and higher likelihood of progression to HUS[304,305]. Dendritic cells, activated in the infected tissue after exposure to MAMPs or DAMPs, travel to colonic patches of gut-associated lymphoid tissue. Here, they present antigen to mature naïve lymphocytes, leading to the activation of an adaptive immune response. Activated B cells differentiate into plasma cells and produce antibodies against specific bacterial antigens. EHEC infection

triggers antibody generation against the O-antigen of lipopolysaccharide, T3SS effector proteins, and more unusually, Stx[306–308].

EHEC proteins can modulate the host immune response in epithelial and immune cells. Several T3SS effectors function to suppress cellular immune pathways[172]. Common effector mechanisms include the inhibition of phagocytosis or apoptosis[172], as well as the suppression of NFKB signaling, which is necessary to alter gene expression in response to pathogen detection[279]. The existence of an immune response to Stx in the colon has been controversial, with claims that flagella and LPS are the main stimulatory factors[299]. More recently, it has become clear that Stx can shape the immune response to infection, although the extent of that remodeling in the prodromal phase of disease in the GI tract is poorly characterized[299]. The interaction between Stx and the colonic mucosa is discussed in detail in Chapter 4, where we show that Stx powerfully shapes the host transcriptional response to infection[26].

Pathophysiology of Hemolytic Uremic Syndrome

After passage through the colonic mucosa, Stx enters the bloodstream. Initially, Stx is detected in patient blood bound to blood cells, including leukocytes[274,309], red blood cells[310,311] and platelets[312,313]. An increase in the concentration of unbound Stx in blood over the course of infection is associated with progression to HUS[314]. Stx-laden blood enters the kidney via the afferent arteriole, where it passes through one of many glomeruli, the microvasculature capillary units of the kidney (Figure 1.5). The endothelial cells of the glomerulus are adjacent to podocyte cells of the Bowman's capsule, which collects filtrate from the blood. This filtrate transits through vessels made of tubule cells and eventually becomes urine. Three cell types in this region, the endothelial cells of the microvasculature, the absorptive podocytes, and tubule cells, all express high levels of Gb3, making them exquisitely sensitive to Stx[315–319]. Glomerular endothelial cells are considered the primary target of Stx[299,316,320]. LPS and cytokines in blood during infection can promote even higher expression of Gb3 on endothelial cells, sensitizing them further to Stx[294,321–323].

Intoxication of endothelial cells results in cell death and detachment from the glomerular basement membrane[299,317]. Destruction of the endothelial barrier allows Stx to move into the Bowman's capsular space,

facilitating interaction with podocytes and tubule cells and generating more inflammation, recruiting immune cells[317]. Exposure to cytokines and Stx damage converts the endothelial layer from a normal 'thromboresistant' phenotype to one that actives platelets, leading to increased fibrin deposition and thrombi formation[317,320]. Growing thrombi sequester platelets and occlude capillaries, impeding blood flow and leading to fragmentation of red blood cells from shear stress[324]. This leads to the clinical manifestation of thrombocytopenia and anemia. Eventually, this immunopathology can lead to acute kidney failure[316,325]. Comparatively less is known about the coagulation processes in the GI mucosa; this is discussed in detail in Chapter 4, where we show that Stx can induce expression of coagulation genes in colonic epithelial cells.

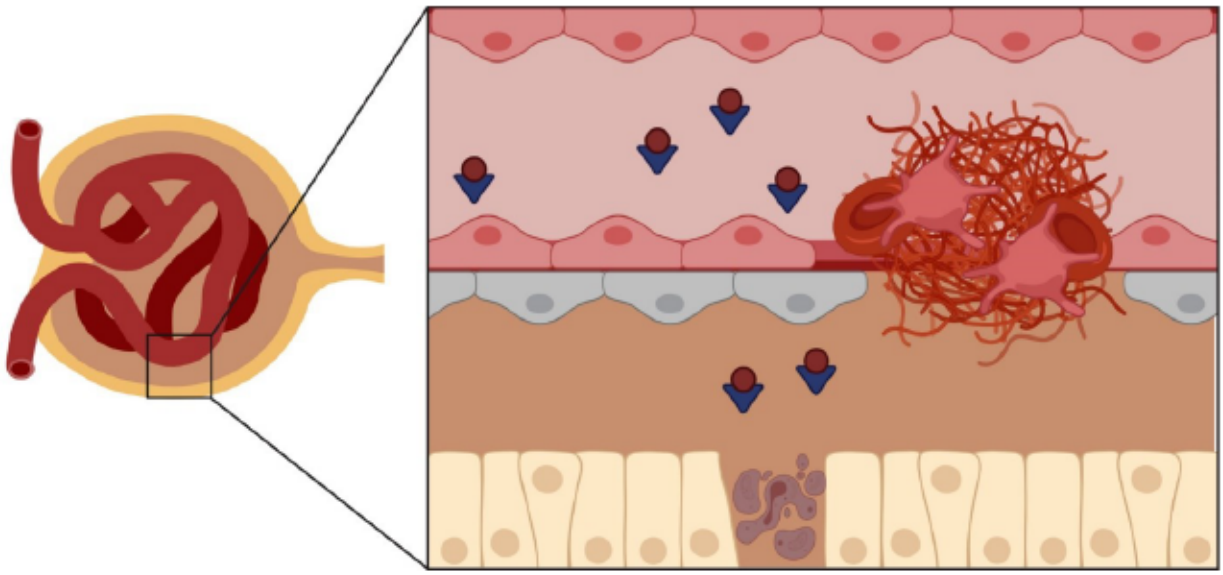


Figure 1.5: Shiga toxin at the glomerulus.

Shiga toxin (red circle-blue triangle) kills endothelial cells of the glomeruli microvasculature and triggers thrombus formation. Shiga toxin can also travel into the Bowman's space, damaging podocytes and proximal tubule cells.

Therapeutics for EHEC infection and HUS

Despite decades of work, there are no available therapeutic options to treat or prevent HUS after EHEC infection[326,327]. The current standard-of-care is to treat patients with IV fluids to prevent dehydration and reduce severity of kidney injury[328]. Although available, antibiotics are not typically used to treat EHEC infection, as they increase the risk for progression to HUS[12,13].

Several therapeutic options have proven to be ineffective in preclinical animal models or human trials. Gb3 analogs were developed that could bind to the Stx B subunit and effectively neutralize the toxin in vitro[326,327]. Phase II trials showed no effect in HUS patients[329]. Other strategies aimed to block retrograde transport of Stx A subunit into cells after binding, but these drugs had no or limited effect in preclinical models[330–333].

Toxin-neutralizing therapies have been successful for diseases caused by other toxins, like diphtheria and tetanus[334]. Administration of anti-Stx antibodies showed promise in preclinical animal studies and were deemed safe after Phase I and Phase II clinical trials[326,335,336]. One major challenge of antibody-neutralization therapy is that there are antigenically distinct subtypes of Stx; an antibody or antibody cocktail would need to neutralize multiple toxin subtypes in order to be widely useful[326]. Additionally, a polysaccharide conjugate vaccine was tested in adults and children and elicited an anti-LPS response[337–339], but has not been developed further. Improvements could be made to include the Stx B subunit or antigenic T3SS effector proteins[337,340,341].

It is difficult to progress past Phase II clinical trials for therapeutics treating HUS. Because of the sporadic nature of EHEC outbreaks and relatively rare progression to HUS, Hall *et al.* describe the "statistics and finances for a clinical trial [as] almost insurmountable"[326]. Targeting cellular processes downstream of Stx intoxication that are shared with other disease mechanisms may be more fruitful for designing a successful large-scale trial. These processes include the unfolded protein response, apoptosis, the ribotoxic stress response, and toxin-induced inflammatory processes[326]. For example, anti-cytokine therapies and kinase inhibitors could block secretion of inflammatory cytokines which drive pathology in the colon and kidney[326]. More work is necessary to determine how inflammatory responses in the colon contribute to the progression of HUS; this is discussed in Chapter 4.

2

Transposon-insertion sequencing screens unveil requirements for EHEC growth and intestinal colonization

Overview

Enterohemorrhagic *Escherichia coli* O157:H7 (EHEC) is an important food-borne pathogen that colonizes the colon. Transposon-insertion sequencing (TIS) was used to identify genes required for EHEC and *E. coli* K-12 growth *in vitro* and for EHEC growth *in vivo* in the infant rabbit colon. There was a restrictive bottleneck for EHEC colonization of the rabbit colon, which complicated identification of EHEC genes facilitating growth *in vivo*. Two analysis methods were used compensate for the effects of the infection bottleneck. These analyses confirmed that the EHEC LEE-encoded type 3 secretion apparatus is required for growth *in vivo* and revealed that only a few effectors are critical for *in vivo* fitness. Over 200 mutants not previously associated with EHEC growth *in vivo* also appeared attenuated *in vivo*, and a subset of these putative *in vivo* fitness factors were validated. Some were found to contribute to efficient type-three secretion while others promote EHEC resistance to host-derived stresses. We identified *cpxA* as required for EHEC, *Vibrio cholerae* and *Vibrio parahaemolyticus* resistance to the bile salt deoxycholate, highlighting the important role of this previously uncharacterized protein in pathogen survival. Collectively, our findings provide a comprehensive framework for understanding EHEC growth in the intestine.

Parts of this work were first published as: Warr AR, Hubbard TP, Munera D, Blondel CJ, Abel zur Wiesch P, Abel S, Wang X, Davis BM, Waldor MK. 2019. Transposon-insertion sequencing screens unveil requirements for EHEC growth and intestinal colonization. Plos Pathogens. 15(8):e1007652

Author contributions: ARW contributed equally to this work with TPH, DM, and CJB. The work was conceptualized by ARW, TPH, DM, CJB, and MKW. The investigation and data analysis were performed by ARW, TPH, DM, CJB, PAzW, SA, XW with supervision from MKW. The figures were made by ARW and TPH. The writing was completed by ARW, TPH, BM and MKW.

Introduction

Enterohemorrhagic Escherichia coli (EHEC) is an important food-borne pathogen that causes gastrointestinal (GI) infections worldwide. EHEC is a non-invasive pathogen that colonizes the human colon and gives rise to sporadic infections as well as large outbreaks[342–344]. The clinical consequences of EHEC infection range from mild diarrhea to hemorrhagic colitis and include the potentially lethal hemolytic uremic syndrome (HUS)[66,345].

The prototypical EHEC O157:H7 strain, EDL933, caused the first recognized EHEC outbreak in 1982[67]. EDL933 has a 5.5 Mb chromosome and a 90 kb virulence plasmid[63,346]. The *E. coli* species pan-genome is large (>16,000 genes), and any given isolate contains a diverse, mosaic genome with approximately 1,500–2,000 conserved 'core' genes[53,54,56,57,347,348] and an additional 3,000–4,000 'accessory' genes. The pathotype *E. coli* O157:H7 specifically contains one or more prophages encoding Shiga toxins and the Locus of Enterocyte Effacement (LEE) pathogenicity island[61]. These two horizontally acquired elements are critical EHEC virulence determinants. Shiga toxins contribute to diarrhea and the development of HUS[21,66,345]. The LEE encodes a type III secretion system (T3SS) and several secreted effectors. EHEC's T3SS mediates attachment of the pathogen to colonic enterocytes, effacement of the brush border microvilli, and the formation of actin-rich pedestal-like structures underneath attached bacteria[184]. Once translocated into the host cell, T3SS effectors, which are encoded both inside and outside the LEE, target diverse signaling pathways and cellular processes[171,185]. A functional LEE T3SS is required for EHEC intestinal colonization in animal models as well as in humans[21,171,177,180,182,186,349].

In addition to the virulence factors that prompt the key symptoms of infection, EHEC also relies on bacterial factors that enable pathogen survival in and adaptation to the host environment. During colonization of the human GI tract, EHEC encounters multiple host barriers to infection, including but not limited to stomach acid, bile, and other host- and microbiota-derived compounds

with antimicrobial properties[189]. EHEC is known to detect intestinal cues derived from the host and the microbiota to activate expression of virulence genes and to modulate gene expression both temporally and spatially[124,194,197,350,351]. However, a comprehensive, genome-wide analysis of bacterial factors that contribute to EHEC survival within the host has not been reported.

The development of transposon-insertion sequencing (TIS, also known as TnSeq, InSeq, TraDIS, or HITS)[352–355] facilitated high-throughput and genome-scale analyses of the genetic requirements for bacterial growth in different conditions, including in animal models of infection[356–365]. In this approach, the relative abundance of transposon-insertion mutants within transposon-insertion libraries provides insight into loci's contributions to bacterial fitness in different environments[366,367]. Potential insertion sites for which corresponding insertion mutants are not recovered frequently correspond to regions of the genome that are required for bacterial growth (often termed 'essential genes'), although the absence of a particular insertion mutant does not always reflect a critical role for the targeted locus in maintaining bacterial growth[368,369]. Comparative analyses of the abundance of mutants in an initial (input) library and after growth in a selective environment (e.g., an animal host) can be used to gauge loci's contributions to fitness in the selective condition.

Here, transposon libraries were created in EHEC EDL933 and the laboratory-adapted *E. coli* K-12 and used to characterize their respective *in vitro* growth requirements. The EHEC library was also passaged through an infant rabbit model to identify genes required for intestinal colonization. Our data indicate that during infection of the GI tract, EHEC populations undergo a severe infection bottleneck that complicates identification of genes with true *in vivo* fitness defects. We used two complementary analytic approaches to mitigate the noise introduced by restrictive bottlenecks to identify over 200 genes required for efficient colonization of the rabbit colon. As expected, these included the LEE-encoded T3SS and *tir*, a LEE-encoded effector necessary for intestinal

colonization[21,22]. In addition, 2 non-LEE effectors and many additional new genes that encode components of the bacterium's metabolic pathways and stress response systems were found to enable bacterial colonization of the colon. Isogenic mutants for 17 loci were constructed, validated in the infant rabbit model and tested *in vitro* under stress conditions that model host-derived challenges encountered within the GI tract. *cpxA* was found to be specifically required for resistance to the bile salt deoxycholate and therefore appears to be a previously unappreciated member of the bile-resistance repertoire of diverse enteric pathogens.

Identification of genes required for EHEC growth *in vitro*

The mariner-based Himar1 transposon, which inserts specifically at the TA dinucleotide[370] was used to generate a transposon-insertion library in EDL933. The library was characterized via high-throughput sequencing of genomic DNA flanking sites of transposon-insertion. To map the reads, we used the most recent EDL933 genome sequence[346] and annotation (NCBI, February 2017). Since this genome, unlike the initial EHEC genome[63], has not been linked to functional information (e.g., the EHEC KEGG database[371]), we generated a correspondence table in which the new gene annotations (RS locus tags) are linked to the original annotations (Z numbers) (Table S1). This correspondence table enabled us to utilize historically valuable resources as well as the updated genomic sequence. 137,805 distinct insertion mutants were identified, which corresponds to 52.5% of potential insertion mutants with an average of ~21 reads per genotype (Figure 2.1A). Sensitivity analysis revealed that nearly all mutants were represented within randomly selected read pools containing ~2 million reads. Increasing sequencing depth to ~3 million reads had a negligible effect on library complexity, suggesting that a sequencing depth of ~3 million reads is sufficient to identify virtually all genotypes within this EHEC library (Figure 2.1A).

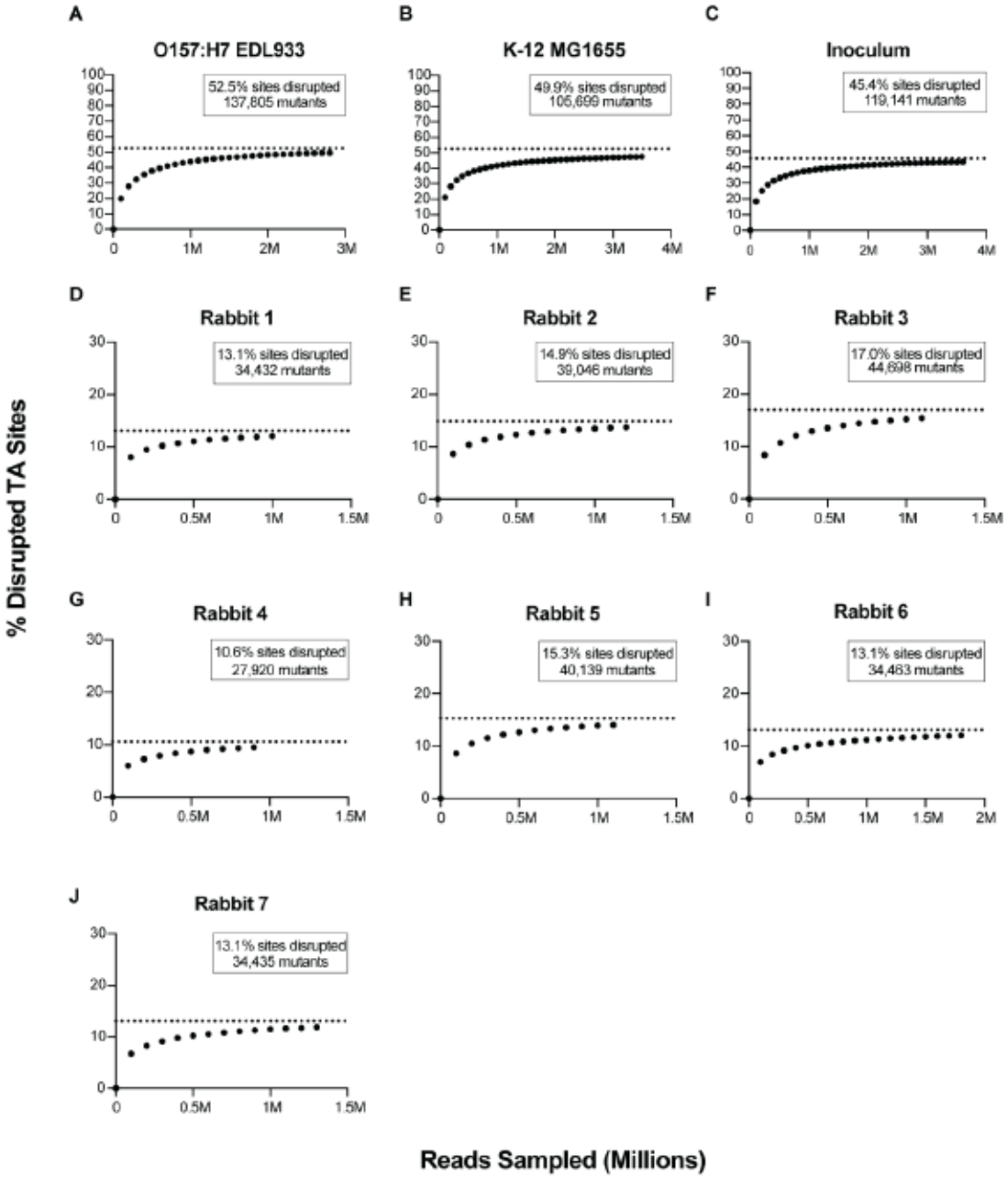


Figure 2.1: Sequencing saturation of TIS libraries.

Reads were randomly sampled from each library and the percentage of TA sites disrupted in each randomly selected pool were plotted for the EDL933 library (A), MG1655 library (B), the inoculum library used to infect infant rabbits (C), and the libraries recovered from 7 rabbit colons (D-J).

EHEC's 6,032 annotated genes were binned according to the percentage of disrupted TA sites within each gene (Figure 2.2A). As expected for a high-density Himar1 transposon-insertion library, which contains insertions at a majority of TA sites, this distribution was bimodal, with a minor peak comprised of genes disrupted in few potential insertion sites (Figure 2.2A, left), and a major peak comprised of genes that are disrupted in most or all potential insertion sites (Figure 2.2A, right)[367]. Based on the center of the major right-side peak, we estimate that ~70% of non-essential insertion sites have been disrupted in this EHEC library, a degree of complexity that enabled high-resolution analysis of transposon-insertion frequency.

Further analysis of insertion site distribution was performed using a hidden Markov model-based analysis pipeline (EL-ARTIST[366]) that classifies loci with a low frequency of transposon-insertion across the entire coding sequence as 'underrepresented' (often referred to as 'essential' genes) or across a portion of the coding sequence as 'regional' (Figure 2.2B). All other loci are classified as 'neutral'. Of EHEC's 6,032 genes, 895 were underrepresented (red), 407 regional (purple), and 4,730 neutral (blue) (Figure 2.2AB, Table S2). Neutral genes are likely dispensable for growth in LB, whereas non-neutral genes (regional and underrepresented genes combined) likely have important functions for growth in this media or are otherwise refractory to transposon-insertion[368,369].

We identified Z Numbers (Table S1) and the linked Clusters of Orthologous Groups (COG)[372,373] and KEGG pathways associated with the 1302 genes classified as non-neutral (underrepresented and regional) (Tables S2 and S3). Each COG category was plotted against its 'COG Enrichment Index', which is calculated as the percentage of non-neutral genes in each COG category divided by the percent of the whole genome with that COG[374]. A subset of COGs,

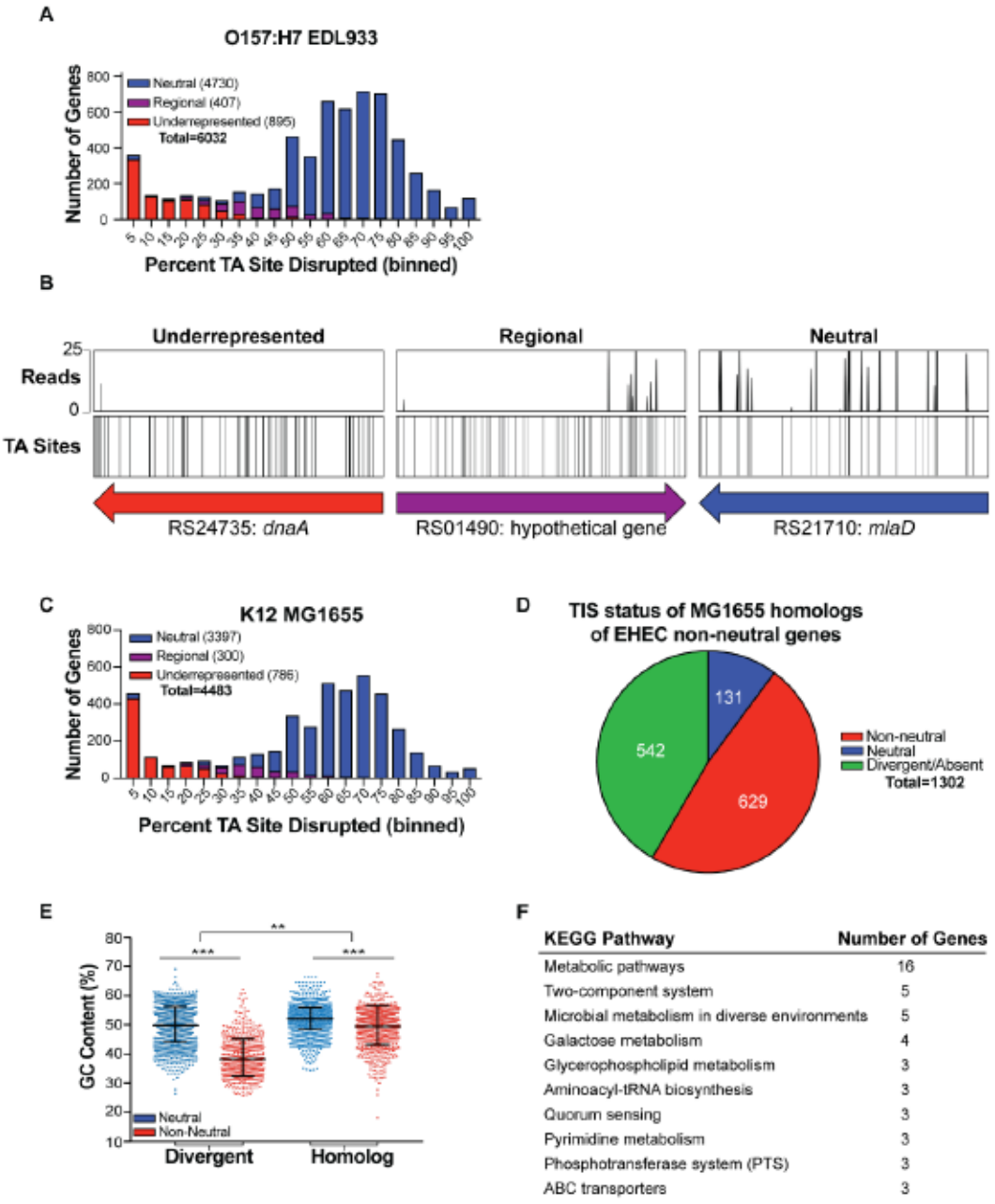


Figure 2.2: Analysis of essential genes in EHEC EDL933 and comparison to K-12 MG1655.

(A) Distribution of percentage TA site disruption for all genes in EHEC EDL933. Genes are classified by EL-ARTIST as either underrepresented (red), regional (purple), or neutral (blue). (B) Transposon-insertion profiles of representative underrepresented, regional, and neutral genes. (C) Distribution of percentage TA site disruption for all genes in K-12 MG1655. Genes are classified by EL-ARTIST as underrepresented (red), regional (purple), or neutral (blue) using the same parameters as for the EDL933 library. (D) Genes classified as non-neutral in the ELD933 TIS library were compared to the MG1655 TIS library and categorized as either lacking a homolog (green), having the same classification in both libraries (red), or being non-neutral in EDL933 and neutral in MG1655 (blue). (E) GC content (%) of EDL933 genes with and without homologs in K-12 MG1655, classified by TIS classification (neutral or non-neutral). Neutral and non-neutral genes within each gene type (divergent or homolog) are compared using a Mann-Whitney U test with a Bonferroni correction. Ratios of neutral to non-neutral genes for each gene type are compared using a Fisher's Exact Test; (**) indicates a p-value of <0.01 and (***) indicate a p-value of <0.001. (F) Top 10 KEGG pathways associated with genes that are non-neutral in EDL933 and neutral in MG1655.

particularly translation, lipid and coenzyme metabolism, and cell wall biogenesis were associated with non-neutral genes at a frequency significantly higher than expected based on their genomic representation (Figure 2.3A). Collectively, the COG and a similar KEGG analysis (Table S3) revealed that EHEC genes with non-neutral transposon-insertion profiles are associated with pathways and processes often linked to essential genes in other organisms[375].

Evaluating the abundance of non-neutral genes in EDL933

Non-neutral genes comprise ~22% of EHEC's annotated genes, a proportion of the genome that is substantially larger than the 8% and 9% observed in analogous TIS-based characterizations of *Vibrio parahaemolyticus* and *Vibrio cholerae*[366,376]. To evaluate whether the abundance of non-neutral loci was specific to EHEC or was characteristic of additional *E. coli* strains, a high-density transposon-insertion library was constructed in *E. coli* K-12 MG1655[377]. EL-ARTIST analysis of the K-12 library was implemented with the same parameters as those for the EHEC library and classified 24% of genes as underrepresented (786 underrepresented, 300 regional and 3397 neutral; Figure 2.2C, Table S4).

We compared gene classifications between homologous loci (Table S2) and found a substantial concordance between the sets of genes with non-neutral insertion profiles: 83% (629/760) of the non-neutral EHEC genes with homologs in K-12 were likewise classified as non-neutral in K-12 (Figure 2.2D). Thus, analyses of non-neutral loci suggest either that most ancestral loci make similar contributions to the survival and/or proliferation of EHEC and K-12 in LB or that they are similarly resistant to transposon-insertion.

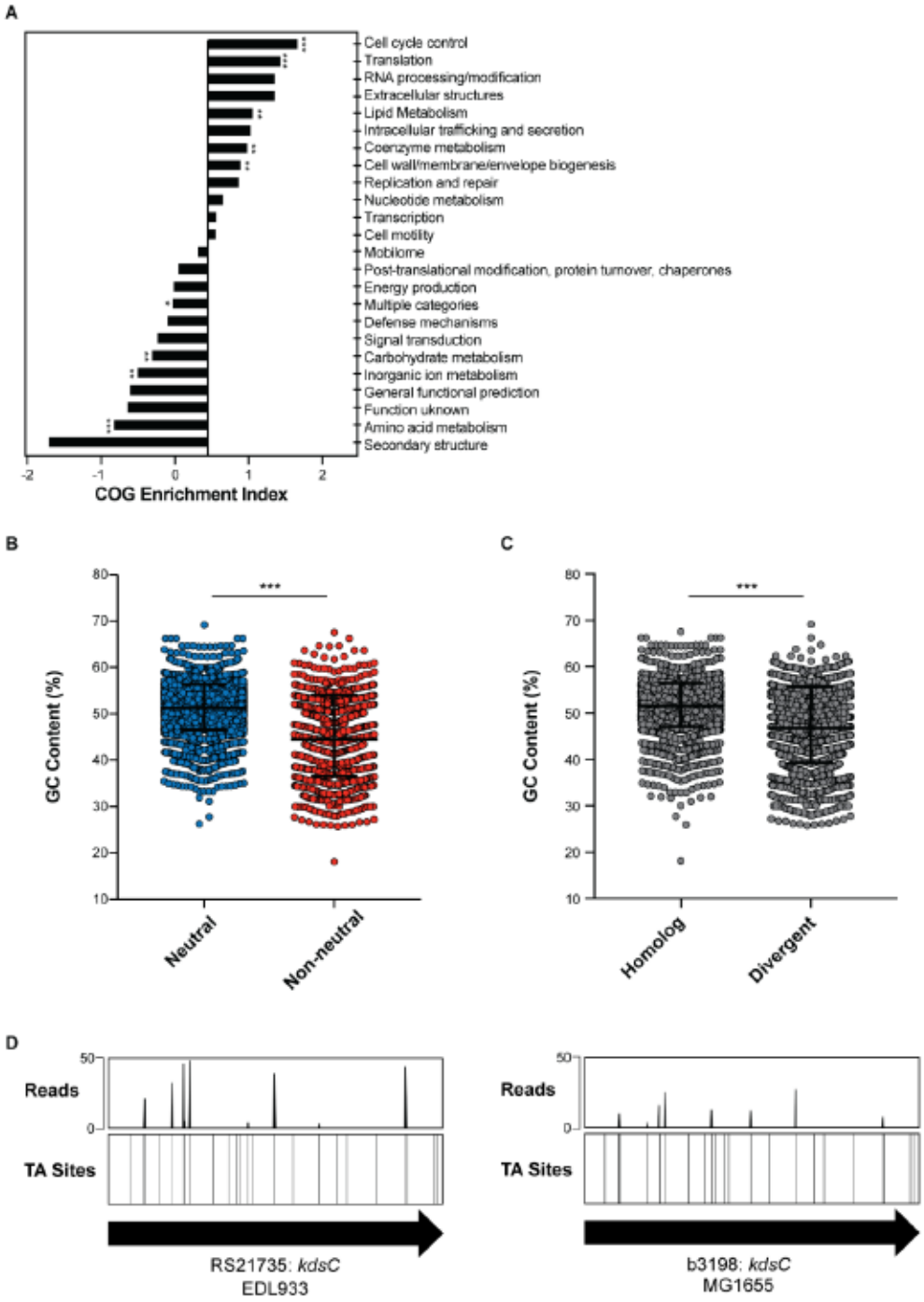


Figure 2.3: Assessment of non-neutral EHEC EDL933 genes.

(A) Non-neutral genes by Clusters of Orthologous Groups (COG) classification. A two-tailed Fisher's exact test with a Bonferroni correction was used to test the null hypothesis that enrichment is independent of TIS status. p-values considered to be significant if <0.002. Single asterisks (*) indicates p-value <0.002, double asterisks (**) indicate p-value <0.001, and triple asterisks (***) indicate p-value <0.0001. (B) GC content (%) of EDL933 genes classified as either neutral (blue) or non-neutral (regional + underrepresented; red) by TIS. Distributions are compared using a Mann-Whitney U test; triple asterisks (***) indicate p-value of <0.0001. (C) GC content (%) of EDL933 genes classified as having a homolog in MG1655 (homolog) or not (divergent). Distributions are compared using a Mann-Whitney U test; triple asterisks (***) indicate p-value of <0.0001. (D) TA insertions across *kdsC* in EDL933 (left) and MG1655 (right).

Previous analyses revealed that nucleoid binding proteins such as HNS, which binds to DNA with low GC content, can hinder Himar1 insertion[368]. Consistent with this observation, EHEC genes classified as non-neutral have a lower average GC content than genes classified as neutral (Figure 2.3B). Interestingly, the disparity in GC content between neutral and non-neutral loci is particularly marked for EHEC genes that do not have a homolog in K-12 (divergent; Figure 2.2E). These analyses suggest that there is an association between GC content and transposon-insertion frequency in EHEC, as in other organisms, and that the prevalence of underrepresented loci among divergent loci may in part stem from the lower average GC content of these loci (Figure 2.3C). Additional studies are necessary to determine if the association between low GC content and reduced transposon-insertion is due to the binding of HNS or other nucleoid-associated proteins, or yet unidentified fitness-independent transposon-insertion biases.

Comparison of TIS and deletion-based gene classification

The sets of genes classified as underrepresented or regional in EHEC and K-12 transposon libraries were compared to the 300 genes classified as essential in the K-12 strain BW25113 based on their absence from a comprehensive library of single gene knockouts[378–380]. 98% of these genes (294/300) were also classified as underrepresented or regional in EDL933 (Table S2) and MG1655 (Table S4). The few loci previously classified as essential but not found to be underrepresented or regional in our analysis include several small genes, whose low number of TA sites hampers confident classification. One gene in this list, *kdsC*, was found to have insertions across the gene in both EDL933 and MG1655 (Figure 2.3D). *kdsC* knockouts have also been reported previously[381], confirming that this locus is not required for K-12 growth despite the absence of an associated mutant within the Keio collection. Thus, underrepresented and regional loci encompass, but are not limited to, loci previously classified as essential.

Several factors likely account for the frequent classification of 'non-essential' loci as underrepresented or regional. First, loci can be classified as underrepresented even when viable mutants are clearly present within the insertion library (Figure 2.2A); insertions simply need to be consistently less abundant across a segment of the gene than insertions at other (neutral) sites. Loci may also be classified as underrepresented due to fitness-independent insertion biases, as discussed above[368,369]. Additional evidence that loci categorized as non-neutral by transposon-insertion studies are not necessarily essential for growth was provided by a recent study of essential genes in K-12[382]. However, the more expansive non-neutral classification can provide insight into loci that enable optimal growth, in addition to those that are required.

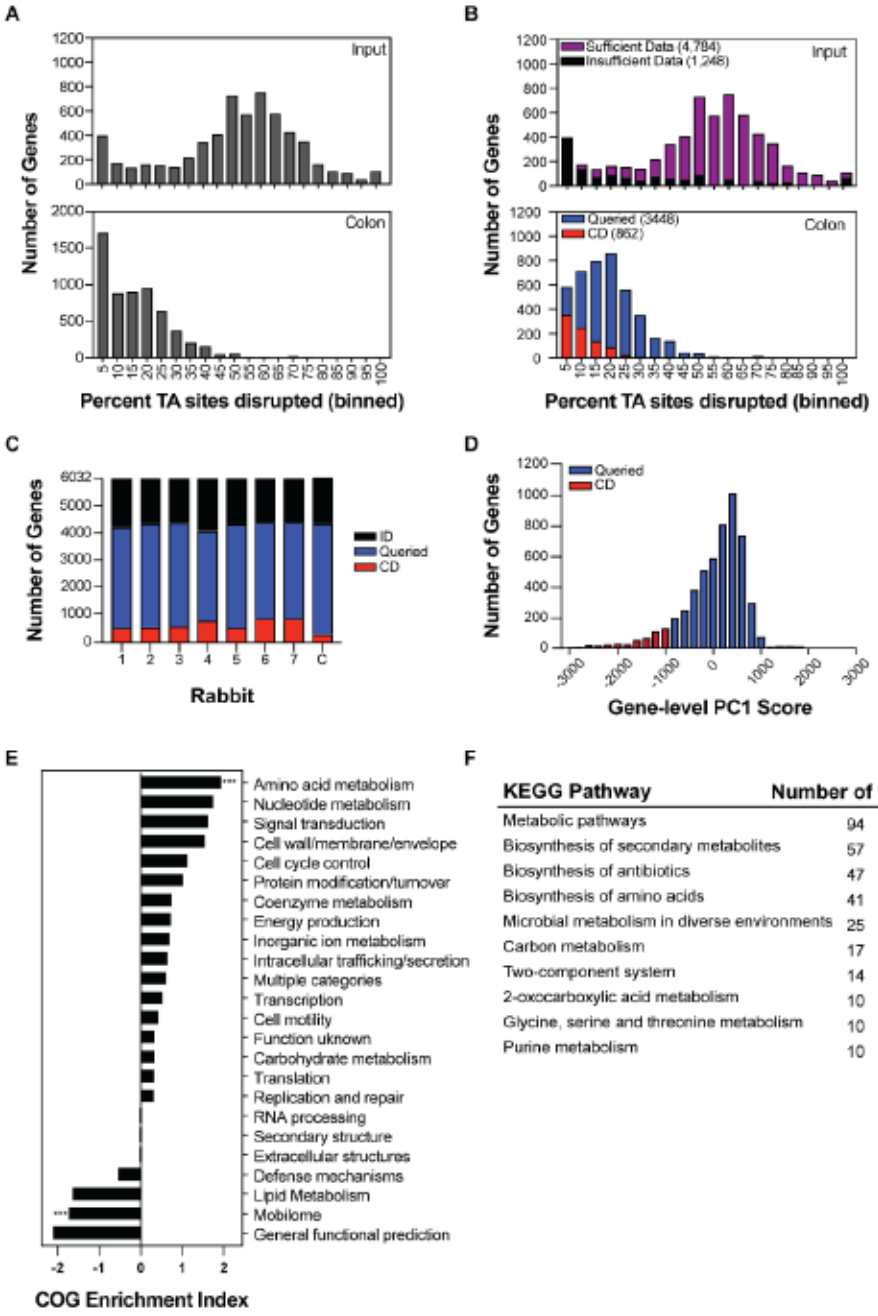
TIS-based comparison of EHEC and E. coli in vitro growth requirements

We further explored the 131 underrepresented EHEC loci (Table S5) that were classified as neutral (able to sustain insertions) in K-12. Most of these genes are linked to KEGG pathways for metabolism, particularly metabolism of galactose, glycerophospholipid, and biosynthesis of secondary metabolites (Table S5, Figure 2.2F). While this divergence could reflect the laboratory adaptation of the K-12 isolate, gene acquisition during EHEC evolution may have heightened the pathogen's reliance on metabolic processes that are not critical for growth of K-12. Such ancestral genes may be useful targets for antimicrobial agents, as they might antagonize EHEC growth without disruption of closely related commensal Enterobacteriaceae populations. TIS studies in additional *E. coli* isolates are required to determine the relative contributions of the core and accessory *E. coli* genome to the list of essential genes.

Identification of EHEC genes required for growth *in vivo*

To identify mutants deficient in their capacity to colonize the mammalian intestine, the EHEC transposon library was orogastrically inoculated into infant rabbits, an established model host for infection studies[21,22,383,384]. EHEC strain EDL933 causes diarrhea and similar pathology in infant rabbits as that previously described for EHEC strain 905[21]. Transposon-insertion mutants were recovered from the colon 2 days post-infection, and the sites and abundance of transposon-insertion mutations were determined via sequencing. The relative abundance of individual transposon-insertion mutants in the library inoculum was compared to samples independently recovered from the colons of 7 animals to identify insertion mutants that were consistently less abundant in libraries recovered from the colon. Under ideal conditions, this signature is indicative of negative selection of the mutant during infection, reflecting that the disrupted locus is necessary for optimal growth within the intestine.

Sequencing and sensitivity analyses of the 7 passaged libraries revealed that they contained substantially fewer unique insertion mutants than the library inoculum (23–38% total mutants recovered, ~30,000 of 120,000) (Figure 2.1D-J). These data are suggestive of population constrictions that could have arisen from 2 distinct but not mutually exclusive causes: 1) negative selection, leading to depletion of mutants deficient at *in vivo* survival or intestinal colonization; and/or 2) infection bottlenecks, population constrictions that lead to stochastic reductions in the average number of insertions per gene independent of genotype or selective pressures. We binned genes according to the percentage of TA sites disrupted within their sequences and plotted the number of genes corresponding to each bin for both the inoculum (Figure 2.4A-top) and a representative rabbit-passaged sample (Figure 2.4A-bottom). The passaged sample exhibited a marked leftward shift relative to the inoculum, a signature indicative of population constriction due to an infection bottleneck[367,385].



F

KEGG Pathway	Number of Genes
Metabolic pathways	94
Biosynthesis of secondary metabolites	57
Biosynthesis of antibiotics	47
Biosynthesis of amino acids	41
Microbial metabolism in diverse environments	25
Carbon metabolism	17
Two-component system	14
2-oxocarboxylic acid metabolism	10
Glycine, serine and threonine metabolism	10
Purine metabolism	10

Figure 2.4: Identification of EHEC genes required for intestinal colonization

(A) Distribution of percentage TA site disruption for all genes in EDL933 in the library used to inoculate infant rabbits (top) and in a representative library recovered from a rabbit colon two days after infection (bottom). (B) Distribution shown in (A) overlaid with the classifications from Con-ARTIST. Genes with insufficient data are removed from the bottom panel. (C) Distribution of Con-ARTIST gene classifications (insufficient data (ID, black), queried (blue), conditionally depleted (CD, red)) in each library recovered from seven infant rabbit colons two days post infection as compared to the inoculum. The last column, C, represents the consensus classification for all genes. (D) Distribution of PC1 scores across all EHEC genes. Red bins fall within the lowest 10% of PC1 scores (CD). Blue bins represent genes classified as queried (Q). (E) Conditionally depleted genes by Clusters of Orthologous Groups (COG) classification. A two-tailed Fisher's exact test with a Bonferroni correction was used to test the null hypothesis that enrichment is independent of TIS classification. (***), p-value <0.0001. (F) Top 10 KEGG pathways of EHEC genes classified as conditionally depleted by both Con-ARTIST and CompTIS (Group 1).

As infection bottlenecks can confound identification of genes with true fitness defects *in vivo,* we analyzed the TIS data using two complementary pipelines that mitigate the effects of bottlenecks: Con-ARTIST[366] and CompTIS[386]. Both Con-ARTIST and CompTIS use iterative simulation-based normalization to compensate for experimental bottlenecks and facilitate discrimination between stochastic reductions in genotype abundance and reductions attributable to genuine selection (mutants for which there was a fitness cost in the host environment); however, they use distinct methodologies to measure phenotypic consistency across animal replicates in order to categorize genes as either conditionally depleted (CD), queried (Q), or insufficient data (ID) as compared to the inoculum library (Figure 2.5).

Con-ARTIST utilizes relatively stringent standards to identify robust candidate genes that facilitate EHEC intestinal growth in each rabbit. Such 'conditionally depleted' (CD) genes must contain sufficient transposon-insertions for Con-ARTIST analysis (at least 5 TA sites disrupted by transposon-insertion within the inoculum and output datasets) and meet a standard of a 4-fold reduction in read abundance (fold-change) that is consistent across TA sites in a gene, where consistency is measured using a Mann Whitney U (MWU) statistical test (Figure 2.5). Queried genes contain sufficient insertions for analysis but fail to meet the fold-change or p-value threshold. In contrast, genes classified as insufficient data (ID) have fewer than 5 TA sites disrupted by transposon-insertion. The output of gene categorization using these thresholds is displayed for a single rabbit in Figure 2.4B (additional animals in Figure 2.6) and summarized for all animals in Figure 2.4C. An additional criterion that genes be classified as CD in 5 or more of the 7 animals analyzed was imposed to create a consensus list of CD genes (Figure 2.4C). In contrast to the >2000 genes classified as conditionally depleted in one or more animals, only 246 genes were classified as conditionally depleted across 5 or more animals (Table S6).

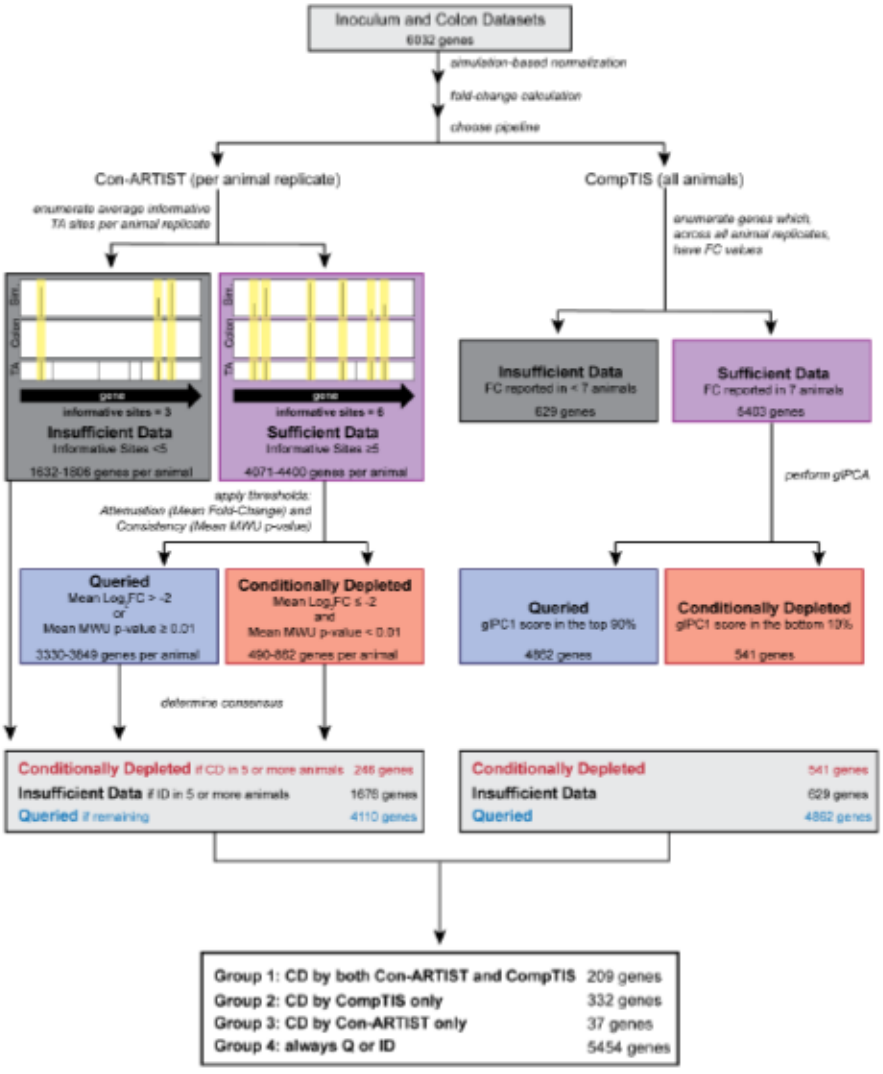


Figure 2.5: Schematic of analytic scheme to identify conditionally depleted genes.

The TIS data was analyzed with Con-ARTIST and CompTIS. The first step is correction for bottleneck effects to facilitate identification of attenuated mutants. To do this, simulation-based normalization is performed on the inoculum dataset to model the stochastic loss observed in the colon datasets. Next, relative abundance of mutants in the normalized inoculum and colon dataset were compared to determine a mean fold-change across each gene. Next, genes with mean fold-change values that are consistent with a signature of attenuation *in vivo* were identified. With Con-ARTIST, genes are first categorized based on their phenotype in each animal replicate, and then a consensus is determined for the phenotype across all replicates. To achieve this, in each animal, genes are first filtered by the number of informative sites—the number of unique TA sites between the input and output that have transposon-insertions. Genes with less than 5 informative sites are classified as insufficient data (ID, black). Genes with sufficient data (≥ 5 TA sites disrupted) are classified based on a dual standard of attenuation (mean $\log_2$ fold change ≤ -2) and consistency (Mann Whitney U p-value <0.01) as either queried (Q, blue) or conditionally depleted (CD, red). To choose genes with consistent phenotypes across the replicates, genes were classified as CD if they were classified as CD in 5 or more replicates. CompTIS synthesizes data across all animal replicates to identify genes important for colonization. Genes are first filtered to remove those without fold-change information across all replicates (ID, black). Gene-level PCA is performed on genes with sufficient data, and genes are classified as Q or CD based on their glPC1 score. Genes with a score in the bottom 10% of the glPC1 score distribution are attenuated (CD).

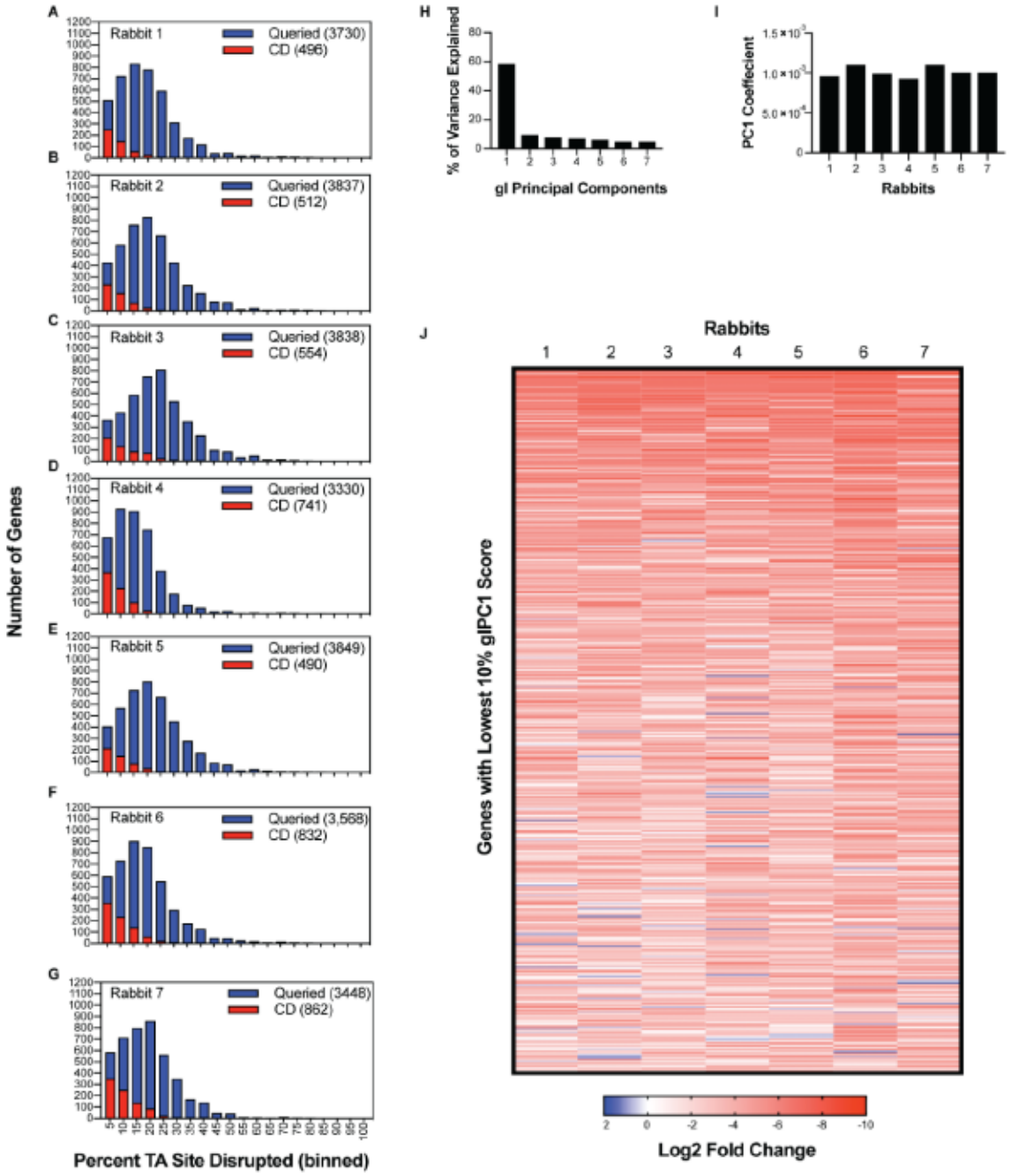


Figure 2.6: Con-ARTIST and CompTIS classification of genes important for colonization.

(A-G) Distribution of percentage TA site disruption in libraries recovered from 7 rabbit colons. These distributions are overlaid with Con-ARTIST classification (queried, blue; CD (conditionally depleted), red). (H) Variance explained by each gene-level (gl) principal component for glPCA performed across the 7 rabbit screens. (I) Gene-level principal component 1 (glPC1) coefficients for each rabbit dataset. (J) Heatmap of the $\log_2$ fold change for each gene with a glPC1 score that falls within the bottom 10% of the distribution. Each column represents genes from a separate rabbit replicate. Genes are ordered by glPC1 score, lowest at the top of the heatmap and highest at the bottom.

CompTIS was also used to compare the seven libraries recovered from rabbit colons. CompTIS relies on PCA, a dimensional reduction approach used to describe the sources of variation in multivariate datasets, and can be used here as an alternative measure of phenotypic consistency between animals. This is particularly beneficial in this case, where the severe bottleneck limits the availability of individual transposon mutant replicates required for Con-ARTIST's MWU p-value thresholds. In the CompTIS analysis, each gene's fold change from the seven colon libraries was subjected to gene level PCA (glPCA)[386]. Genes for which fold change information is not reported in all seven animal replicates are classified as ID. glPC1 describes most of the variation in the animals (Figure 2.6H) and represents a weighted average of the fold change values for each gene across the 7 animals (Table S6). The signs and magnitudes of glPC1 were all similar (Figure 2.6I), indicating that each rabbit contributes approximately equally to glPC1, as expected for biological replicates. The distribution of glPC1 scores is continuous (Figure 2.4D) and describes each gene's contribution to EHEC intestinal colonization. Most genes have a glPC1 score close to zero, suggesting that they do not contribute to colonization. However, the non-symmetrical distribution includes a marked left tail encompassing the lowest 10% of scores, which were classed as CD (Figure 2.4D, 2.6J). The list of 541 CD genes includes nearly all (85%) of the genes classified as CD by the more conservative Con-ARTIST analysis outlined above (Figure 2.5).

These analyses yield four groups of genes (Figure 2.5). Group 1 includes the 209 genes categorized as CD by both ConARTIST and CompTIS and represents the highest confidence candidate genes required for colonization. Group 2 includes 332 genes identified as CD by CompTIS but classified as Q or ID by Con-ARTIST; it includes *ler* (glPC1 = -2290), a critical activator of the LEE T3SS[387], which is classified as queried by Con-ARTIST due to the relative paucity of unique insertion mutants. The lack of *ler* mutants is likely due to the small size of the gene (few TA sites) and to HNS binding occluding transposon insertion[387]. Group 3 includes the 37 genes

identified as CD by ConARTIST but not by CompTIS, due to the absence of fold-change information for all seven replicate rabbits. Lastly, Group 4 includes the 5454 genes always identified as queried or insufficient data. Due to the severe bottleneck, we do not conclude that these genes are *not* attenuated relative to the wild type strain *in vivo*; it is likely that the list of CD loci called by either method is incomplete. Below, we primarily focused on Group 1 genes for further analysis, as they represent the most robust candidates for factors promoting EHEC intestinal colonization.

Using our Z correspondence table (Table S1), 89% (186/209) of Group 1 genes were assigned to a COG functional category. CD genes were frequently associated with amino acid and nucleotide metabolism, signal transduction, and cell wall/envelope biogenesis, but only amino acid metabolism reached statistical significance after correction for multiple hypothesis testing (Figure 2.4E). These genes are also associated with KEGG metabolic pathways (particularly amino acid metabolism), several two-component systems, including *qseC*, which has previously been implicated in EHEC virulence gene regulation, and lipopolysaccharide biosynthesis[194] (Table S7, Figure 2.4F). 30 of the 209 CD genes are EHEC specific, whereas the remaining 179 have homologs in K-12 (Table S6), highlighting the importance of conserved metabolic pathways in the pathogen's capacity to successfully colonize its colonic niche. Similar metabolic pathways were also found to be important for *V. cholerae* growth in the infant rabbit small intestine[366,388], raising the possibility of targeting metabolic pathways such as those for amino acid biosynthesis with antibiotics[389–391].

Analyses of the requirement for T3SS and effectors in colonization

To assess the accuracy of our gene classifications using Con-ARTIST and CompTIS, we examined classifications within the LEE pathogenicity island, which encodes the EHEC T3SS and plays a critical role in intestinal colonization[21,177,180,182,186]. The LEE is comprised of 40 genes, including genes encoding the structural components of the T3SS, some of the pathogen's effectors, their

chaperones, and Intimin, the adhesin that binds to the translocated Tir protein. In infant rabbits, previous studies using single deletion mutants revealed that *tir*, *eae*, and *escN*, the T3SS ATPase, were all required for colonization[21,22]. We observed a marked reduction in the abundance of insertions across nearly the entire LEE in the samples from the rabbit colons relative to the simulation-normalized input reads, indicating this locus is required for colonization (Figure 2.7A). However, the LEE has low GC content and is regulated through HNS binding[387] which when coupled with the infection bottleneck are expected to hamper assessment of gene contributions to intestinal fitness.

12/40 LEE-encoded genes were categorized as Group 1 (CD by both ConARTIST and CompTIS), including 3 genes previously found to be required for colonization (*tir*, *eae* and *escN*) and 8 additional genes critical for T3SS activity, including translocon T3SS components (*espB*, *espD*, and *espA*) and structural components (*escD*, *escQ*, *escV*, *escI*, and *escC*) (Figure 2.8AB, Table S6)[180,392–396]. Additionally, 15 genes were categorized as Group 2, including many encoding chaperones critical for T3SS assembly or T3SS structural components.

We also assessed the contribution of EHEC T3SS effectors on colonization. EHEC has 49 effectors: 6 genes encoded within the LEE (*espF*, *espG*, *espH*, *espZ*, and *map*) and 43 non-LEE encoded effectors (Nle). Nearly all effectors (5/6 LEE-encoded effectors and 41/43 Nle genes) were categorized into Group 4 (not CD). Consistent with these results, previous studies have shown that the LEE-encoded effectors *espG* and *map* are dispensable for robust colonization, and that Δ*espH* and Δ*espF* only had modest colonization defects[22]. The only effectors found to be important for colonization by either Con-ARTIST or CompTIS were *tir* (as expected) and 2 Nle genes: *nleA* and *espM1*. NleA was previously reported to be important for colonic colonization by a related enteric pathogen, *Citrobacter rodentium*[397], and is thought to suppress inflammasome activity[398]; EspM1 is thought to modulate host actin cytoskeletal dynamics[399,400]. Interestingly, mutants in *nleA* and *espM1* were also scored as attenuated in a TraDIS-based study of EDL933 growth in

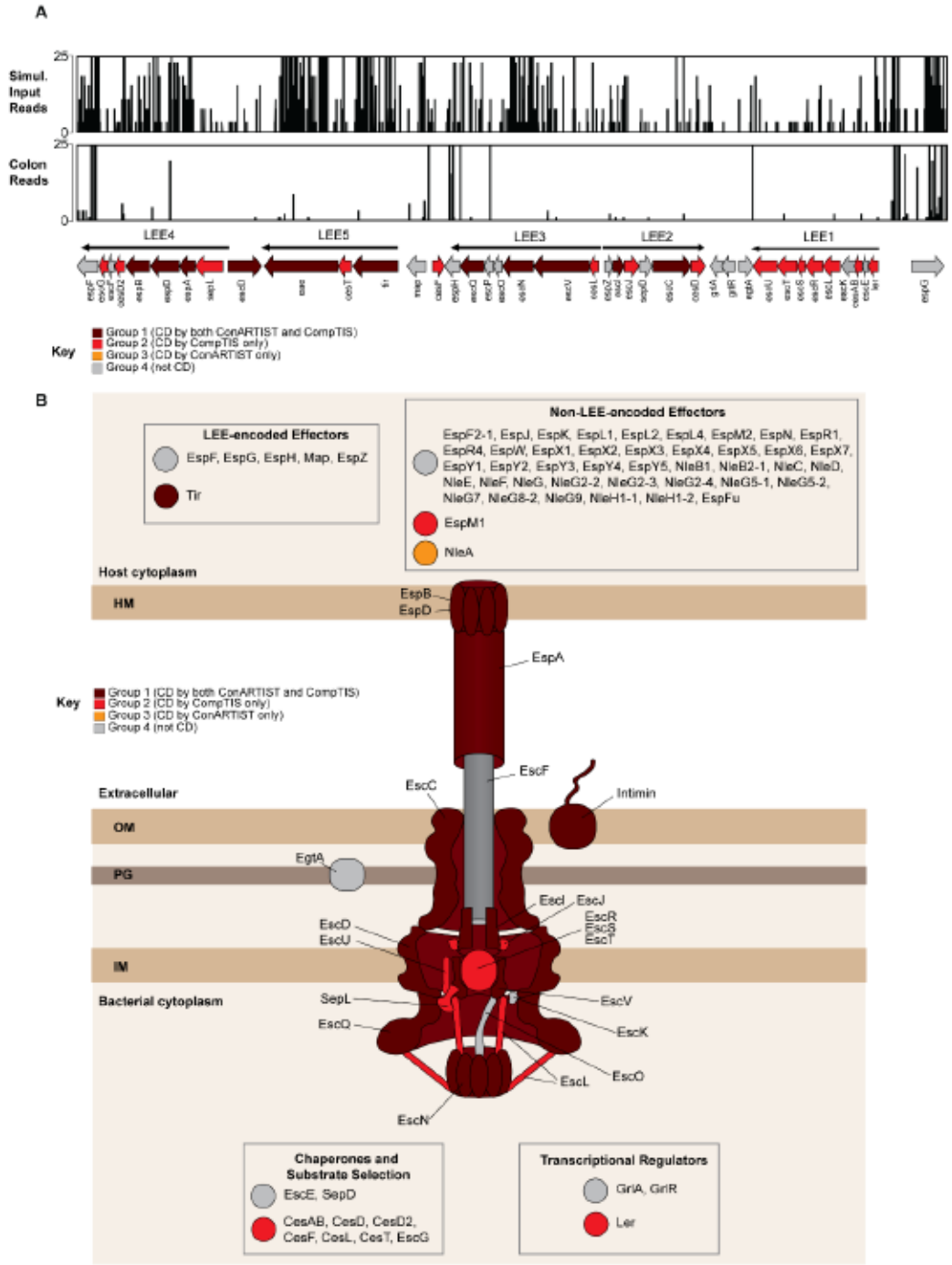


Figure 2.7: Con-ARTIST and CompTIS-based classification of LEE genes and T3SS effectors.

(A) Artemis plots of reads in the LEE pathogenicity island in the control-simulated inoculum library (top) and a representative library recovered from the rabbit colon (middle). The genes in the LEE are displayed at the bottom. The key indicates which classification group each gene belongs: maroon genes belong to Group 1, which are classified as conditionally depleted (CD) by both Con-ARTIST and CompTIS; red genes belong to Group 2, which are classified as CD by CompTIS alone; orange genes belong to Group 3, which are classified as CD by Con-ARTIST alone; and gray genes belong to Group 4, which are not CD. (B) Schematic showing classification of the LEE genes and Nle effectors.

calves, where abundance of mutants in feces was analyzed[401]; however, this study used a much smaller miniTn5 library (covering 855 genes) and relied on a different analytic approach that used a less stringent definition of attenuation (solely based on 2-fold reduction in transposon mutant abundance in output vs input), making comparison to our study difficult. Additional studies are warranted to confirm and further explore how these 2 Nle effectors play pivotal roles promoting intestinal colonization.

Validation of colonization defects in non-LEE encoded genes

We performed further studies of 17 conditionally depleted genes/operons that had not previously been demonstrated to promote EHEC intestinal colonization. All genes were part of Group 1, except *hupB*, which was identified as CD only by CompTIS (Table S6). Mutants with in-frame deletions of either single loci (*agaR*, *cvpA*, *envC*, *htrA*, *hupB*, *mgtA*, *oxyR*, *prc*, *sspA*, *sufI*, *tolC*, and RS09610) or operons classified as conditionally depleted (*acrAB*, *clpPX*, *envZompR*, *phoPQ*, *tatABC*) were generated. Then, each mutant strain was barcoded with unique sequence tags integrated into a neutral locus to enable multiplexed analysis. The barcoded mutants, along with the barcoded WT EHEC, were co-inoculated into infant rabbits to compare the colonization. The relative frequencies of WT and mutant EHEC recovered from infected animals was enumerated by barcode sequencing, and these frequencies were used to calculate competitive indices (CI) for each mutant (i.e., relative abundance of mutant/WT). 14 of the 17 mutants tested had CI values significantly lower than 1, validating the colonization defects inferred from the TIS data (Figure 2.8). The *in vitro* growth of the barcoded mutants was indistinguishable from that of the WT strain (Figure 2.9), suggesting that the *in vivo* attenuation is not explained by a generalized growth deficiency. In aggregate, these observations support our experimental and analytical approaches and provide confirmation that many of the genes classified as CD *in vivo* contribute to intestinal colonization.

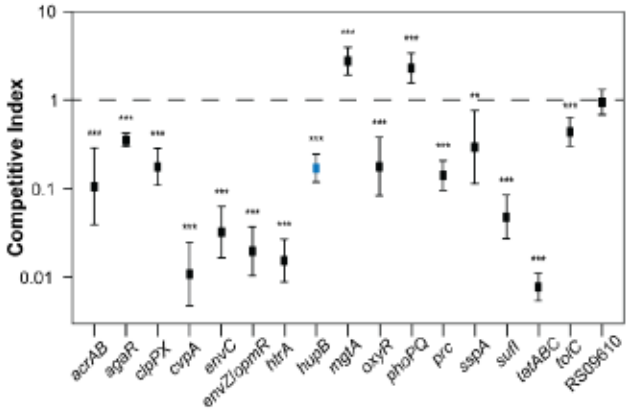


Figure 2.8: Validation of colonization defects in selected mutants.

Competitive indices of indicated mutants vs wild type EHEC. Barcoded mutants were co-inoculated with barcoded wild-type EHEC into infant rabbits and recovered two days later from the colons of infected rabbits. Relative abundance of each mutant was determined by sequencing the barcodes. (**) p-value < 0.01 and (***) p-value <0.001. Δ*hupB*, which had a glPC1 score in the bottom 10%, but was not classified as CD by Con-ARTIST, is highlighted in blue.

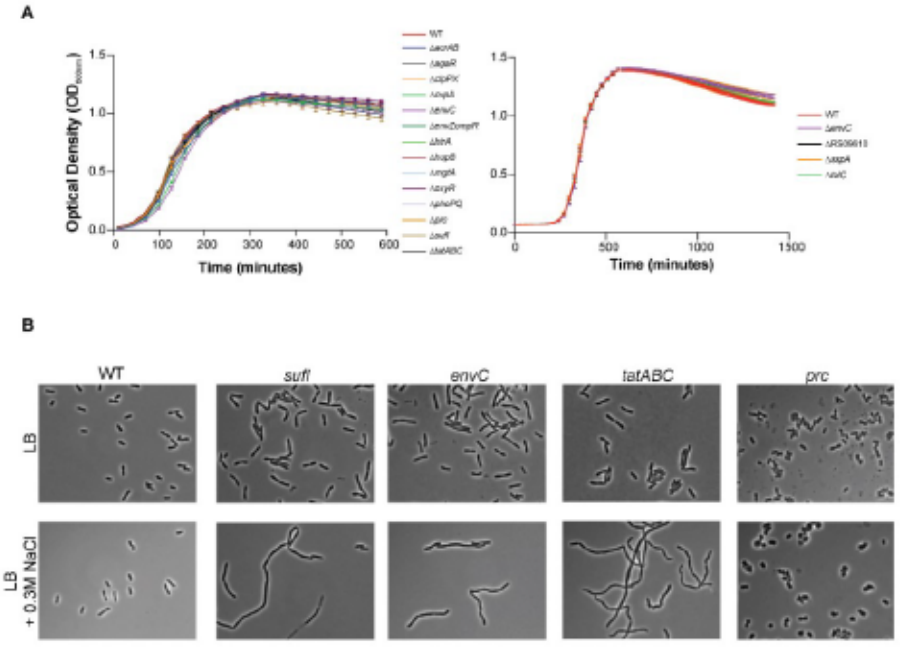


Figure 2.9: *In vitro* growth and morphology of mutants.

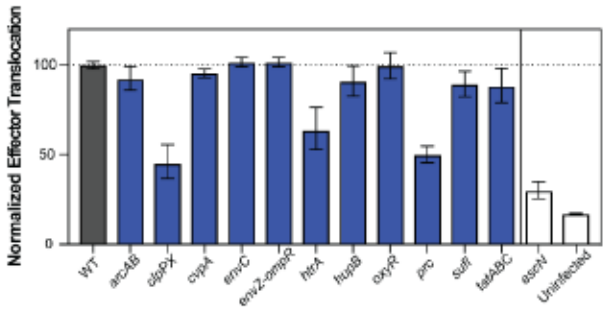
(A) 17 mutant strains plus the wild-type were grown in LB and turbidity measured by optical density. The average of three readings with the standard deviation is plotted. (B) Cell-shape defects of Δ*sufI*, Δ*envC*, Δ*tatABC*, and Δ*prc* mutants in high osmolality media. Morphology in LB (top) or LB supplemented with 0.3M NaCl (bottom) is shown.

Many conditionally depleted loci exhibit reduced T3SS effector translocation and/or increased sensitivity to extracellular stressors

The many new genes implicated in EHEC colonization by the TIS data could contribute to the pathogen's survival and growth *in vivo* by a large variety of mechanisms. Given the pivotal role of EHEC's T3SS in intestinal colonization, as well as previous observations that factors outside the LEE can regulate T3SS gene expression and/or activity[387], we assessed whether T3SS function was impaired in the 11 mutants with CIs <0.3 (Figure 2.10A). Translocation of EspF (an effector protein) fused to a TEM-1 beta-lactamase reporter into HeLa cells was used as an indicator of T3SS functionality[402]. An Δ*escN* mutant, which lacks the ATPase required for T3SS function, was used as a negative control.

Deletions in three protease-encoded genes, *clpPX*, *htrA*, and *prc*, were associated with reduced EspF translocation (Figure 2.10A). Both ClpXP and HtrA have been implicated in T3SS expression/activity in previous reports[403–406]. The ClpXP protease controls LEE gene expression indirectly by degrading LEE-regulating proteins RpoS and GrlR[407]. The periplasmic protease HtrA (also known as DegP) has been implicated in post-translational regulation of T3SS as part of the Cpx-envelope stress response[405,406]. Interestingly, *prc*, which also encodes a periplasmic protease[407], also appears required for robust EspF translocation. Prc has been implicated in the maintenance of cell envelope integrity under low and high salt conditions in *E. coli* K-12[408]. Consistent with this observation, in high osmolarity media a Δ*prc* EHEC mutant exhibited cell shape defects (Figure 2.9B). Deficiencies in the cell envelope associated with absence of Prc may impair T3SS assembly and/or function, perhaps also by triggering the Cpx-envelope stress response. Together, these observations suggest that *in vivo* these three proteases modulate T3SS expression/function, thereby promoting EHEC intestinal colonization.

A



B

Strain	NaCl (mM)	Bile (%)
WT	1500	12
Δ*acrAB*	1500	**0.05**
Δ*clpPX*	1500	**6**
Δ*cvpA*	1500	**3**
Δ*envC*	**750**	**0.05**
Δ*envZ-ompR*	**750**	**1.5**
Δ*htrA*	1500	12
Δ*hupB*	1500	12
Δ*oxyR*	1500	**0.75**
Δ*prc*	**750**	**6**
Δ*sufI*	1500	12
Δ*tatABC*	**750**	**6**

C

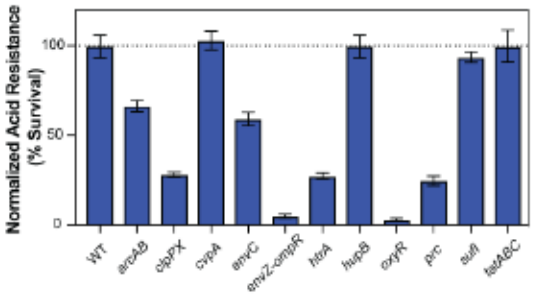


Figure 2.10: Effector translocation and survival in response to various gastro-intestinal stressors by mutants defective in colonization.

(A) Normalized effector translocation of mutants compared to WT. ΔescN, a mutant that abrogates T3SS activity, was used as a control. Mutants were tested for their ability to translocate EspF-TEM1 into HeLa cells, as measured by a shift in emission spectra from 520 to 450 nm. Fluorescence was normalized to WT levels. Geometric means and geometric standard deviations are plotted. (B) MIC for NaCl (osmotic stress) and crude bile for the indicated mutants. Bold text highlights values differing from the wild-type. (C) Normalized acid resistance. Mutants were tested for their ability to survive low acid shock. Survival is shown as a percentage of the acid resistance of the WT. Geometric mean and geometric standard deviation

We also investigated the capacity of each of the 11 mutant strains to survive challenge with three stressors–low pH, bile, and high salt (osmotic challenge)–that the pathogen may encounter in the gastrointestinal tract. Relative to the WT strain, all but one (*sufI*) of the mutant strains exhibited reduced survival following one or more of these challenges (Figure 2.10BC), suggesting that exposure to these host environmental factors may contribute to the *in vivo* attenuation of these mutants. Many of the EHEC mutants exhibited sensitivities to external stressors that are consistent with previously described phenotypes in other organisms and experimental systems. For example, the EHEC Δ*acrAB* locus, which was associated with bile sensitivity in EHEC (Figure 2.10B), is

known to contribute to a multidrug efflux system that can extrude bile salts, antibiotics, and detergents[409]. Our observation that mutants lacking the oxidative stress response gene *oxyR* are sensitive to bile and to acid pH is also concordant with previous reports linking both stimuli to oxidative stress[114,410,411]. Furthermore, the heightened sensitivity to bile, acid, and elevated osmolarity of EHEC lacking the two-component regulatory system EnvZ/OmpR is consistent with previous reports that EnvZ/OmpR is a critical determinant of membrane permeability, due to its regulation of outer membrane porins OmpF and OmpC. Mutations that activate this signaling system (in contrast to the deletions tested here) have been found to promote *E. coli* viability *in vivo* and to enhance resistance to bile salts[412].

The EHEC Δ*tatABC* mutant exhibited a marked colonization defect and a modest increase in bile sensitivity. The twin-arginine translocation (Tat) protein secretion system, which transports folded protein substrates across the cytoplasmic membrane[413,414], has been implicated in the pathogenicity of a variety of Gram-negative pathogens, including enteric pathogens such as *Salmonella enterica* serovar Typhimurium[415–417], *Yersinia pseudotuberculosis*[418,419], *Campylobacter jejuni*[420], and *Vibrio cholerae*[421]. Attenuation of Tat mutants can reflect the combined absence of a variety of secreted factors. For example, the virulence defect of *S. enterica* Typhimurium *tat* mutants are likely due to cell envelope defects caused by the inability to secrete the periplasmic cell division proteins AmiA, AmiC and SufI[416]. Notably, single knock-outs of any of these genes does not cause *S. enterica* attenuation[416], but altogether their absence renders the cell-envelope defective and more sensitive to cell-envelope stressors, such as bile acids[417].

In EHEC, the Tat system has been implicated in Stx1 export[422], but because Stx1 was not a hit in our screen and is not thought to modulate intestinal colonization in infant rabbits[21], it is not likely to explain the marked colonization defect of the EHEC Δ*tatABC* mutant. The suite of EHEC Tat substrates has not been experimentally defined, although putative Tat substrates can be identified by

a characteristic signal sequence[413,414]. A few substrates, including SufI, OsmY, OppA, MglB, and H7 flagellin, have been detected experimentally[422]. *sufI*, interestingly, was also a validated hit in our screen, and is the only CD gene that has a predicted Tat-secretion signal. However, the Δ*sufI* mutant did not display enhanced bile sensitivity, suggesting that attenuation of this mutant, and perhaps of the Δ*tatABC* mutant as well, reflects deficiencies in other processes. SufI is a periplasmic cell division protein that localizes to the divisome and may be important for maintaining divisome assembly during stress conditions[423,424]. *E. coli tat* mutants have septation defects[425], presumably from loss of SufI at the divisome. Interestingly, *envC*, another validated CD gene, encodes a septal murein hydrolase[426] that is required for cell division, and the Δ*envC* mutant also displayed increased bile sensitivity. Consistent with this hypothesis, in high osmolarity media, the Δ*sufI*, Δ*envC*, and Δ*tatABC* mutants exhibited septation or cell shape defects (Figure 2.10B). Collectively, these data suggest that an impaired capacity for cell division may reduce EHEC's fitness for intraintestinal growth, and that at times this may reflect increased susceptibility to clearance by host factors such as bile.

CvpA promotes EHEC resistance to deoxycholate

We further characterized EHEC Δ*cvpA* because other TIS-based studies of the requirements for colonization by diverse enteric pathogens (*Vibrio cholerae*, *Vibrio parahaemolyticus* and *Salmonella enterica* serovar Typhimurium) also classified *cvpA* as important for colonization, but did not explore the reasons for mutant attenuation[366,376,427]. *cvpA* has been linked to colicin V export in *E. coli* K-12[428] as well as curli production and biofilm formation in UPEC[429]. The EHEC Δ*cvpA* mutant did not exhibit an obvious defect in biofilm formation or curli production (Figure 2.11AB), suggesting that *cvpA* may have a distinct and previously unappreciated role in pathogenicity.

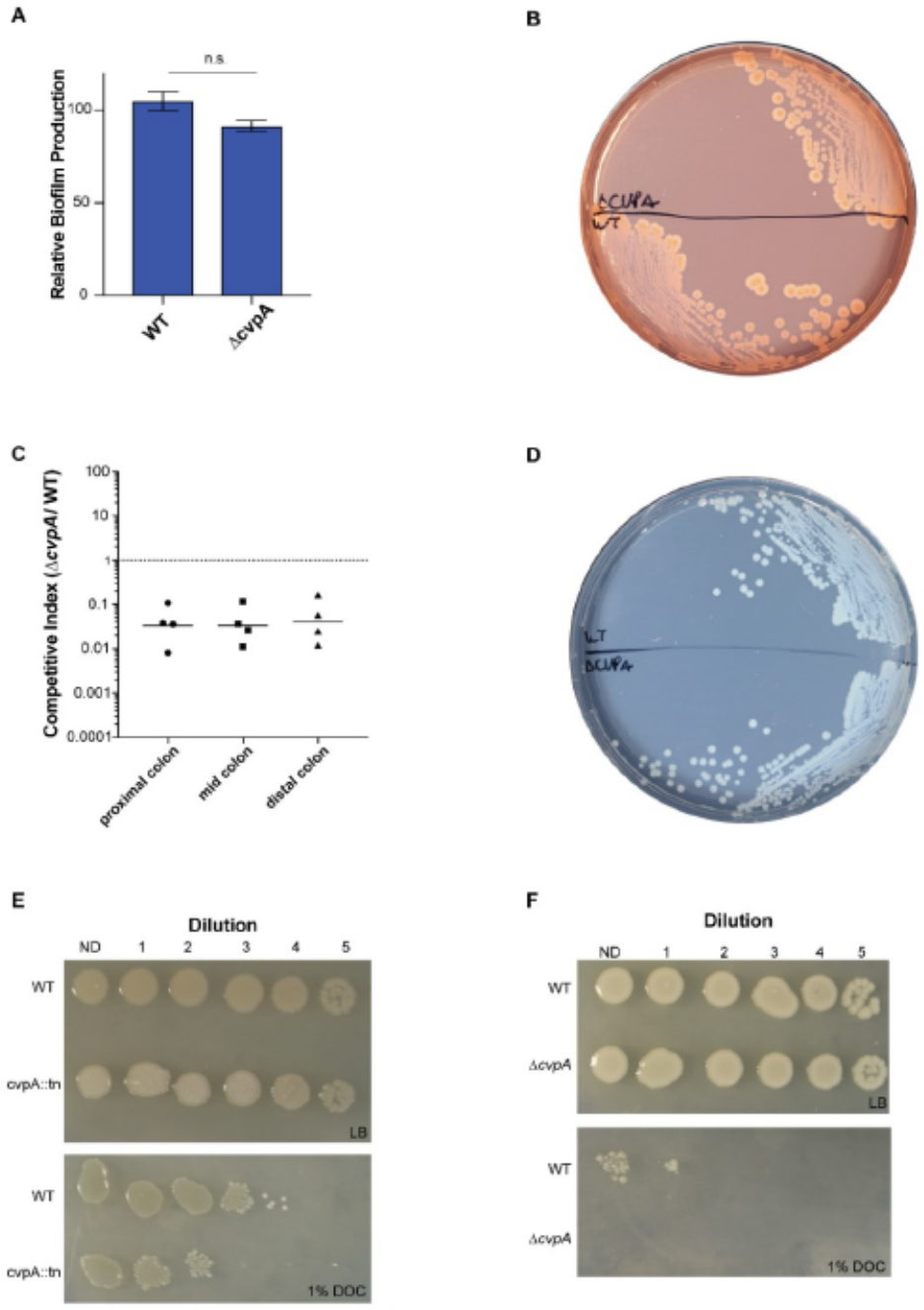
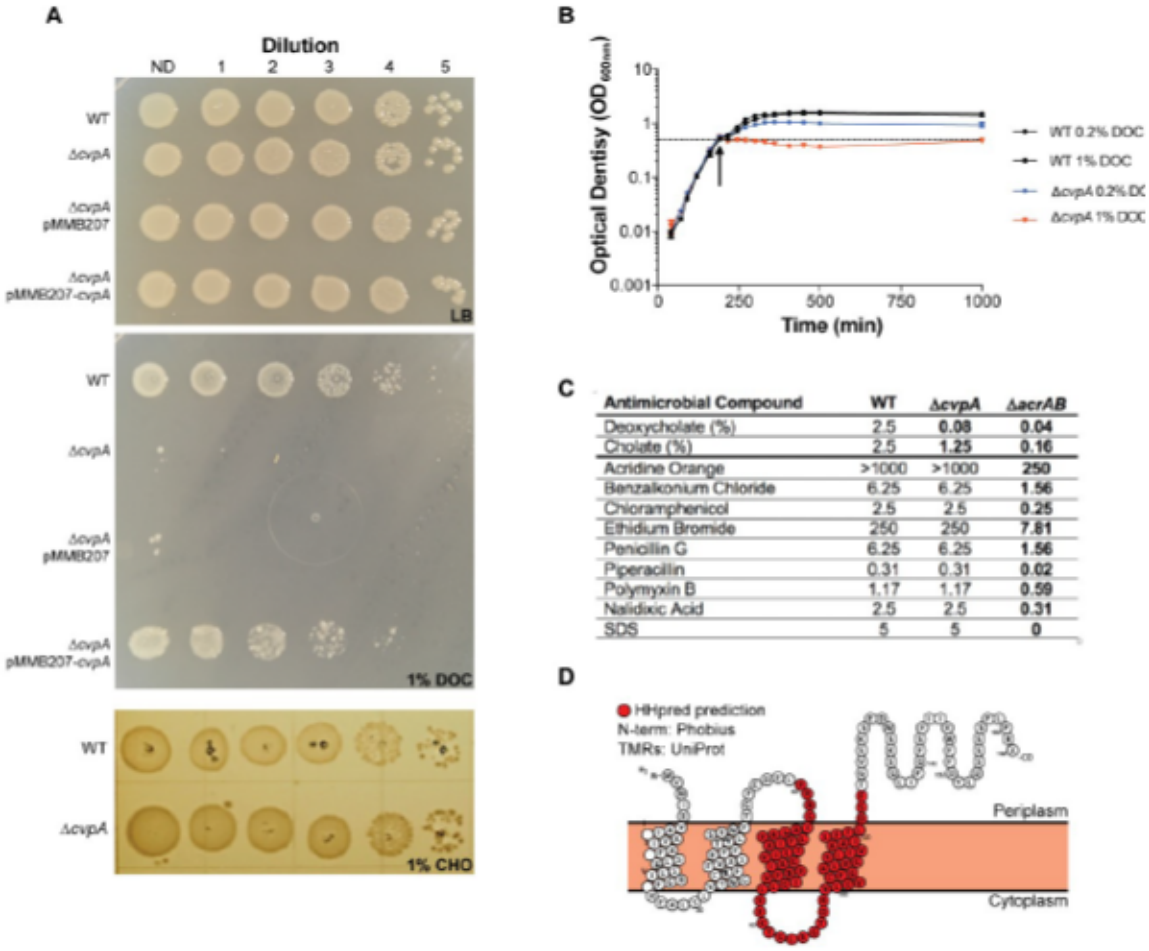


Figure 2.11: Characterization of Δ*cvpA*.

(A) Biofilm production in WT and Δ*cvpA* using crystal violet staining and absorption. Levels were normalized to a percent of the WT value; three samples were analyzed and the geometric means and geometric standard deviation are plotted. The differences between the two groups were not significant (n.s.) by Mann-Whitney U. (B) WT and Δ*cvpA* struck to single colonies on an agar plate made with YESCA media supplemented with Congo Red to detect curli fibers. (C) 1:1 competitive infection between Δ*cvpA* and Δ*lacZ*. (D) WT and Δ*cvpA* struck to single colonies on an agar plate containing minimal media with no exogenous purines. (E) Dilution series of *Vibrio cholerae* C6706 WT and *cvpA*::tn plated on LB and LB 1% deoxycholate (DOC). (F) Dilution series of *Vibrio parahaemolyticus* WT and Δ*cvpA* plated on LB and LB 1% deoxycholate (DOC).

Initially, we confirmed that the Δ*cpxA* mutant exhibits an intestinal colonization defect in 1:1 competition vs the wild type strain (~20-fold defect, Figure 2.12C). To further characterize the sensitivity of the EHEC Δ*cpxA* mutant to bile (Figure 2.12B), we exposed the mutant to the two major bile salts found in the gastrointestinal tract, cholate (CHO) and deoxycholate (DOC)[114,115] (Figure 2.12AC). In contrast to WT EHEC, which displayed equivalent sensitivity to the two bile salts in MIC assays (MIC = 2.5% for both), the Δ*cpxA* mutant was much more sensitive to DOC than to CHO (MIC = 0.08% versus 1.25%). The Δ*cpxA* mutant's sensitivity to deoxycholate was present both in liquid cultures and during growth on solid media (Figure 2.12ABC). *cpxA* lies upstream of the purine biosynthesis locus *purF*, and some Δ*cpxA* mutant phenotypes have been attributed to reduced expression of *purF* due to polar effects[428,430], but the growth of our Δ*cpxA* mutant was not impaired in the absence of exogenous purines (Figure 2.11D), suggesting the *cpxA* deletion does not adversely modify *purF* expression. Moreover, the DOC sensitivity phenotype of Δ*cpxA* was restored by plasmid-based expression of *cpxA*, confirming that this phenotype is linked to the absence of *cpxA* (Figure 2.12A).

Bile sensitivity has been associated with defects in the bacterial envelope or with reduced efflux capacity[115]. We assessed the growth of the Δ*cpxA* mutant in the presence of a variety of agents that perturb the cell envelope to assess the range of the defects associated with the absence of *cpxA*. The MICs of WT and Δ*cpxA* EHEC were compared to those of an Δ*acrAB* mutant, whose lack of a broad-spectrum efflux system provided a positive control for these assays. Notably, the Δ*cpxA* mutant did not exhibit enhanced sensitivity to any of the compounds tested other than bile salts. In marked contrast, the Δ*acrAB* mutant displayed increased sensitivity to all agents assayed (Figure 2.12C). These observations suggest that the sensitivity of the *cpxA* mutant to DOC is not likely attributable to a general cell envelope defect in this strain. *Vibrio cholerae* and *Vibrio*



Antimicrobial Compound	WT	*ΔcvpA*	*ΔacrAB*
Deoxycholate (%)	2.5	**0.08**	**0.04**
Cholate (%)	2.5	**1.25**	**0.16**
Acridine Orange	>1000	>1000	**250**
Benzalkonium Chloride	6.25	6.25	**1.56**
Chloramphenicol	2.5	2.5	**0.25**
Ethidium Bromide	250	250	**7.81**
Penicillin G	6.25	6.25	**1.56**
Piperacillin	0.31	0.31	**0.02**
Polymyxin B	1.17	1.17	**0.59**
Nalidixic Acid	2.5	2.5	**0.31**
SDS	5	5	**0**

Figure 2.12: CvpA promotes EHEC resistance to deoxycholate.

(A) Dilution series of WT, Δ*cvpA* mutant, and Δ*cvpA* mutant with arabinose-inducible *cvpA* complementation plasmid plated on LB, LB 1% cholate (CHO) or LB 1% deoxycholate (DOC). (B) Optical density of WT and Δ*cvpA* grown in LB and two concentrations of DOC, added at the indicated arrow. The average of three readings is plotted with errors bars indicating standard deviation. (C) MIC of antimicrobial compounds for WT and Δ*cvpA* and Δ*acrAB* mutants. Units are mg/mL unless specified otherwise. Bolded values are those different than the wild-type. (D) Schematic of predicted CvpA topology.

parahaemolyticus Δ*cvpA* mutants also exhibited sensitivity to DOC (Figure 2.12EF), implying a similar role in bile resistance in these distantly related enteric pathogens.

A variety of bioinformatic algorithms (PSLPred, HHPred, Phobius, Phyre2) suggest that CvpA is an inner membrane protein with 4–5 transmembrane elements similar to small solute transporter proteins (Figure 2.12D). Phyre2 and HHPred reveal CvpA's partial similarity to inner membrane transporters in the Major Facilitator Superfamily of transporters (MFS) and the small-conductance mechanosensitive channels family (MscC). The PFAM database groups CvpA (PF02674) in the LysE transporter superfamily (CL0292), a set of proteins known to enable solute export. In conjunction with findings presented above, these predictions raise the possibility that CvpA is important for the export of a limited set of substrates that includes DOC. Additional studies to confirm this hypothesis and to establish how CvpA enables export are warranted, particularly because this protein is widespread amongst enteric pathogens.

Conclusions

Here, we created a highly saturated transposon library in EHEC EDL933 to identify the genes required for *in vitro* and *in vivo* growth of this important food-borne pathogen using TIS. This approach has transformed our capacity to rapidly and comprehensively assess the contribution an organism's genes to growth in different environments[367,431,432]. However, technical and biologic issues can confound interpretation of genome-scale transposon-insertion profiles. For example, we found that EHEC genes with low GC content or those without homologs in K-12 were less likely to contain transposon-insertions, which may reflect processes other than reduced biological fitness. For example, the presence of nucleoid binding proteins could impair transposon-insertion into these loci. Unexpectedly, more than 100 of the genes conserved between EHEC and K-12 appear to promote the growth of the pathogen but not that of K-12, suggesting that strain-specific processes

may have enabled divergence of the metabolic roles of ancestral *E. coli* genes in these backgrounds. More comprehensive comparisons between additional *E. coli* isolates are needed to confirm this hypothesis.

Materials and Methods

Ethics statement

All animal experiments were conducted in accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health and the Animal Welfare Act of the United States Department of Agriculture using protocols reviewed and approved by Brigham and Women's Hospital Committee on Animals (Institutional Animal Care and Use Committee protocol number 2016N000334 and Animal Welfare Assurance of Compliance number A4752-01)

Bacterial strains, plasmids and growth conditions

Strains, plasmids and primers used in this study are listed in Tables S8 and S9. Strains were cultured in LB medium or on LB agar plates at 37°C unless specified otherwise. Antibiotics and supplements were used at the following concentrations: 20 μg/mL chloramphenicol (Cm), 50 μg/mL kanamycin (Km), 10 μg/mL gentamicin (Gm), 50 μg/mL carbenicillin (Cb), and 0.3 mM diaminopimelic acid (DAP).

A gentamicin-resistant mutant of *E. coli* O157:H7 EDL933 (Δ*lacI*::*aacC1*) and a chloramphenicol-resistant mutant of *E. coli* K-12 MG1655 (Δ*lacI*::*cat*) were used in this study for all experiments, and all mutations were constructed in these strain backgrounds except where specified otherwise. The Δ*lacI*::*aacC1* and Δ*lacI*::*cat* mutations were constructed by standard allelic exchange techniques[433] using a derivative of the suicide vector pCVD442 harboring a gentamicin resistance cassette

amplified from strain TP997 (Addgene strain #13055)[434] or a chloramphenicol resistance cassette from plasmid pKD3 (Addgene plasmid #45604)[378] flanked by the 5' and 3' DNA regions of the *lacI* gene. Isogenic mutants of EDL933 Δ*lacI*::*aacC1* were also constructed by standard allelic exchange using derivatives of suicide vector pDM4 harboring DNA regions flanking the gene(s) targeted for deletion. *E. coli* MFDλpir[435] was used as the donor strain to deliver allelic exchange vectors into recipient strains by conjugation. Sequencing was used to confirm mutations.
A Δ*vpA* strain was also constructed using standard allelic exchange in a streptomycin- resistant mutant of *V. parahaemolyticus* RIMD 2210633. A *vpA*::tn mutant was used from a *Vibrio cholerae* C6706 arrayed transposon library[436].

Eukaryotic cell line and growth conditions

HeLa cells were obtained from The Harvard Digestive Diseases Center (HDDC) Core Facilities at Boston Children's Hospital and were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS). Cells were grown at 37°C with 5% CO_2 and routinely passaged at 70 to 80% confluence; medium was replenished every 2 to 3 days.

Transposon-insertion library construction

To create transposon-insertion mutant libraries in EHEC EDL933 Δ*lacI*::*aacC1*, conjugation was performed to transfer the transposon-containing suicide vector pSC189[437] from a donor strain (*E. coli* MFDλpir) into the EDL933 recipient. Briefly, 100 µL of overnight cultures of donor and recipient were pelleted, washed with LB, and combined in 20 µL of LB. These conjugation mixtures were spotted onto a 0.45 µm HA filter (Millipore) on an LB agar plate and incubated at 37°C for 1 h. The filters were washed in 8 mL of LB and immediately spread across three 245x245 mm (Corning) LB-agar plates containing Gm and Kn. Plates were incubated at 37°C for 16 h and

then individually scraped to collect colonies. Colonies were resuspended in LB and stored in 20% glycerol (v/v) at -80°C as three separate library stocks. The three libraries were pooled to perform essential genes analysis, and one library aliquot was used to as an inoculum for infant rabbit infection studies.

To create TIS mutant libraries in *E. coli* K-12 MG1655 Δ*lacI*::cat, conjugation was performed as above. 200 uL of overnight culture of the donor strain (*E. coli* MFDλpir carrying pSC189) and the recipient strain (MG1655 Δ*lacI*::cat) were pelleted, washed, combined and spotted on 0.45 μm HA filters at 37°C for 5.5 hours. Cells were collected from the filter, washed, plated on selective media (LB Km, Cm), and incubated overnight at 30°C. Colonies were resuspended in LB and frozen in 20% glycerol (v/v). An aliquot was thawed and gDNA isolated for analysis.

Infant rabbit infection with EHEC transposon-insertion library

Mixed gender litters of 2-day-old New Zealand White infant rabbits were co-housed with a lactating mother (Charles River). To prepare the EHEC transposon-insertion library for infection of infant rabbits, 1 mL from one library aliquot was thawed and added to 20 mL of LB. After growing the culture for 3 h at 37°C with shaking, the OD_{600} was measured and 30 units of culture at OD_{600} = 1 (about 8 mL) were pelleted and resuspended in 10 mL PBS. Dilutions of the inoculum were plated on LB agar plates with Gm and Km for precise dose determination. An aliquot of the inoculum was saved for subsequent gDNA extraction and sequencing (input). Each infant rabbit was infected orogastrically with 500 μl of the inoculum ($1x10^9$ CFU) using a size 4 French catheter. Following inoculation, the infant rabbits were monitored at least 2x/day for signs of illness and euthanized 2 days postinfection. The entire intestinal tract was removed from euthanatized rabbits, and sections of the mid-colon were removed and homogenized in 1 mL of sterile PBS using a minibeadbeater-16 (BioSpec Products, Inc.). 200 μL of tissue homogenate from the colon were plated on LB agar

containing Gm and Km to recover viable transposon-insertion mutants. Plates were grown for 16 h at 37°C. The next day, colonies were scraped and resuspended in PBS. A 5 mL aliquot of cells was used for genomic DNA extraction and subsequent sequencing (Rabbits 1–7).

Characterization of transposon-insertion libraries

Transposon-insertion libraries were characterized as described previously[376,438]. Briefly, for each library, gDNA was isolated using the Wizard Genomic DNA extraction kit (Promega). gDNA was then fragmented to 400–600 bp by sonication (Covaris E220) and end repaired (Quick Blunting Kit, NEB). Transposon junctions were amplified from gDNA by PCR. PCR products were gel purified to isolate 200-500bp fragments. To estimate input and ensure equal multiplexing in downstream sequencing, purified PCR products were subjected to qPCR using primers against the Illumina P5 and P7 hybridization sequence. Equimolar DNA fragments for each library were combined and sequenced with a MiSeq. Reads for all TIS libraries have been deposited in the SRA database (Accession Number: PRJNA548905).

Reads were first trimmed of transposon and adaptor sequences using CLC Genomics Workbench (QIAGEN) and then mapped to *Escherichia coli* O157:H7 strain EDL933 (NCBI Accession Numbers: chromosome, NZ_CP008957.1; pO157 plasmid, NZ_CP008958.1) using Bowtie without allowing mismatches. Reads were discarded if they did not align to any TA sites, and reads that mapped to multiple TA sites were randomly distributed between the multiple sites. After mapping, sensitivity analysis was performed on each library to ensure adequate sequencing depth by sub-sampling reads and assessing how many unique transposon mutants were detected. Next, the data was normalized for chromosomal replication biases and differences in sequencing depth using a LOESS correction of 100,000-bp and 10,000-bp windows for the chromosome and plasmid,

respectively. The number of reads at each TA site was tallied and binned by gene and the percentage of disrupted TA sites was calculated. Genes were binned by percentage of TA sites disrupted.
For essential gene analysis, EL-ARTIST was used as described previously[376]. Protein-coding genes, RNA-coding genes, and pseudogenes were included in this analysis. Briefly, EL-ARTIST classifies genes into one of three categories (underrepresented, regional, or neutral), based on their transposon-insertion profile. Classifications are obtained using a hidden Markov model (HMM) analysis following sliding window (SW) training ($p < 0.05$, 10 TA sites). Insertion-profiles for example genes were visualized with Artemis.

For identification of mutants conditionally depleted in the rabbit colon as compared to the input inoculum, Con-ARTIST was used as described previously[438]. First, the input library was normalized to simulate the severity of the bottleneck as observed in the libraries recovered from rabbit colons using multinomial distribution-based random sampling ($n = 100$). Next, a modified version of the Mann-Whitney U (MWU) function was applied to compare these 100 simulated control data sets to the libraries recovered from the rabbit colon. All genes were analyzed, but classification as 'conditionally depleted' (CD) was restricted to genes that had sufficient data (≥5 informative TA sites), met our standard of attenuation (mean $\log_2$ fold change ≤ -2), met our standard of phenotypic consistency (MWU p-value of ≤0.01), and had a consensus classification in 5 or more of the 7 animals analyzed. Genes with ≥5 informative TA sites that fail to exceed both standards of attenuation and consistency are classified as 'queried' (Q, blue), whereas genes with less than 5 informative TA sites are classified as 'insufficient data' (ID).

Gene-level PCA (glPCA) was performed using CompTIS, a principal component analysis-based TIS pipeline, as described in[386]. Briefly, $\log_2$ fold change values were derived by comparing read abundance in each sample to 100 control-simulated datasets as in Con-ARTIST. These fold change values were weighted to minimize noise due to variability[386]. Next, genes that did not have a fold

change reported for all 7 animals were discarded. The fold change values were then z-score normalized. Weighted PCA was performed in Matlab (Mathworks) with the PCA algorithm (pca). Genes were categorized into 4 groups based on their classifications in Con-ARTIST and CompTIS.

GC content

The GC content of classified genes was compared using a Mann-Whitney U statistical test and a Bonferroni correction for multiple hypothesis correction when more than one comparison was made. A p-value <0.05 was considered significant for one comparison, p<0.025 for two. A Fisher's exact two-tailed t-test was used to compare ratios of classifications between groups, where a p-value of <0.01 was considered significant.

In vivo competitive infection: pooled mutants versus wild-type

Barcodes were introduced into Δ*lacI*::*aacC1*and isogenic mutant strains as described previously[376,439]. Briefly, a 991bp fragment of *cynX* (RS02015) that included 51bp of the intergenic region between *cynX* and *lacA* (RS02020) was amplified using primers that contained a 30 bp stretch of random sequence and cloned into SacI and XbaI digested pGP704. The resulting pSoA176.mix was transformed into *E. coli* MFDλpir. Individual colonies carrying unique tag sequences were isolated and used as donors to deliver pSoA176 barcoded derivatives to EDL933 Δ*lacI*::*aacC1*and each isogenic mutant strain. Three barcodes were independently integrated into EDL933 Δ*lacI*::*aacC1*, and three barcodes into each isogenic mutant via homologous recombination in the intergenic region between *cynX* and *lacA*, which tolerates transposon-insertion *in vitro* and *in vivo*, indicating this locus is neutral for the fitness of the bacteria. Correct insertion of barcodes was confirmed by PCR and sequencing.

To prepare the culture of mixed EHEC-barcoded strains for the multi-coinfection experiment, 100 µl of overnight cultures of the barcoded strains were mixed in a flask and 1 mL of this mix was added to 20 mL LB. After growing the culture for 3 h at 37°C with shaking, the OD_{600} was measured and 30 units of culture at OD_{600} = 1 (about 8 mL) were pelleted and resuspended in 10 mL PBS. Dilutions of the inoculum were plated on LB agar plates with Gm and Cb for precise CFU determination. 10 infant rabbits were inoculated and monitored as described above, and colon samples collected. Tissue homogenate was plated, and CFU were collected the following day. gDNA was extracted and prepared for sequencing as described previously[439].

The quantification of sequence tags was done as described previously[439]. In brief, sequence tags were amplified from the inoculum culture and libraries recovered from rabbit colons. The relative *in vivo* fitness of each mutant was assessed by calculating the competitive index (CI) as follows: We compare two strains (Δ*lacI*::*aacC1* and isogenic mutant) in a population with frequencies f_{wt} and $f_{mut,x}$ respectively where x is one of 17 mutant strains with a deletion in gene x. For simplicity, we assume here that both expand exponentially from a time point t_0 to a sampling time point t_s, their relative fitness (offspring/generation) is proportional to the competitive index CI: $ln\left(\frac{f_{mut,x,s}/f_{mut,x,0}}{f_{wt,s}/f_{wt,0}}\right) = ln\,(CI)$. Here, $f_{wt,0}$ and $f_{mut,x,0}$ are the frequencies of the strains in the inoculum, measured in triplicates, and $f_{wt,s}$ and $f_{mut,x,s}$ describe the frequencies at the sampling time point in the animal host. Because the WT strain was tagged with 3 individual tags and the inoculum was measured in triplicate, we have 3x3=9 measurements of the ratio $f_{wt,s}/f_{wt,0}$. The same is true for all mutant strains, such that we have 9 measurements of the ratio $f_{mut,x,s}/f_{mut,x,0}$. In total, we therefore have 3x3x3x3= 81CI measurements for each mutant per animal. To determine intra-host variance in these 81 measurements, a 95% confidence interval of the CI in single animal hosts was

determined by bootstrapping. For combining the CIs measured across all 10 animal hosts, we performed a random-effects meta-analysis using the metafor package [121] in the statistical software package R (version 3.0.2). The pooled rate proportions and 95% confidence intervals were calculated using the estimates and the variance of CIs in each animal determined by bootstrapping and corrected for multiple testing using the Benjamini-Hochberg procedure.

In vitro growth

Each bacterial strain was grown at 37°C overnight. The next day, cultures were diluted 1:1000 into 100 uL of LB in 96-well growth curve plates in triplicate. Plates were left shaking at 37°C for 10–24 hours. Absorbance readings at 600nm were normalized to a blank, and the average of each triplicate was taken as the optical density.

T3SS translocation assays

T3SS functionality was assessed by translocation of the known EHEC T3SS effector protein EspF into HeLa cells as described previously[396]. Briefly, the plasmid encoding the effector protein EspF fused to TEM-1 beta-lactamase was transformed into each of the bacterial strains to be tested. Overnight cultures of each bacterial strain were diluted 1/50 in DMEM supplemented with HEPES (25mM), 10% FBS and L-glutamine (2mM) and incubated statically at 37°C with 5% CO2 for two hours. This media is known to induce T3SS expression[440]. HeLa cells were seeded at a density of $2x10^4$ cells in 96-well clear bottom black plates and infected for 30 minutes at an MOI of 100. After 30 minutes of infection IPTG was added at a final concentration of 1mM to induce the plasmid-encoded T3SS effector. After an additional hour of incubation, monolayers were washed in HBSS solution and loaded with fluorescent substrate CCF2/AM solution (Invitrogen) as recommended by the manufacturer. After 90 minutes, fluorescence was quantified in a plate fluorescence reader with

excitation at 410nm and emission was detected at 450nm. Translocation was expressed as the emission ratio at 450/520nm to normalize beta-lactamase activity to cell loading and the number of cells presented at each well, and then normalized to WT levels of translocation.

Biofilm, curli production, and purine assays

Biofilm and curli production assays were performed as described previously[429]. For biofilm assays, bacterial cultures were grown in yeast extract-Casamino Acids (YESCA) medium until they reached an OD_{600} ~ 0.5 and 1/1000 dilution of this culture was used to seed 96-well PVC plates. The cultures were grown at 30°C for 48 hours and biofilm production was quantitatively measured using crystal violet staining and absorbance reading at 595nm. Relative biofilm production was normalized to the average of three WT samples. A two-tailed Mann Whitney U test was used to determine if biofilm production was significantly different. A p-value < 0.05 was considered to be significant. To test curli production, bacterial cultures were grown in YESCA medium until they reached an OD_{600} ~ 0.5 and then were struck to single colonies onto YESCA agar plates supplemented with Congo Red. Red colonies indicate curli production. To test if our Δ*apA* deletion had polar effects on *purF*, the mutant and WT were struck onto minimal media lacking exogenous purines.

Acid shock assays

An adaptation of the acid shock method described previously[441] was performed. Briefly, bacterial cultures were grown until mid-exponential phase (OD_{600} ~ 0.6), then diluted 20-fold in LB pH 5.5 and incubated for 1 hour before preparing serial dilutions and plating each culture to determine the relative percentage of survival in comparison to the wild-type EDL933 strain. The pH of the LB broth was adjusted using sterilized 1mM HCl and buffered with 10% MES. Values are expressed as percent survival normalized to WT.

MIC assays

MIC assays were performed using an adaptation of a standard methodology with exponential-phase cultures[442]. Briefly, the different compounds to be tested (see Figure 2.6B and 2.7C) were prepared in serial 2-fold dilutions in 50 ul of LB in broth in a 96-well plate format. To each well was added 50 ul of a culture prepared by diluting an overnight culture 1,000-fold into fresh LB broth, growing it for 1 h at 37°C, and again diluting it 1,000-fold into fresh medium. The plates were then incubated without shaking for 24 h at 37°C.

Bile salts survival assays

Bile salt sensitivity assays were adapted from previous reports[443]. For plate sensitivity assays, each bacterial strain was grown at 37°C until they reached mid-exponential phase of growth (OD_{600} of 0.5) and the culture was serially diluted and spot-titered onto LB agar plates supplemented with either 1% DOC or 1% CHO. Spots were air dried and plates incubated at 37°C for 24 h. For sensitivity assays done in liquid culture, each bacterial strain was grown at 37°C until it reached mid-exponential phase of growth (OD_{600} of 0.5) and then cultures were split and supplemented with either DOC, CHO or buffer (PBS) and bacterial growth was assessed by absorbance at 600nm.

Growth in high-salt media

Bacterial strains were grown in either LB or LB supplemented with 0.3M NaCl until mid-exponential phase and analyzed by phase microscopy at 100x magnification.

1:1 competitive infection of Δ*cpsA* and wild-type

To test the Δ*cpsA* *in vivo* fitness defect, we competed Δ*cpsA* against a Δ*lacZ* mutant in the infant rabbits. To prepare the cultures for infection, 100 µl of overnight cultures of each strain were

inoculated into 100 mL of LB Gm for. After 3 hours of growth at 37°C with shaking, the OD_{600} was measured and 30 units of culture at OD_{600} = 1 were pelleted and resuspend in 10 mL PBS. Then, 5 mL of the Δ*lacZ* culture was combined with 5 mL Δ*cpxA* to make a 1:1 mixture. Dilutions of the inoculum were plated on LB agar plates containing Gm and X-Gal (60 mg/mL) for precise CFU determination and determining the input ratio of Δ*lacZ* (white) to Δ*cpxA* (blue). 4 infant rabbits were inoculated and inoculated and monitored as described above, and colon samples collected. Tissue homogenate was plated on LB agar plates containing Gm and X-Gal (60 mg/mL) and CFU were counted the following day. A competitive index was calculated by dividing the burden of Δ*cpxA* divided by the Δ*lacZ* burden, adjusted for the input ratio.

Δ*cpxA* complementation assays

cpxA was inserted into the expression vector pMMB207[444] (ATCC #37809) downstream of the IPTG-inducible promoter using isothermal assembly. The resulting plasmid, pMMB207-*cpxA*, as well as an empty vector control, were transformed into Δ*cpxA*. To rescue the Δ*cpxA* DOC sensitivity phenotype, we used these strains in a bile salt sensitivity assay as described above. We found that expression from this plasmid was very leaky at basal (non-induced) conditions and could rescue the *cpxA* DOC sensitivity even without addition of IPTG.

Computational analysis

To enable comprehensive functional/pathway analyses in EHEC we carried out BLAST-based comparisons between the old EHEC genome sequence and annotation system[63] (NCBI Accession Numbers AE005174 and AF074613) and the new sequence and annotation system[346] (NZ_CP008957.1 and NZ_CP008958.1). This comparison links the new annotations (RS locus tags) to the original 'Z numbers' and their associated function and pathway annotation.

To make the correspondence table (Table S1) between the old EHEC annotation system (Z Numbers) and the new system (RS Numbers), local nucleotide BLAST with output format 6 was used. First, a reference nucleotide database was generated from the newest EHEC sequence and annotation (NZ_CP008957.1 and NZ_CP008958.1). The EHEC genome sequence containing Z number annotations (AE005174 and AF074613) was used as the query. Of the 6032 EDL933 genes, 5508 have 85% or greater nucleotide identity to a Z Numbered locus. 4796 of these genes match only one Z Numbered locus and 712 genes match multiple Z Numbers. These cases are frequently due to repetitive genomic sequences (such as cryptic phage genes, transposons, or insertion elements) or situations cases where two loci have been merged into one locus. Due to gaps and ambiguity present in the original EDL933 sequence, we did not filter on alignment length in order to find the best matches. In cases where there was one matching locus, the alignment was always ≥90% of the length of the gene. For genes with multiple matches, the length of the alignment varied. There are 524 genes with no corresponding Z Number, presumably loci that are newly recognized genes, and also 71 Z Numbers not assigned to an RS Number, which are loci now recognized as intergenic regions or not present in the final assembly.

To find the K-12 homolog for EHEC genes (Table S2), local BLAST was also used. A reference nucleotide and amino acid database was generated from MG1655 K-12 (NC_000913.3), and the newest EHEC genome sequence was used as the query. For pseudogenes and genes coding for RNA, ≥90% nucleotide identity across ≥90% of the gene length was considered a homolog. For protein coding genes, ≥90% amino acid identity across ≥90% of the amino acid sequence was considered a homolog.

To find KEGG pathways and COG assignments for genes of interest, the Z correspondence table was used to look up the Z number of each gene. The Z number and corresponding functional information was searched on the EHEC KEGG database.

To determine if COGs were enriched in certain groups of genes (such as conditionally depleted genes), a COG enrichment index was calculated[374]. The COG Enrichment Index is the percentage of the genes of a certain category (essential genes or CD genes) assigned to a specific COG divided by the percentage of genes in that COG in the entire genome. A two-tailed Fisher's exact test was used to determine if this ratio was independent of grouping. A Bonferroni correction was applied for multiple hypothesis testing. A p-value of <0.002 was considered to be significant.

Sequencing saturation of TIS libraries was determined by randomly sampling 100,000 reads from each library and identifying the number of unique mutants in that pool. Libraries are sequenced to saturation when no new mutants are identified as additional reads are added. 2–4 million reads are sufficient to capture the depth of libraries used here.

3

Genetic analyses link the conserved inner membrane protein CvpA to the σ^E extracytoplasmic stress response

Overview

The function of *cvpA*, a bacterial gene predicted to encode an inner membrane protein, is largely unknown. Early studies in *E. coli* linked *cvpA* to Colicin V secretion and recent work revealed that it is required for robust intestinal colonization by diverse enteric pathogens. In enterohemorrhagic *E. coli* (EHEC), *cvpA* is required for resistance to the bile salt deoxycholate (DOC). Here, we carried out genome-scale transposon-insertion mutagenesis and spontaneous suppressor analysis to uncover *cvpA*'s genetic interactions and identify common pathways that rescue the sensitivity of a Δ*cvpA* EHEC mutant to DOC. Collectively, these screens led to the hypothesis that the Δ*cvpA* mutant is impaired in its capacity to activate the σ^E-mediated stress response. This idea was supported by showing that mutations that activate σ^E, either indirectly or through its direct overexpression, can restore the Δ*cvpA* mutant's resistance to DOC. Analysis of the distribution of CvpA homologs revealed that this inner membrane protein is conserved across bacterial phyla, in both enteric and non-enteric bacteria that are not exposed to bile. Together, our findings suggest that CvpA may function in triggering activation of the σ^E stress response pathway in response to DOC as well as additional stimuli.

This work is being developed for publication as: Warr AR, Giorgio RT, Waldor MK. Genetic analyses link the conserved inner membrane protein CvpA to the σ^E extracytoplasmic stress response.

Author contributions: The work was conceptualized by ARW and MKW. The investigation and data analysis were performed by ARW and RTG with supervision from MKW. The figures were made by ARW. The writing was completed by ARW with editing from MKW.

Introduction

Enteric pathogens encounter a broad range of host-derived stressors during their transit through the gastrointestinal (GI) tract. These pathogens enter the GI tract through consumption of contaminated food or water, and the diverse challenges they face in the host environment include substantial fluctuations in osmolarity, oxygen concentration, pH, and nutrient availability, as well as mechanical shear force from peristalsis and the bactericidal activities of antimicrobial peptides and bile salts. Pathogens must also compete with the microbiota and face the threat of immune cells and effectors[112,113]. To survive and successfully colonize the host, enteric pathogens must rapidly sense and respond to these stressors.

Bile in particular poses complex threat to enteric pathogens. Bile is an aqueous solution of bile salts, bilirubin, fats, and inorganic salts produced by the liver and secreted into the intestine, where it plays a critical role in digestion of fats[445,446]. Bile salts are potent antimicrobial compounds that can damage diverse components of the bacterial cell. Bile salts are thought to enter the bacterial cell by both passive diffusion and through outer membrane proteins (OMPs) such as porins[114,115]. Enteric pathogens employ a variety of strategies to combat the perturbations caused by bile salts, including utilizing efflux pumps to remove bile salts from the cell cytosol, and expressing chaperones and proteases to ameliorate protein folding stress[189–191].

To coordinate responses to bile, bacteria utilize a host of global stress-response programs. For example, bile-induced DNA damage activates the expression of the SOS regulon, a large suite of genes that assist in repairing and responding to DNA damage[447–449]. Bile also triggers the activation of alternate sigma factors, which enable RNA polymerase (RNAP) to selectively transcribe sets of genes in response to environmental conditions. During normal growth conditions, RNAP associates with the housekeeping sigma subunit, RpoD (σ^{70}). The cellular stresses provoked by bile promote utilization of the alternative stress response sigma factor RpoS (σ^{S}), which governs expression of a

large number (>500) of stress-response genes[450–452]. RpoE (σ^E), a sigma factor which mediates the response to 'extracytoplasmic' stress, such as misfolded proteins in the periplasmic space or disruptions in LPS biosynthesis, may also be important for resistance to the membrane-damaging effects of bile[453,454]. σ^E function intersects with a variety of other cell envelope stress response pathways in gram-negative bacteria such as the two-component systems CpxRA and BaeSR; together, these pathways create a complex regulatory network mediating bacterial envelope homeostasis[453,455].

The gene *cvpA* was identified in screens for genes required for intestinal colonization in several enteric pathogens, including *Vibrio cholerae*, *V. parahaemolyticus*, and *Salmonella* Typhimurium[366,376,427], but the mechanism(s) that account for the colonization defects have not been characterized. We recently found that the protein CvpA is required for the foodborne intestinal pathogen enterohemorrhagic *Escherichia coli* O157:H7 (EHEC) to optimally colonize the colon of infant rabbits; moreover, EHEC *cvpA* deletion mutants were highly sensitive to bile, and in particular the bile salt deoxycholate (DOC)[24]. EHEC causes hemorrhagic colitis and in some cases the severe complication hemolytic uremic syndrome, a condition in which damage to blood vessels of the kidney microvasculature leads to renal failure. Previous studies have led to the idea that EHEC uses bile as a cue to induce expression of virulence genes, flagella, and genes promoting survival *in vivo*, including the AcrAB efflux pump which enables removal of bile salts from the cell interior[449,456–462], but the role of CvpA in contributing to bile resistance in EHEC is unknown.

cvpA was originally described as a gene required for the production of Colicin V (ColV), a small peptide antibiotic produced by some strains of pathogenic *E. coli*[428]. Although the ColV secretion system was characterized[463], the specific role for CvpA in this mechanism was not identified. Here, we provide evidence that CvpA likely provides EHEC resistance to DOC by modulating the σ^E stress response pathway. We also show that CvpA is widely conserved across diverse bacterial phyla

and hypothesize that it may function to coordinate stress responses to perturbations in ion homeostasis.

CvpA localizes to the EHEC cell periphery

CvpA bears structural similarity to the inner membrane mechanosensitive ion channels MscS and MscL (HHPred, Phyre2)[24], and whole-proteome localization analysis in *E. coli* assigned CvpA to the inner membrane[464]. Furthermore, several protein modeling algorithms (PSLPred, HHPred, Phobius, Phyre2, OCTOPUS)[24] predict that CvpA is a 4-pass inner membrane protein with periplasmic C-terminus (Figure 3.1A). To experimentally verify these assignments, we generated two plasmids that inducibly express sfGFP-CvpA fusions, with sfGFP on either the N- or C-terminus, enabling visualization of the localization of CvpA in EHEC. Only the sfGFP fusion to the N-terminus of CvpA yielded a construct that complemented the growth defect of an EHEC strain lacking the *cvpA* gene (Δ*cvpA*) on plates containing DOC, implying this construct contains a functional CvpA protein (Figure 3.1B). Induction of this fusion protein revealed a clear signal that outlined the EHEC cell membrane in a similar pattern as the membrane stain FM4-64 (Figure 3.1CD). These data support the predicted localization of CvpA, and suggest that the availability of its periplasmic C-terminus may be necessary for function.

EHEC *cvpA* mutants do not secrete functional ColV and are sensitive to the bile salt DOC

A previous report suggested that *E. coli* CvpA is required for the export of the peptide antibiotic Colicin V (ColV)[428]. We sought to reproduce this observation in EHEC. The entire ColV-production and immunity operon and endogenous promoters derived from the naturally occurring

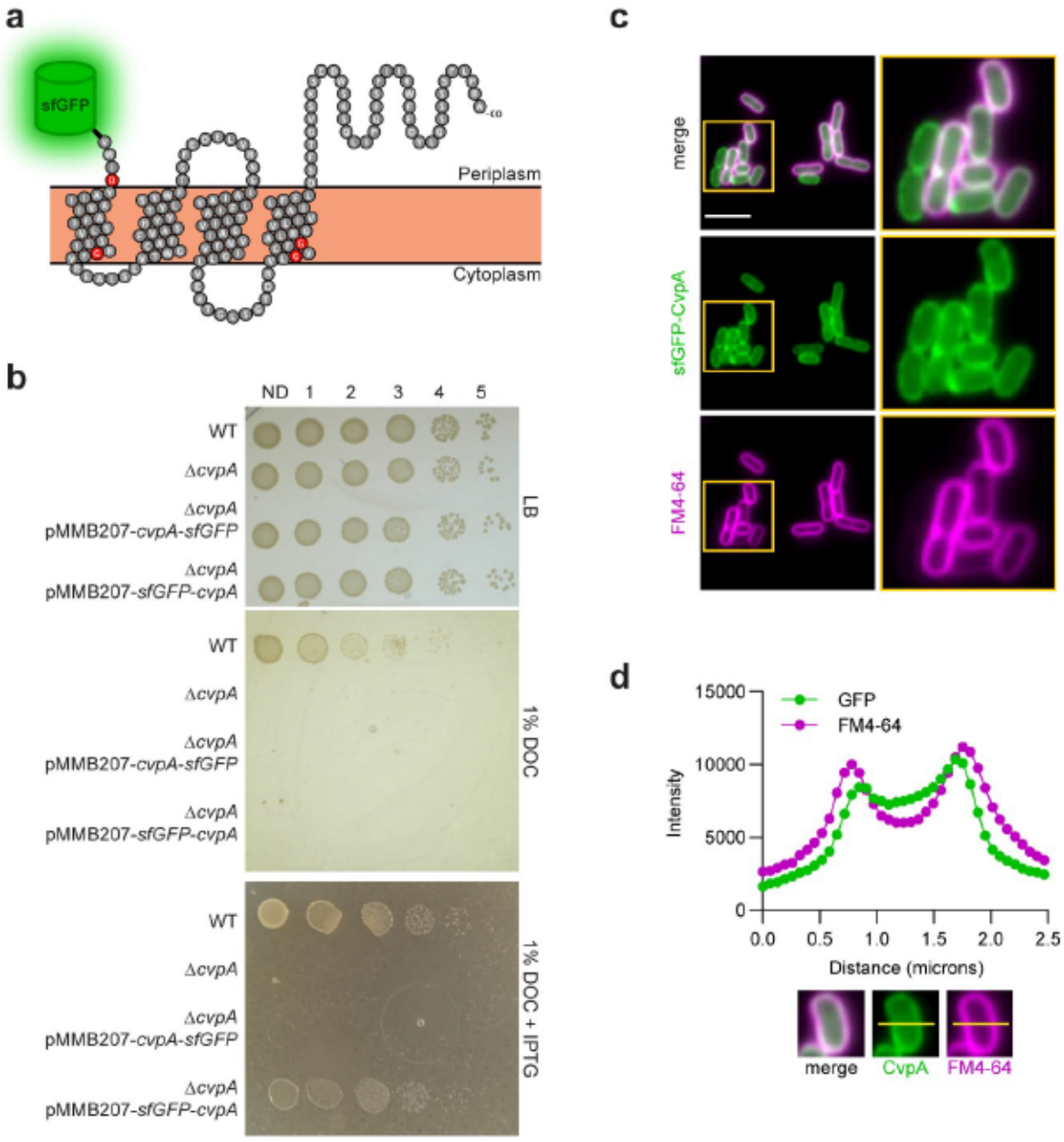


Figure 3.1: CvpA localizes to the EHEC cell periphery.

a) Schematic of the predicted CvpA topology, with location of N-terminal sfGFP fusion shown. Adapted from (22). Red residues are nearly 100% conserved across bacterial phyla. b) Dilution series of WT, Δ*cvpA* mutant, and Δ*cvpA* mutant with an IPTG-inducible *sfGFP-cvpA* fusion complementation plasmid plated on LB, LB 1% deoxycholate (DOC), and LB 1% DOC 1 mM IPTG. c) Micrographs of an N-terminal sfGFP-CvpA fusion protein expressed in EHEC and stained with the FM4-64 membrane dye. Scale bar indicates 5 μM. d) Horizontal line scan of GFP and FM4-64 signal.

ColV plasmid, pColV-K30, was inserted into the cloning vector pBR322, yielding pBR322-ColV. Although both WT (EDL933) and Δ*cvpA* EHEC were transformable with pBR322-ColV, only the WT strain grew robustly in LB broth while carrying pBR322-ColV. In contrast, the Δ*cvpA* mutantelongated and lysed (Figure 3.2A). The abilities of these strains to secrete functional ColV was evaluated by testing whether they create a zone of growth inhibition on a sensitive indicator strain. The WT strain carrying pBR322-ColV killed the indicator strain, creating a distinct zone of clearing (Figure 3.2B), whereas the Δ*cvpA* strain failed to establish such a zone (Figure 3.2B). These data are consistent with previous reports[428] and suggest that Δ*cvpA* mutants are unable to secrete functional ColV and may be sensitive to self-intoxication.

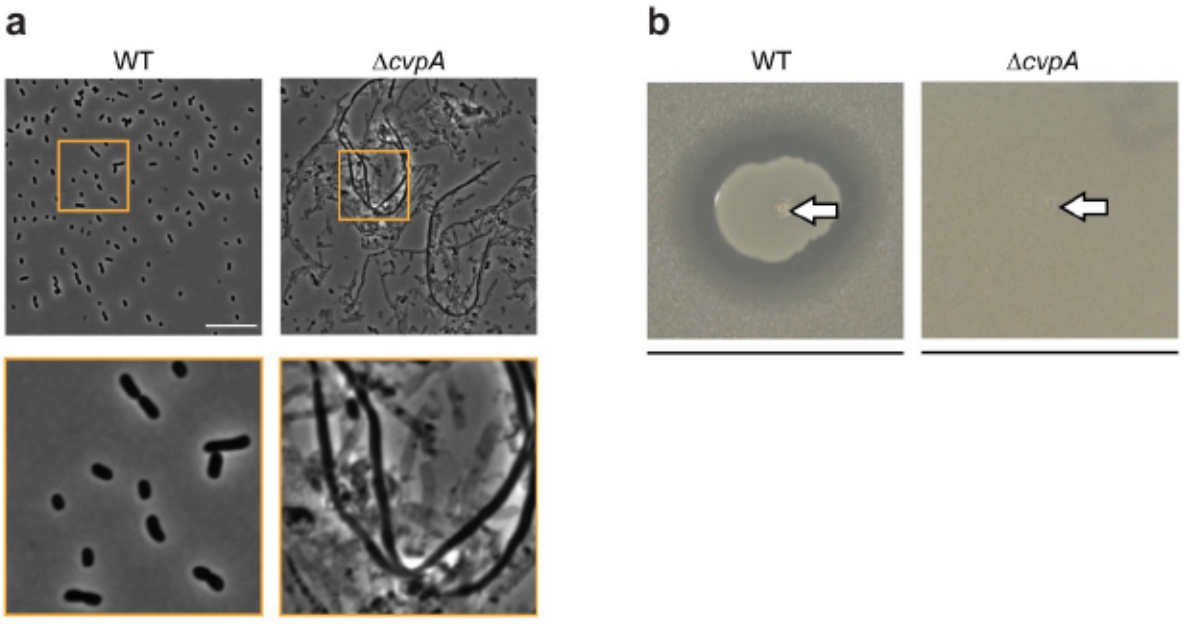


Figure 3.2: Δ*cvpA* EHEC is sensitive to Colicin V production and does not appear to secrete it.

a) Micrograph of overnight culture of indicated strains expressing ColV from plasmid pBR322-ColV. Scale bar is 10 μM. Orange box denotes inset. b) 5 uL jab of overnight culture of indicated strain in soft agar matrix containing 100 μL of indicator (sensitive EHEC) strain. Arrow indicates position of jab. Line represents 20 mm.

Our previous work revealed that CvpA is required for EHEC to robustly colonize the infant rabbit colon, and that Δ*cvpA* mutants are sensitive to the bile salt DOC but not to the bile salt cholate (CHO)[24]. Bile salts like DOC are thought to cause a wide range of deleterious effects in bacteria, including damage to the cell envelope and DNA, and widespread protein misfolding and redox stress[114,115]. The exact mechanisms through which bile salts mediate these diverse deleterious effects are not completely understood[114,115]. We attempted to determine which aspect of DOC toxicity was most damaging to Δ*cvpA* mutants by testing a variety of agents that cause individual aspects of bile stress, including the membrane-damaging detergent SDS, DNA damaging antibiotics, cell wall perturbing antibiotics, protein synthesis perturbing antibiotics, and the redox-stress inducing compounds diamide and hydrogen peroxide[24] (Table 3.1). However, the Δ*cvpA* mutant was not more sensitive than the WT strain to any of these agents. These observations argue against the idea that the sensitivity of the Δ*cvpA* mutant to DOC results from a general cell envelope defect, and instead suggest that DOC toxicity (and potentially the toxicity of ColV accumulation) reflects a more specific defect in cells lacking *cvpA*. Below we moved forward with DOC as a tool to probe the function of CvpA.

Table 3.1: MIC for various compounds in WT and Δ*cvpA* EHEC

Type of stress	Compound (unit)	WT MIC	Δ*cvpA* MIC
Several	DOC (%)	2.5	0.08
Disulfide bond formation	Diamide (mM)	0.625	0.625
Reactive oxygen species	H_2O_2 (mM)	3.2	3.2
Cell wall damage	Ampicillin (μg/mL)	7.8	7.8
Cell wall damage	Piperacillin (μg/mL)	0.31	0.31
Detergent	SDS (%)	5	5
Detergent	Polymyxin B (μg/mL)	1.17	1.17
DNA damage	Nalidixic acid (μg/mL)	2.5	2.5

Genetic screens link CvpA to the σ^E response

To further assess CvpA's function, we carried out three genetic screens in EHEC to define *cvpA's* network of genetic interactions. Two of the screens relied on transposon-insertion mutagenesis, and one used spontaneous suppressor analysis. Similar genetic approaches have been fruitful for other membrane proteins[465–468], which are often difficult to probe using biochemical approaches to identify protein-protein interactions.

For the transposon screens, we began by generating high density transposon-insertion mutant libraries in the WT and Δ*cvpA* EHEC strains by mutagenizing the strains with the mariner transposon Himar1, which inserts randomly at TA dinucleotides. The abundance of each transposon-insertion mutant in the two libraries was quantified by deep sequencing the sites of transposon insertion. Genes were binned by the percentage of TA sites disrupted (Figure 3.3A), and as expected for libraries of high insertion density, this distribution was bimodal. The left-most minor peak represents genes with few to no insertions, and the right-most major peak represents genes with a majority of TA sites disrupted[367]. The distribution of transposon insertions in the WT and Δ*cvpA* libraries was similar and displayed complexity enabling high-resolution analysis of transposon-insertion frequency[366] (Figure 3.3A).

We first completed a synthetic screen in which we compared the distributions of the transposon insertions across the genome in the WT and Δ*cvpA* libraries. This comparison yields three categories of genes: (1) loci with similar insertion profiles in both backgrounds are deemed 'neutral'; (2) loci with fewer insertions in the Δ*cvpA* vs the WT background are considered 'synthetic sick'; and (3) loci with more insertions in the Δ*cvpA* compared to the WT background are deemed 'synthetic healthy'. This approach has been used to identify genetic interactions in several other organisms[469–472]. We identified 32 candidate synthetic healthy genes, for which transposon-insertion mutants were overrepresented at least 4-fold in the Δ*cvpA* mutant background compared to the WT (Figure 3.3B,

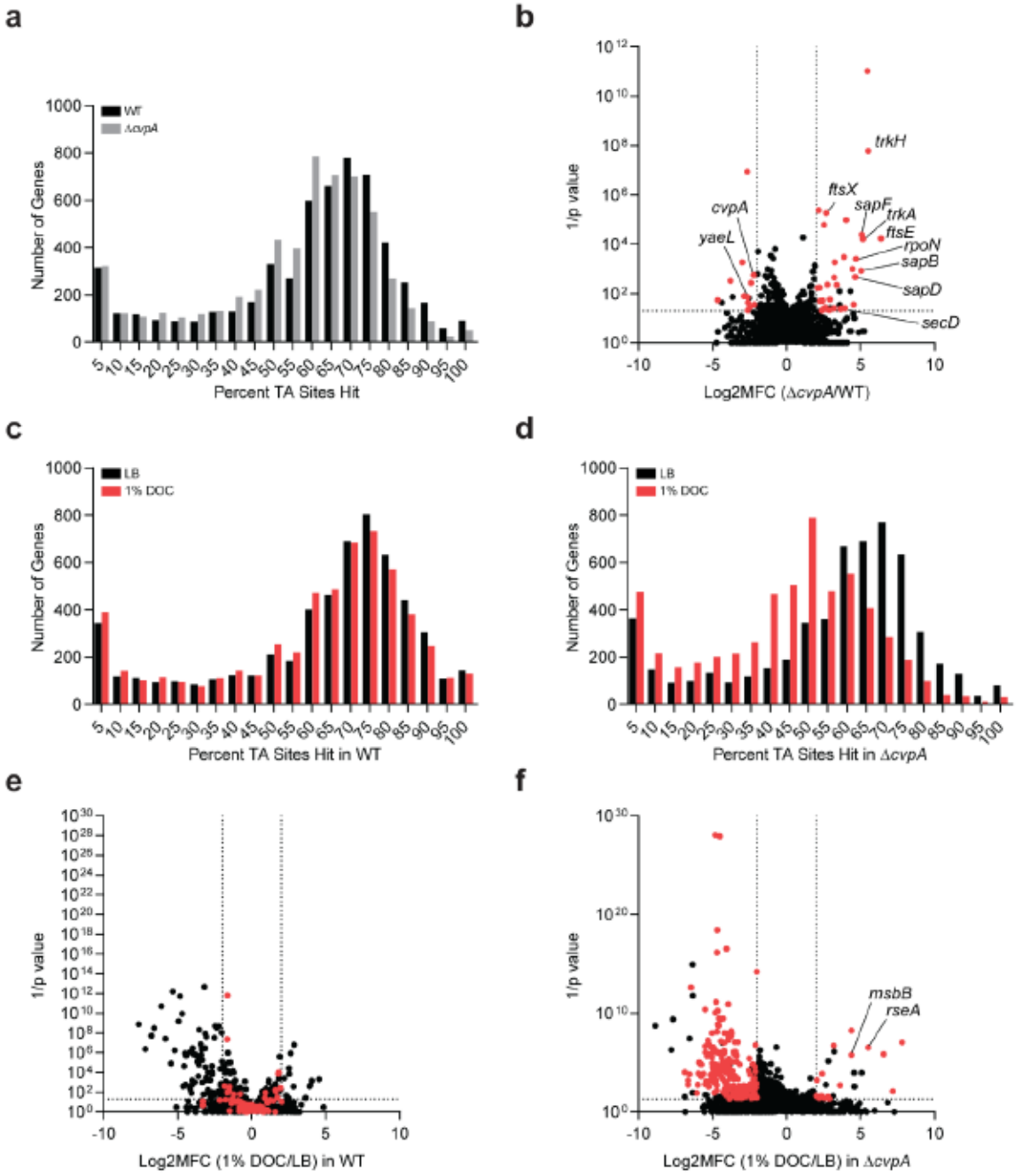


Figure 3.3: Transposon-insertion sequencing screens reveal CvpA's genetic interactions.

a) Distribution of the percentage of TA sites disrupted for all genes in the library constructed in the WT (black bars) and Δ*cvpA* (gray bars) backgrounds. b) Volcano plot depicting the relative abundance of read counts mapped to individual genes in transposon libraries in WT and Δ*cvpA* mutant backgrounds. The mean log2 fold change (x axis) and Mann-Whitney U p-value (y axis) are shown for each gene. Genes highlighted in red have greater than 5 unique transposon mutants, a $\log_2$FC>2 or <-2, and p-value < 0.05. c-d) Distribution of the percentage of TA sites disrupted for all genes in the library constructed in the WT (c) or Δ*cvpA* background (d) after growth in LB (black bars) or 1% DOC (red bars). e-f) Volcano plot depicting the relative abundance of read counts mapped to individual genes in transposon libraries in WT (e) or Δ*cvpA* background (f) in LB vs in 1% DOC. The log2 mean fold change (x axis) and Mann-Whitney U p-value (y axis) are shown for each gene. Genes highlighted in red are those for which mean log2FC >2 or <-2 and p-value < 0.05 in the Δ*cvpA* mutant background but not the WT background.

Table S10). Several of these genes were components of ion transport systems, including *trkH* and *trkA*, which are components of the cellular potassium uptake system, and three genes of the *sap* transporter complex (*sapB, sapD,* and *sapF).* SapD is also known to interact with the Trk potassium transport system[473]. Several chaperones, including *secB*, the alternative sigma factor *rpoN*, and several cell shape/septal structure genes, including *ftsEX,* were also on the list of synthetic healthy genes. 14 candidate synthetic sick genes, which were underrepresented at least 4-fold in the Δ*vpA* mutant background compared to the WT (Figure 3.3B, Table S1), were also identified. These genes included *yaeL*, a protease that works with DegS to activate σ^E by degrading the anti-sigma factor RseA. Collectively, these results suggest that inactivating certain ion transport processes bolster the fitness of the Δ*vpA* strain, while impairing the cell's ability to activate the σ^E -mediated stress response is particularly deleterious in the Δ*vpA* background. Interestingly, there is evidence that cellular potassium levels can regulate promoter selectivity of sigma factors[474].

A second transposon screen was designed to identify insertion mutations in genes that suppress the DOC sensitivity phenotype in the Δ*vpA* mutant. The WT and Δ*vpA* transposon-insertion mutant libraries were screened on plates containing either LB or LB supplemented with 1% DOC. As above, after enumerating the abundance of transposon-insertion mutants with sequencing, genes were binned according to the percentage of disrupted TA sites (Figure 3.3CD). The distribution of disrupted TA sites per gene in the WT library was not markedly different between the LB and DOC selection, as expected because a majority of mutants do not have a phenotype in DOC (Figure 3.3C) (Figure 3.3C). However, the distribution in the Δ*vpA* library was different between the two conditions; there was a leftward shift in the distribution of insertions after the Δ*vpA* library was exposed to DOC compared to LB (Figure 3.3D). This leftward shift is indicative of a bottleneck, a stochastic, genotype-independent constriction of population size, and reflects the profound DOC sensitivity of the Δ*vpA* mutant. To identify mutants in both libraries which have markedly better or

worse phenotypes than expected in DOC, we used Con-ARTIST, an analytical pipeline that mitigates the effects of bottlenecks to identify differentially abundant mutants[366]. For each library, the abundance of each mutant was compared between the LB and DOC condition to identify three categories of genes: (1) loci with similar insertion profiles in both LB and DOC are 'neutral'; (2) loci with fewer insertions in the DOC condition than expected when compared to LB are 'depleted'; and (3) loci with more insertions in the DOC condition than expected when compared to LB are 'enriched' (Figure 3.3EF). In the Δ*cpxA* background, we identified 233 'depleted' genes in which mutations are expected to increase baseline DOC sensitivity, and 26 'enriched' genes in which mutations are expected to suppress DOC sensitivity (Figure 3.3F). 19 of these 26 genes did not show differential abundance in DOC vs. LB comparison in the WT background and were classified as 'neutral' (Figure 3.3EF). These genes are of particular interest as candidate suppressors, as they have *cpxA*-specific DOC phenotypes. Many of these genes were linked to stress response pathways and cell envelope integrity, such as the anti-sigma factor *rseA*, the σ^E -effector chaperone *skp,* the two-component system *basSR*, the redox genes *sodC* and *cyoC,* the disulfide/redox bond formation genes *dsbA* and *dsbB*, and the cell membrane synthesis genes *msbB, wcaM,* and *rfaC*. Mutations in several of these genes are known to increase the level of σ^E activity[475–477].

In addition to transposon-based screens, we also characterized spontaneous suppressors of DOC sensitivity in the Δ*cpxA* mutant. In this screen, 2 mL of OD 0.5 Δ*cpxA* broth culture were plated on large plates and 16 putative suppressor colonies were picked after 24 hours. After re-patching on 1% DOC, which confirmed their resistance, whole-genome sequencing of the suppressors was carried out and these genomes sequences were compared to that of the Δ*cpxA* parent strain to identify putative suppressor mutations. SNPs or deletions that were not present in Δ*cpxA* parent strain were found in 14 of the 16 strains. Notably 12 of the 14 strains had mutations in genes linked to LPS biogenesis, including several genes in the lipopolysaccharide transport (LPT)

system (*lptB, lptE, lptF, lptG*) and an enzyme required for lipid A synthesis, *msbB;* genes related to redox including *nuoH* and *cyoDC* were also identified (Table 3.2). Interestingly, *msbB* was also identified as a candidate suppressor in the transposon DOC screen above. LPT genes are essential, and thus not highly represented in our transposon libraries. Although not statistically robust, we analyzed the transposon DOC screen data for the LPT genes with at least 1 insertion mutant and found that *lptCDFG* had large, positive fold-changes in 1% DOC vs LB in the Δ*cpxA* background and not WT (Table S11), supporting the findings of the spontaneous suppressor screen.

Table 3.2: Mutations in suppressors of Δ*cpxA* DOC sensitivity[a]

Suppressor	Mutation	Locus	Gene	Function
1	SNP	RS00350	*araC*	Arabinose operon transcriptional regulator
	SNP	RS02475	*yajR*	Putative transport protein
	SNP	RS13840	*msbB*	Lipid A biosynthesis acyltransferase
2	SNP	RS16885	*nuoH*	NADH-quinone oxidoreductase subunit
	SNP	RS27795	*lptG*	LPS export permease
3	SNP	RS16885	*nuoH*	NADH-quinone oxidoreductase subunit
	SNP	RS21750	*lptB*	LPS export ATP-binding protein
4	SNP	RS13840	*msbB*	Lipid A biosynthesis acyltransferase
5	SNP	RS21750	*lptB*	LPS export ATP-binding protein
6	SNP	RS27795	*lptG*	LPS export permease
7	SNP	RS27795	*lptG*	LPS export permease
8	SNP	Intergenic	upstream *lptE*	LPS assembly lipoprotein
	DEL	RS02485	*cyoD*	Cytochrome bo subunit IV
9	DEL	RS13840	*msbB*	Lipid A biosynthesis acyltransferase
10	SNP	RS16885	*nuoH*	NADH-quinone oxidoreductase subunit
	SNP	RS27795	*lptG*	LPS export permease
11	SNP	RS27790	*lptF*	LPS export permease
14	SNP	RS12645	*dtpA*	Proton-dependent di/tripeptide symporter
	SNP	RS02495	*cyoB*	Cytochrome bo3 subunit 1
15	SNP	RS13840	*msbB*	Lipid A biosynthesis acyltransferase
	SNP	RS16885	*nuoH*	NADH-quinone oxidoreductase subunit
16	SNP	RS12645	*dtpA*	Di/tri-peptide permease
	SNP	RS02495	*cyoB*	Cytochrome bo3 subunit 1

[a] SNP indicates a single-nucleotide polymorphism, DEL indicates a larger deletion mutation.

Increased σ^E activity rescues the sensitivity of the Δ*cvpA* mutant to DOC

Several independent pieces of evidence connecting *cvpA* to the σ^E response were found in the 3 genetic screens described above. σ^E is normally sequestered at the inner membrane by association with an anti-sigma factor, RseA. When the appropriate stimuli is detected, proteases YaeL and DegS sequentially cleave RseA, liberating σ^E to activate expression of ~100 genes in response[478,479]. A variety of perturbations to the cell envelope can trigger this pathway, including misfolded OMPs, disruptions in LPS biosynthesis leading to precursor accumulation, and changes in cellular redox state[480,475–477]. Notably, the genetic screens revealed that the Δ*cvpA* mutant is more sensitive than the WT to mutations in *yaeL* and the suppressor screens uncovered a variety of mutations, including in *rseA*, *lptBCDEFG*, and *msbB*, which likely trigger the σ^E response, that appeared to rescue the sensitivity of Δ*cvpA* to DOC. Inactivating mutations in *rseA* increase the basal activity of σ^E by preventing its membrane sequestration, and mutations in LPS biosynthesis genes *msbB* and the LPS transport system trigger σ^E through accumulation of LPS precursors[475,476].

To test the linkage between *cvpA* and σ^E, deletion mutants of *rseA and msbB* in the WT and Δ*cvpA* backgrounds were created. These deletions both restored the resistance of the Δ*cvpA* mutant to 1% DOC to WT levels (Figure 3.4A). Similarly, direct overexpression of *rpoE* from a plasmid rescued the sensitivity of the Δ*cvpA* mutant to DOC (Figure 3.4B). These data suggest that mutations that activate σ^E, either indirectly or through its direct overexpression, can restore the Δ*cvpA* mutant's resistance to DOC, implying that the mutant may be otherwise unable to activate σ^E in response to DOC.

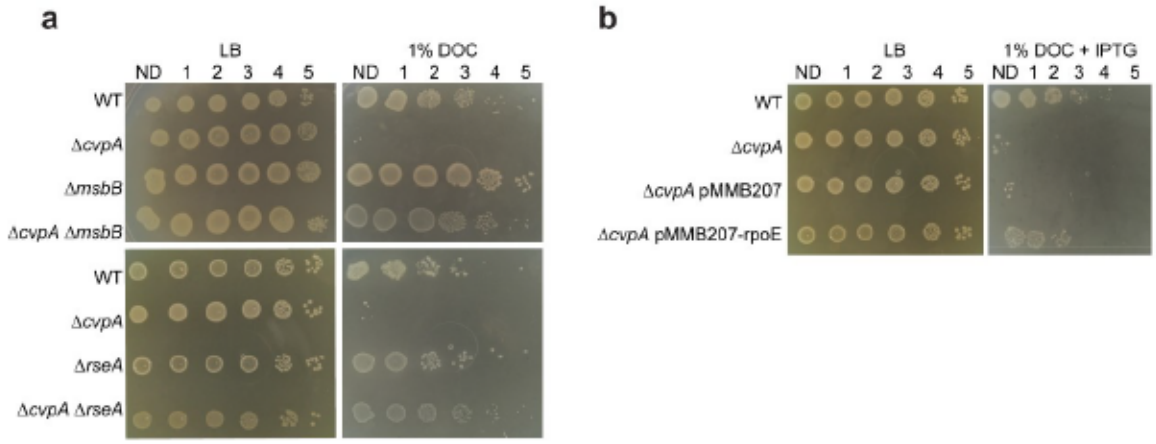


Figure 3.4: Activation of RpoE rescues the Δ*cvpA* mutant's sensitivity to DOC.

a) Dilution series of WT, Δ*cvpA*, Δ*cvpA* Δ*msbB* and Δ*cvpA* Δ*rseA* plated on LB or LB 1% deoxycholate (DOC). b) Dilution series of WT, Δ*cvpA*, Δ*cvpA* transformed with empty expression vector pMMB207, and Δ*cvpA* transformed with pMMB207 carrying IPTG inducible *rpoE* plated on LB and LB 1% DOC with 0.1 mM IPTG.

CvpA is highly conserved across bacterial phyla

The phylogeny tool Annotree[481] was used to analyze the distribution of CvpA homologs across more than 27,000 bacterial genomes. CvpA homologs (≥30% amino acid identity, ≥70% alignment) were found in 9,830 genomes distributed across 55 phyla (Figure 3.5A, Table S12). Homologs are present in the majority of represented genomes in the Campylobacterota (236/254, 93%), Deferribacterota (13/14, 93%), and Proteobacteria (6923/7630, 91%) phyla (Figure 3.5B). In contrast, only 2 of 3118 strains in Actinobacteria have a CvpA homolog (0.06%) (Figure 3.5B).

Sequence alignment of the primary amino acid sequences of 20 CvpA homologs from bacterial species from diverse phyla revealed several highly conserved residues, including an aspartate and three glycines with nearly 100% conservation (Figure 3.6). The motif GXXXG, common in transmembrane segments[482] is also present. All 20 CvpA homologs are predicted to localize to the inner membrane by PSLPred. All homologs share greater similarity in alignments of their respective

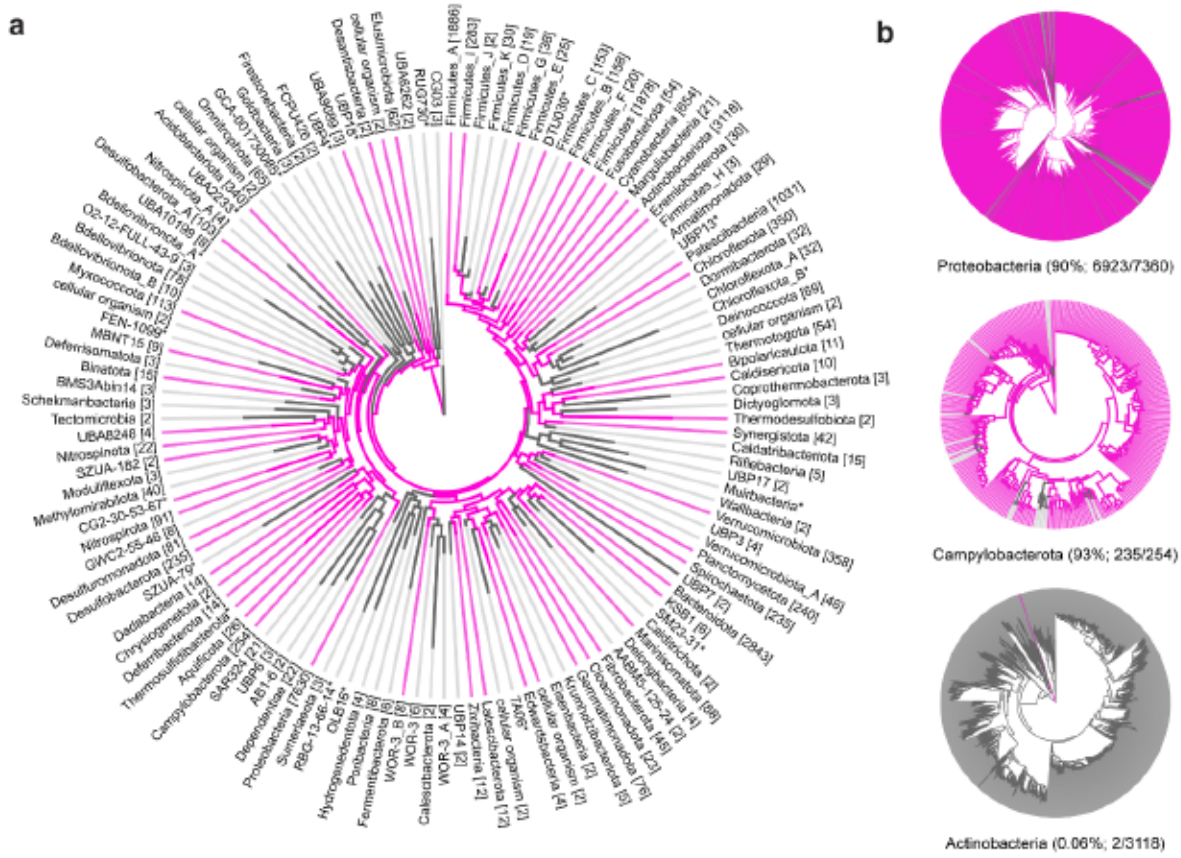


Figure 3.5: CvpA is highly conserved across many bacterial phyla.

a) Phylogeny of CvpA distribution across bacterial phyla. Pink lines represent phyla where at least one species has a CvpA homolog of at least 30% amino acid identity and at least 70% alignment to *E. coli* CvpA. The numbers shown in brackets denote the number of individual genomes in that phylum. b) Right, phyla with varying degrees of CvpA conservation at the species level. Pink lines represent species that contain a CvpA homolog. Numbers in parenthesis after phylum name indicate the percentage of species in that phylum with a CvpA homolog.

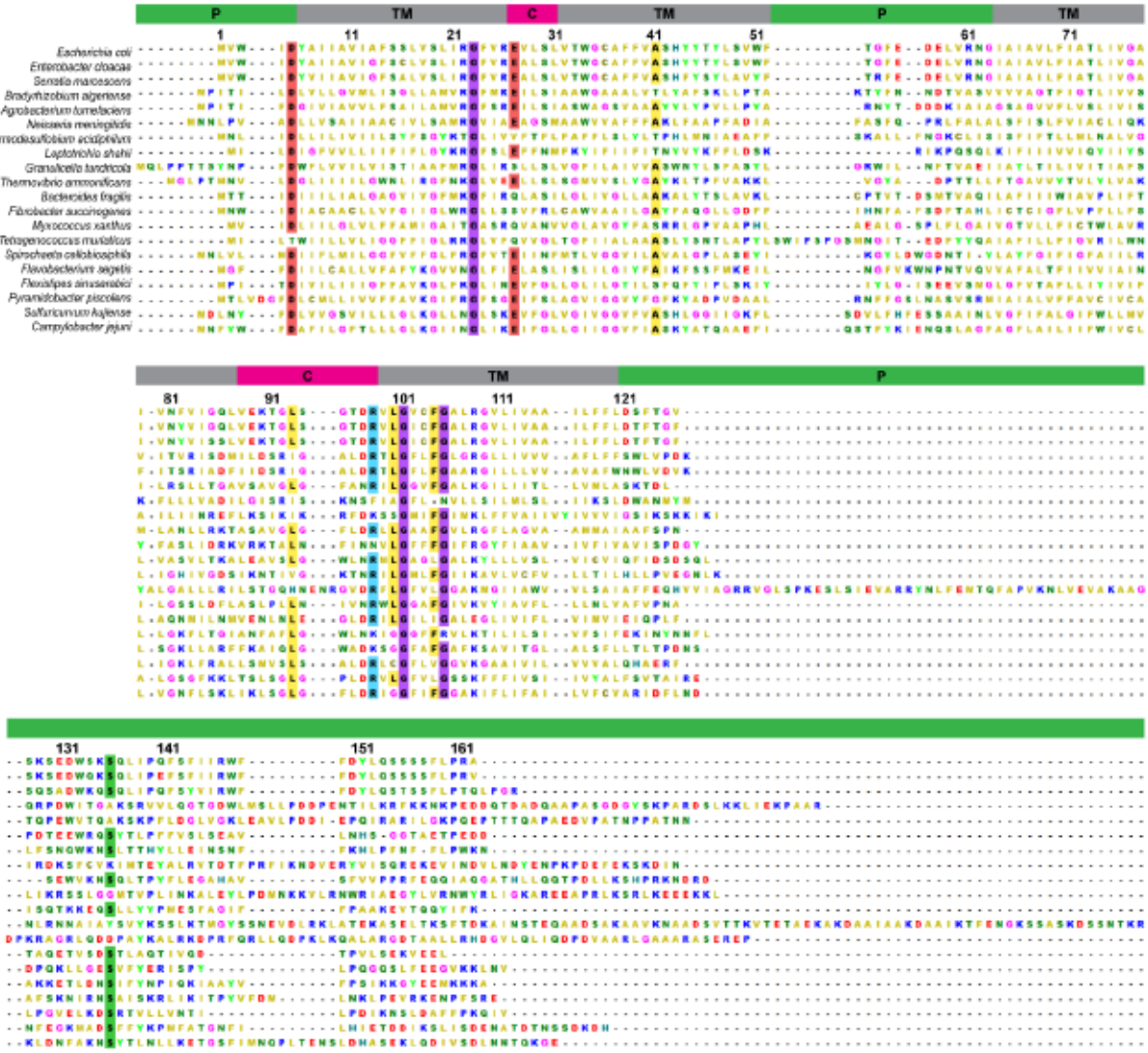


Figure 3.6: MUSCLE alignment of 20 CvpA homologs from bacterial species across diverse phyla.

Residues with >70% conservation across species are indicated by a colored box. Residue numbering and predicted membrane topology for the *E. coli* CvpA is shown above the alignment. The green box indicates predicted periplasmic (P) segments, the gray box indicates transmembrane (TM) segments, and the pink box represents cytoplasmic (C) segments.

N-terminal regions, and have more variable lengths in their final periplasmic segments, which could reflect different functions or binding partners at the C-terminus.

This conservation across diverse phyla highlights that the function of CvpA is not strictly related to bile. While homologs are present in enteric pathogenic and commensal strains, such as *Campylobacter jejuni, Yersinia enterolitica, Fibrobacter succinogenes*, and *Bacteroides fragilis,* many strains with a CvpA homolog inhabit non-intestinal niches. These strains include non-enteric pathogens such as *Neisseria meningitidis,* plant symbionts such as *Rhizobium tropici,* and non-pathogenic environmental isolates from diverse environments such as tundra soil (*Granulicella tundricola*), hydrothermal vents (*Thermovibrio ammonificans*), and marine environments (*Spirochaeta cellobiosiphila*).

Discussion

Since the original report describing *cvpA* as a gene required for production of the plasmid-borne ColV peptide from *E. coli* more than 30 years ago[428], very little progress has been made determining the function of this widely conserved inner membrane protein. More recently, *cvpA* has been found to be important for intestinal colonization of *Vibrio cholerae*, *V. parahaemolyticus*, and *Salmonella* Typhimurium and EHEC in studies based on *in vivo* transposon screens[24,366,376,427]. In EHEC, we found that a Δ*cvpA* mutant was highly sensitive to DOC[24], a bile component, thus presenting a possible explanation for *cvpA*'s role in the colonization of distinct niches in the small and large intestine by diverse enteric pathogens. Here, we used genetic screens to further our understanding of *cvpA* function. These screens led to the hypothesis that *cvpA* is an upstream component of the σ^E 'extracytoplasmic' stress response pathway that responds to DOC and other stimuli, such as ColV secretion.

A variety of stimuli can lead to the activation of the σ^E response. Misfolded outer membrane proteins, disruptions in LPS biogenesis, and perturbations in redox state have all been linked to

increased σ^E activity[480,475–477]. Mutations in the regulators of these pathways, for example inactivating *rseA*, the anti-sigma factor which controls σ^E activity[480], *msbB*, an enzyme required for the biosynthesis of lipid A[475], or the redox enzymes *cyoC*, *sodC*, and *dsbAB*[477] can also trigger the activation of σ^E in the absence of external stimuli. Conversely, inactivating mutations in *degS* and *yaeL*, the proteases required to release sequestered σ^E from the membrane, inhibit σ^E activation[483–485]. Nearly all of the σ^E-interacting genes listed above were identified as potential modulators of the fitness Δ*cvpA* mutant in DOC. These data suggest that in response to DOC, CvpA may be required to activate the σ^E response. In a Δ*cvpA* mutant, this pathway is unable to respond to deleterious stimuli, and additional σ^E-activating mutations (deletions in *rseA*, *msbB*, *lpt* genes, direct overexpression of *rpoE*) are required to bypass this block. Mutations in *cyoC*, *sodC*, and *dsbAB* were unable to suppress the sensitivity of Δ*cvpA* DOC mutations, despite these loci having strong phenotypes in the DOC transposon-insertion screen. It is possible that under the conditions of our *in vitro* validation experiments, deletions in these enzymes are not sufficient to rescue DOC sensitivity.

The conservation of CvpA across diverse bacterial phyla, including in species which are never exposed to bile, suggests that CvpA function is not restricted to responding to bile stress. Bile can impair several processes in the bacterial cell, including inflicting damage to the cell envelope and DNA, and generating protein folding stress and alterations in redox state[114,115]. We were unable to isolate which individual aspect of DOC stress kills Δ*cvpA* mutants, raising the possibility that either these deleterious effects need to occur together, or that bile impairs additional processes that were not tested. One possibility is that bile, and DOC in particular, perturbs ion homeostasis, which could in turn disrupt membrane potential. The increased fitness of insertions in several genes linked to K^+ transport links *cvpA* and ion transport processes and supports this idea. Furthermore, in eukaryotic cells, DOC exposure leads to changes in mitochondrial membrane potential[486]. Interestingly, Δ*cvpA*

mutants are not sensitive to the very similar bile salt cholate (CHO)[24], which may reflect a difference in the ability of these salts to move through membranes[487] or the increased hydrophobicity of DOC[443].

Given the similarity of CvpA to the mechanosensitive solute transporters MscS and MscL and the numerous genes related to ion transport that were identified in the synthetic lethal screen, it is tempting to speculate that CvpA is itself an ion channel that monitors and promotes membrane potential homeostasis. Notably, Colicin V mediates its bactericidal action by disrupting the membrane potential of the inner membrane[488,489]; the heightened sensitivity of Δ*cvpA* mutants to perturbations of membrane potential may explain why they cannot produce functional Colicin V. Maintenance of ion homeostasis and membrane potential is a critical capability for all bacterial species. DOC may represent a stimulus that leads to perturbation of ion levels; other stimuli are likely encountered in the disparate niches inhabited by the diverse bacterial species encoding CvpA homologs.

Finally, our observations suggest that the C-terminus of CvpA is required for its function. A fusion of sfGFP to this end of the protein was not functional. The C-terminal region of CvpA varies in length between species. Albeit speculative, we propose that this region of the protein may be critical for binding to periplasmic partner(s) or sensing/responding to periplasmic stimuli. Clearly, additional biochemical studies to elucidate the mechanistic basis of CvpA function are warranted.

Materials and Methods

Bacterial strains and growth conditions

Bacterial strains were cultured in LB medium or on LB agar plates at 37°C. Antibiotics and supplements were used at the following concentrations: 20 μg/mL chloramphenicol (Cm), 50 μg/mL kanamycin (Km), 10 μg/mL gentamicin (Gm), 50 μg/mL carbenicillin (Cb), 0.3 mM

diaminopimelic acid (DAP), 1% deoxycholate (DOC), 0.1 mM Isopropyl β-d-1-thiogalactopyranoside (IPTG).

A gentamicin-resistant mutant of *E. coli* O157:H7 strain EDL933 (Δ*lacI*::*aacC1*)[24] was used in all experiments in this study as the wild-type (WT). All mutants were constructed in this strain background using standard allelic exchange techniques[433] with the pTOX plasmid system[490] or lambda-red recombineering[378].

Colicin V production and immunity operons and endogenous promoters were cloned from plasmid pColV-K30:Tn10 (gift from Dr. Roberto Kolter) into pBR322 (yielding 'pBR322-ColV') using isothermal assembly. The resulting plasmids were transformed into WT or Δ*cvpA strains.*

To construct inducible N- or C-terminal sfGFP fusions to CvpA, the sfGFP sequence was amplified from pBad-sfGFP (Addgene plasmid #85482) and linked to the *cvpA* sequence amplified from EHEC genomic DNA with the linker sequence gcagcggccggcggaggg through isothermal assembly into pMMB207[444] (ATCC #37809) multiple cloning site (MCS). Similarly, to construct the *rpoE* expression plasmid, the *rpoE* sequence was amplified from genomic DNA and cloned into the pMMB207 MCS. The resulting plasmids were transformed into WT or Δ*cvpA* strains by electroporation.

ColV production assay

pBR322-ColV was transformed into WT or Δ*cvpA* strains. Individual colonies of either strain were selected and grown overnight in LB-Carb, and then normalized to OD 0.5. To prepare indicator plates, 100 µL of WT EHEC was added to 5 mL of 0.8% molten LB agar, vortexed to mix, then poured quickly onto a pre-warmed LB plates, which were allowed to solidify undisturbed. 5 *µ*L of OD 0.5 ColV producing strain (WT or Δ*cvpA* transformed with pBR322-ColV) was jabbed into soft

agar plate and incubated upright overnight at 37°C. The next day, plates were examined for zones of clearing around the ColV producing strain.

DOC sensitivity assay

To determine sensitivity to DOC, each strain was grown at 37°C until mid-exponential phase (OD_{600} = 0.5). For inducible expression, at mid-exponential phase, IPTG was added to a final concentration of 1 mM and cells were incubated at 37°C shaking for 1 hour. Cultures were normalized to OD_{600} = 0.5, serially diluted and plated onto LB agar plates with and without 1% DOC and 0.1 mM IPTG.

Microscopy

To prepare sfGFP-CvpA fusion cells for microscopy, strains were grown to mid-exponential phase (OD_{600} = 0.5). IPTG was then added to a final concentration of 1 mM. After one hour, cells were stained with the membrane dye FM4-64 as described previously[469,491]. Briefly, 1 µg/mL FM4-64 was added to 100 µL of culture and incubated at room temperature for 5 minutes. To prepare pColV-containing cells for microscopy, overnight culture was first normalized to OD_{600} = 0.5. 3 µL culture were then immobilized on 0.8% agarose pads and imaged with a Nikon Ti Eclipse microscope equipped with a widefield Andor NeoZyla camera and a 100x oil phase 3 1.4 numerical aperture objective. Images were processed in ImageJ/FIJI[492].

Minimum inhibitory concentration (MIC)

To determine MIC of various compounds listed in Table 1, an assay was performed as described previously[24,442]. Briefly, compounds to be tested were prepared in serial 2-fold dilutions in LB (50 µL total volume) in a 96 well plate. Overnight cultures of bacterial strains were diluted 1:1000 in LB,

grown for 1 hour at 37˚C, and then diluted again 1:1000 in LB. 50 µL of this diluted culture was added to each well. Plates were incubated statically for 24 hours at 37˚C.

Transposon-insertion library construction

TIS libraries were generated in EHEC EDL933 Δ*lacI*::*aacC1 and* Δ*cpxA* mutant as described previously[24]. Briefly, conjugation was performed to transfer the transposon-containing suicide vector pSC189[437] from a DAP-auxotrophic donor strain (*E. coli* MFDλpir) into the recipient strain. 600 µL of overnight culture of donor and recipient were pelleted, washed with 1 mL LB, and resuspended each in 60 µL LB. The cultures were mixed and spotted onto a 0.45 µM HA filter (Milipore) on an LB-DAP agar plant and incubated at 37˚C for 1 hour. The filters were then washed in 24 mL LB and immediately spread across three 245x245 mm^2 (Corning) LB-agar plates containing Gm and Km and incubated overnight at 37˚C. Plates were scraped to collect colonies, which were resuspended in LB and stored in LB + 20% glycerol (v/v) 1 mL aliquots at -80˚C. An aliquot was thawed and gDNA isolated for analysis.

DOC-selected TIS library

1 aliquot of WT and Δ*cpxA* library were thawed and diluted to OD_{600} = 1 with LB. 5 mL of this diluted culture was added to a flask with 75 mL of LB. The culture was grown shaking at 37˚C for 1 hour, at which their OD_{600} measured ~0.25. 5 mL of the culture was plated on each of two 245x245 mm^2 (Corning) LB-agar plates containing Gm and Km and on each of two plates containing Gm and Km and 1% deoxycholate (DOC). Plates were incubated overnight (~16 hours) at 37˚C and immediately scraped to collect colonies. Colonies were resuspended in LB and stored in LB + 20% glycerol (v/v) 1 mL aliquots at -80C.

Characterization of transposon-insertion libraries

Transposon-insertion libraries were characterized as described previously [24]. Briefly, for each library, gDNA was isolated using the Wizard Genomic DNA extraction kit (Promega). gDNA was fragmented to 400-600 bp using a Covaris E200 sonicator and end-repaired using the NEB Quick Blunting Kit. PCR was used to amplify transposon junctions, and PCR products were gel purified to recover 200-500 bp fragments. To estimate library concentration, purified PCR products were subjected to qPCR using primers designed to the Illumina P5 and P7 hybridization sequences. Libraries were mixed in an equimolar fashion and sequenced with a MiSeq.

Reads were trimmed of transposon and adapter sequences using CLC Genomics Workbench (Qiagen) and mapped to *E. coli* O157:H7 EDL933 (NCBI Accession Numbers: chromosome, NZ_CP008957.1; pO157 plasmid, NZ_CP008958.1) using Bowtie, allowing for no mismatches. Reads were discarded if they did not map to any sites in the genome, and reads mapping to multiple sites were randomly distributed. The data was normalized for chromosomal replication biases and differences in sequence depth using a LOESS correction of 100,000 bp (chromosome) and 10,000 bp (plasmid) windows. Reads at each TA site were tallied and binned by protein-coding gene.

For identification of loci with synthetic fitness between WT and Δ*cpxA mutant,* the Con-ARTIST pipeline was used [24,366]. First, the WT library normalized to simulate any bottlenecks or differences in sequencing depth as observed in the Δ*cpxA* library using multinomial distribution-based random sampling (n=100). Next, a Mann-Whitney U test was applied to compare these 100 simulated data sets to the Δ*cpxA* mutant library. Genes for which there were greater than 5 individual transposon-insertion mutants were considered to have sufficient data for analysis (otherwise 'insufficient data'). Genes for which mutant abundance in the Δ*cpxA* background was mean log2FC ≥ 2 compared to the WT with a p-value <0.05 were considered to be 'synthetic healthy', while genes with mutant

abundance log2FC ≤ -2 compared to the WT with a p-value < 0.05 were considered to be 'synthetic sick'. The remaining genes were classified as 'neutral'.

A similar analysis was performed to identify mutants conditionally depleted or enriched in 1% DOC compared to LB for each strain. First, for each strain, the LB library was normalized to simulate any bottlenecks or differences in sequencing depth as observed in the corresponding 1% DOC library using multinomial distribution-based random sampling (n=100). Next, a Mann-Whitney U test was applied to compare these 100 simulated data sets to the 1% DOC mutant library. Genes for which there were greater than 5 individual transposon-insertion mutants were considered to have sufficient data for analysis (otherwise 'insufficient data'). Genes for which mutant abundance in the 1% DOC background was mean log2FC ≥ 2 compared to the WT with a p-value < 0.05 were considered to be 'enriched', while genes with mutant abundance log2FC ≤ -2 compared to the WT with a p-value < 0.05 were considered to be 'depleted'. The remaining genes were classified as 'neutral'.

Spontaneous suppressor isolation and sequencing

To isolate and identify spontaneous suppressor mutations, the Δ*cpxA* mutant was first grown to mid-exponential phase (OD_{600} = 0.5). 2 mL of this culture was plated on 150x15mm (Fisher) LB-agar plate containing 1% DOC and incubated for 24 hours at 37°C. The next day, suppressor colonies were picked and re-streaked on LB agar plates containing 1% DOC. 15 suppressor colonies were picked, grown overnight, and subjected to gDNA isolation with the Wizard Genomic DNA extraction kit (Promega). gDNA from the ΔcvpA mutant was also isolated as the parental control strain. gDNA was either submitted for sequencing at the Microbial Genome Sequencing Center at the University of Pittsburgh or prepared for whole genome sequencing in house using the Nextera

XT DNA Library Preparation Kit (Illumina). Libraries were sequenced using the NextSeq 550 or MiSeq platforms, respectively.

Reads were mapped to the *E. coli* O157:H7 EDL933 genome (NCBI Accession Numbers: chromosome, NZ_CP008957.1; pO157 plasmid, NZ_CP008958.1) *using* CLC Genomics Workbench (Qiagen) with default parameters. Basic variant detection was performed, assuming a ploidy of one and filtering out variants with a frequency of less than 35%. Known variants were filtered against the Δ*cvpA* strain. Variant calls were then manually curated to identify SNPs with a frequency greater than 90%.

Bioinformatics/protein prediction

Protein predictions about CvpA structure were inferred from several programs: PSLPred, HHPred, Phobius, Phyre2, and OCTOPUS[493–497]. Protter[498] was used to create a topological diagram. This figure originally appeared in[24].

Topology diagram for CvpA was made using Protter. To examine the conservation of CvpA amino acid sequence across bacterial phyla, we used Annotree version 89[481]. Query term 'CvpA' (KEGG K03558) was used to find bacterial proteins with at least 30% identity and 70% alignment across 27,000 genomes. 20 CvpA sequences from genomes across bacterial phyla were selected for comparison and aligned using MUSCLE[499] within the MEGA X platform[500] using default parameters.

4

Shiga toxin remodels the intestinal epithelial transcriptional response to Enterohemorrhagic *Escherichia coli*

Overview

Enterohemorrhagic Escherichia coli (EHEC) is a food-borne pathogen that causes diarrheal disease and the potentially lethal hemolytic uremic syndrome. We used an infant rabbit model of EHEC infection that recapitulates many aspects of human intestinal disease to comprehensively assess colonic transcriptional responses to this pathogen. Cellular compartment-specific RNA-sequencing of intestinal tissue from animals infected with EHEC strains containing or lacking Shiga toxins (Stx) revealed that EHEC infection elicits a robust response that is dramatically shaped by Stx, particularly in epithelial cells. Many of the differences in the transcriptional responses elicited by these strains were in genes involved in immune signaling pathways, such as IL23A, and coagulation, including F3, the gene encoding Tissue Factor. RNA FISH confirmed that these elevated transcripts were found almost exclusively in epithelial cells. Collectively, these findings suggest that the potent Stx-mediated modulation of innate immune signaling skews the host response to EHEC towards type 3 immunity.

This work is being developed for publication as: Warr AR, Kuehl CJ, Waldor MK. 2019. Shiga toxin remodels the intestinal epithelial transcriptional response to Enterohemorrhagic Escherichia coli.

Author contributions: The work was conceptualized by ARW and MKW. The investigation and data analysis were performed by ARW and CJK with supervision from MKW. The figures were made by ARW. The writing was completed by ARW with editing from MKW.

Introduction

Enterohemorrhagic *Escherichia coli* (EHEC) is an important foodborne pathogen responsible for up to 2 million annual cases of acute gastrointestinal illness[16]. Chiefly colonizing the colon, EHEC typically leads to self-limiting hemorrhagic colitis; however, 5-10% of infected individuals also develop hemolytic uremic syndrome (HUS), a potentially life-threatening complication that can lead to renal failure[66,227]. Supportive rehydration therapy remains the primary treatment for EHEC infection, while antibiotics are associated with elevated frequencies of HUS and therefore contraindicated [12]. There is no vaccine available. EHEC infection is associated with an inflammatory response in the colon, with patients demonstrating elevated fecal leukocytes and calprotectin levels. Colonic biopsy samples from patients with EHEC infection exhibit inflammation, edema, fibrin deposition, neutrophil invasion, and hemorrhage[266,267,289–291].

E. coli O157:H7 is the most common EHEC serotype, but other serotypes have been described. All serotypes share two primary virulence factors: the 'LEE' pathogenicity island that encodes a type III secretion system (T3SS), and prophages that encode one or more Shiga toxins[343]. The activities of EHEC's T3SS effector proteins mediate the pathogen's tight adherence to the colonic mucosa[387], and can promote or antagonize the inflammatory response in epithelial cells[501–504].

During infection, EHEC produces and releases Shiga toxins (Stx) into the intestinal lumen. Stxs are potent AB_5 subunit exotoxins which can bind to the host cell surface glycosphingolipid globotriaosylceramide (Gb3). Once internalized into the eukaryotic cell cytosol, Stx catalyzes a site-specific depurination of the 28s rRNA, which leads to inhibition of protein synthesis and triggers the ribotoxic stress response, production of cytokines, and cell death[252,257,505,251,249]. Absorption of Stx into the blood results in its systemic circulation and damage to endothelial cells, particularly in the renal microvasculature, leading to the characteristic findings of HUS.

The role of Stx in EHEC pathogenicity in the colon is controversial. Although Stx has been associated with colonic pathology[21,23,254,266,267], the mechanisms that explain these observations are unclear because the presence of Gb3 in the colonic epithelium has been disputed[268,269,506,254]. Some suggest that Stx does not directly act on the colonic epithelium, but instead transverses through the epithelial layer to primarily act on endothelial and immune cells[272,273,507,274,275]. Stx may also enter colon epithelial cells through binding alternative receptors[271,506]. Despite the ambiguity surrounding the mechanisms by which Stx exerts toxic effects in the colon during EHEC infection, purified Stx can stimulate inflammatory responses in cultured cells[258,255,508,256,257,254,259–262,249,251]. Analyses of the transcriptional responses to EHEC infection in cultured epithelial cells and organoids have also demonstrated that processes linked to inflammatory signaling, cytoskeletal organization, and apoptosis are altered[284,509,286,263]. However, to date, our knowledge of the host response to EHEC infection is almost exclusively derived from tissue-culture based studies. Because mice do not develop overt diarrhea or colonic pathology during EHEC infection[17], no comprehensive in vivo analyses of how EHEC modifies colonic gene expression patterns during infection have been reported. Moreover, the extent to which Stx contributes to such gene expression changes in vivo is unclear.

Immune responses to bacterial pathogens have been grouped into type 1 or type 3 immunity[510]. Type 1 responses are usually triggered by intracellular pathogens, chiefly through IL-12 stimulated IFNγ-producing cells, which lead to activation of cytotoxic lymphocytes that can destroy infected cells. Type 3 responses are usually triggered by extracellular pathogens. Typically, exposure to pathogen-associated molecular patterns (PAMPs) stimulates dendritic cells to produce IL1 and IL23, which stimulate immune effector cells to produce IL17 and IL22. These cytokines lead to changes in epithelial cells, promoting antimicrobial peptide production and barrier maintenance, and trigger infiltration of phagocytic cells to engulf extracellular pathogens. However, the associations of type 1

and type 3 immunity with intracellular and extracellular pathogens respectively are not strict. For example, *Citrobacter rodentium,* an extracellular pathogen that encodes a T3SS highly similar to EHEC, elicits a type 1 immune response[511–513]; it is unknown if the same is true for Stx-producing EHEC.

Here, we used infant rabbits, a small animal model of EHEC infection where orogastric inoculation of the pathogen leads to a disease that closely mimics the intestinal manifestations of human EHEC disease[21–23], to investigate how Stx production in the gut modifies the cellular response of the colonic mucosa during EHEC infection. We compared the colonic epithelial and lamina propria transcriptional responses to WT and mutant EHEC lacking Stx genes.

Collectively, our findings provide a comprehensive profile of the colonic transcriptional responses to EHEC and suggest that Stx, which primarily impacts the gene expression of epithelial cells, shifts innate immune responses to the pathogen from type 1 toward type 3 immunity.

Shiga toxin promotes apoptosis and hemorrhage in the colonic mucosa

We used EDL933, a prototypical *E. coli* O157:H7 clinical isolate (WT)[63,346], to investigate how Shiga toxins modify the response of the colonic mucosa to EHEC infection. This strain encodes two Stx variants, Stx1 and Stx2, which were both deleted to yield strain ΔΔ*stx*. WT and ΔΔ*stx* were administered orogastrically to infant rabbits to determine if Stx influences EHEC intestinal colonization or the colonic mucosal response to infection. As described previously in experiments with a different EHEC clinical isolate [21], the burden of WT and ΔΔ*stx* in the colon did not differ (Figure 4.1), suggesting that Stx does not alter EHEC colonization in this model. Nonetheless, Stx appeared to contribute to the development of diarrhea as described[21].

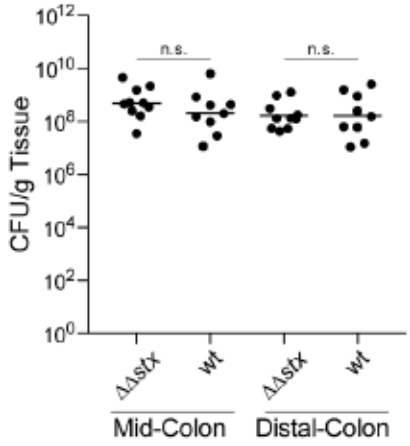


Figure 4.1: Intestinal colonization of WT and ΔΔ*stx* EHEC are similar.

CFU recovered from mid or distal rabbit colon 36 hours post inoculation with either WT or ΔΔ*stx* EHEC. Lines indicate geometric mean. n.s. (not significant) by a two-tailed Mann-Whitney U.

Histopathologic analyses of colon samples from animals inoculated with WT, ΔΔ*stx*, or PBS (mock) were carried out at peak colonization (36-40 hours post inoculation). Compared to mock-treated control animals, colon samples from both WT and ΔΔ*stx* infected rabbits had prominent pathologic changes in the mid and distal colon, sites of maximal colonization. WT and ΔΔ*stx* infections led to similar levels of overall colonic inflammation, characterized by an increased number of small mononuclear cells in the submucosal tissue, as well as comparable levels of heterophil (lapine neutrophil) infiltration (Figure 4.2). Both WT and ΔΔ*stx* infection elicited minor epithelial sloughing in the colon (Figure 4.3). However, compared to the colonic pathology associated with the ΔΔ*stx* strain, there was significantly more apoptosis, indicated by widespread fragmented nuclei, and edema/hemorrhage, indicated by widespread blood and fluid accumulation in tissue, observed in samples from animals infected with WT EHEC (Figure 4.4). These observations support the hypothesis that EHEC production of Stx in the intestine provokes local pathology including apoptosis and hemorrhage in the colonic mucosa.

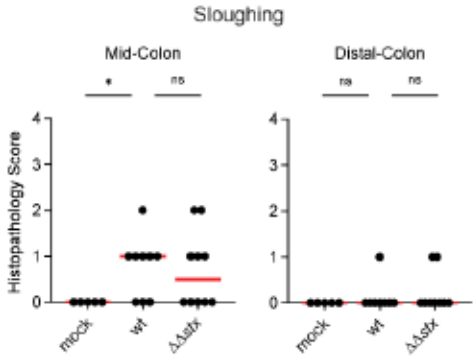


Figure 4.2: Epithelial sloughing is similar in animals infected with WT or ΔΔ*stx* EHEC.

Sloughing in colon sections from infant rabbits inoculated with PBS (mock), WT or ΔΔ*stx* EHEC 36 hours post inoculation. Scores for individual tissue sections are plotted along with the median (red line). Comparisons between groups was made using a Mann-Whitney U test. The Bejmamini-Hochberg Procedure was used to control for the false discovery rate with multiple comparisons at 20%. P-values were considered significant at less than 0.05 (*). n.s. indicates a non-significant difference.

SHIGA TOXIN SHAPES THE COLONIC MUCOSAL TRANSCRIPTOMIC RESPONSE TO EHEC

To further investigate how Stx modifies the colonic mucosa's response to EHEC infection, we used RNA-seq to characterize the transcriptomes of colonic epithelial and lamina propria cells derived from infant rabbits orogastrically inoculated with WT, ΔΔ*stx*, or PBS (mock). In the colon, epithelial cells make initial contact with EHEC and Stx; beneath the epithelial layer, cells in the lamina propria, which include immune cells, respond to signals from the epithelial cells or potentially from direct contact with PAMPs to trigger additional immune responses[510]. To profile transcriptional changes in these cell populations, we harvested colons at the time of peak colonization and collected epithelial and lamina propria cells from 3 rabbits per inoculum type.

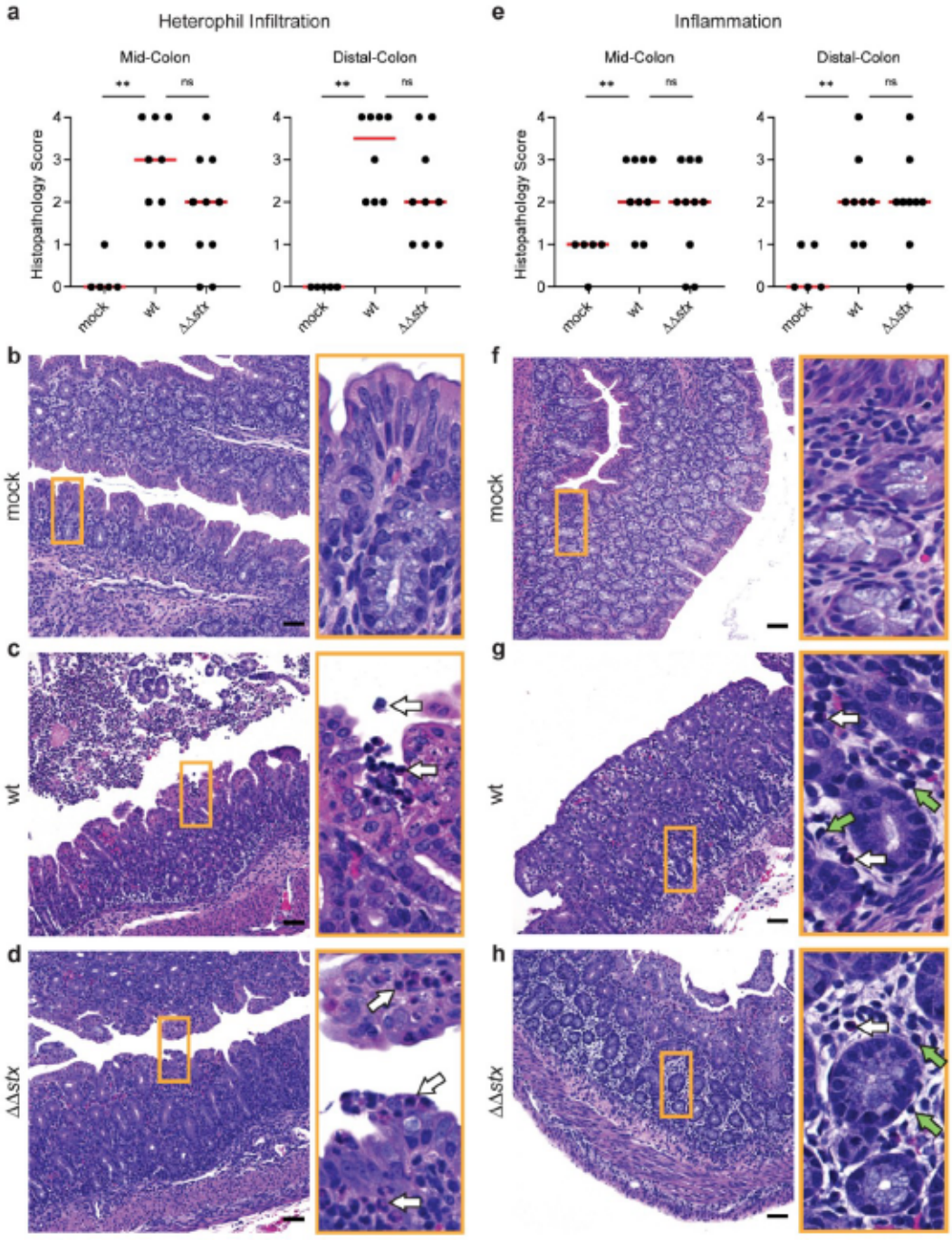


Figure 4.3: Immune cell infiltration is similar in animals infected with WT or ΔΔ*stx* EHEC.

(a) Heterophil infiltration in colon sections from infant rabbits inoculated with PBS (mock), WT or ΔΔ*stx* EHEC 36 hours post inoculation. Scores for individual tissue sections are plotted along with the median (red line). Comparisons between groups was made using a two-tailed Mann-Whitney U test. P-values were considered significant at less than 0.05 (*) or 0.01 (**). n.s. indicates a non-significant difference. (b-d): Representative images from mock (score=0), WT (score=4), and ΔΔ*stx* (score=2) infected colons. Scale bars indicate 50 μm. Orange box denotes inset displayed to the right. White arrows indicate heterophils. (e): Severity of hemorrhage/edema. (f-h): Example images from mock (score=0), WT (score=3), and ΔΔ*stx* (score=3) EHEC-infected colons. Scale bars indicate 50 μm. Orange box denotes inset displayed to the right. White arrows indicate heterophils. Green arrows indicate lymphocytes.

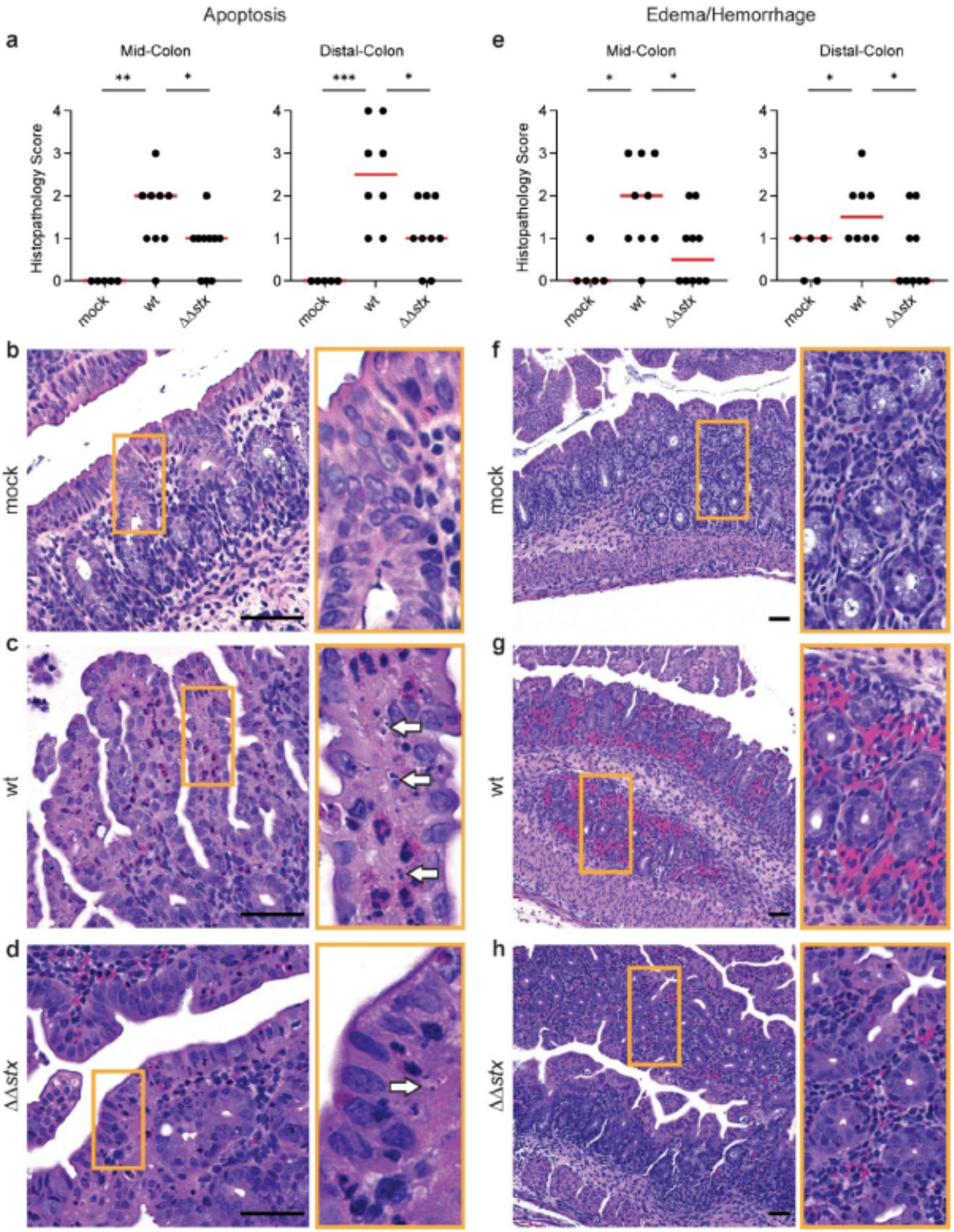


Figure 4.4: Apoptosis and hemorrhage/edema is more prominent in animals infected with WT vs ΔΔ*stx* EHEC.

(a) Apoptotic nuclei in colon sections from infant rabbits inoculated with PBS (mock), WT or ΔΔ*stx* EHEC. Scores for individual animals are plotted with the median (red line). Statistical comparisons were made using a Mann-Whitney U, $p<0.05$ (*), 0.01 (**), or 0.001 (***). (b-d): Example images from mock (score=0), WT (score=4), and ΔΔ*stx* (score=2) infected colons. White arrows indicate apoptotic nuclei. (e): Severity of hemorrhage/edema. (f-h): Representative example images from mock (score=0), WT (score=3) and ΔΔ*stx* (score=1) infected colons. Scale bars indicate 50 μm. Orange boxes denotes inset.

RNA was extracted, subjected to next-generation sequencing, and mapped to the rabbit genome. Expression of two marker genes—*EPCAM,* which encodes for an epithelial cell adhesion protein, and *PTPRC,* which encodes the immune cell surface marker CD45—confirmed the enrichment of epithelial and immune cells respectively (Figure 4.5A). Next, the global gene expression profiles for individual samples were evaluated and compared using the differential gene expression package DESeq2. Principal component analysis (PCA) revealed that epithelial samples from each group clustered separately, highlighting the specificity of the transcriptomic responses elicited by each inoculation type in this compartment (Figure 4.6A). PCA did not segregate the lamina propria samples as neatly based on the presence or absence of *stx* (Figure 4.5B), suggesting that the transcriptional response to Stx is concentrated in the epithelial compartment.

Relative to the mock infected samples, WT infection stimulated more transcriptomic changes than ΔΔ*stx* infection in colonic epithelial cells. Of the ~30,000 rabbit genes surveyed, 1126 had significantly different expression in WT/mock; for ΔΔ*stx*/mock, 497 genes were differentially expressed (Figure 4.7, Tables S13 and S14). 321 genes were in common between the two infection types (Figure 4.8A). Most of these shared genes were associated with GO Molecular Function terms commonly linked to bacterial infection, including cytokine/chemokine signaling and binding, and LPS recognition through TLR4 signaling (Figure 4.8B), implying that many typical pathogen-response pathways are activated similarly in the presence or absence of Stx. Similar patterns were observed in the lamina propria transcriptome (Figure 4.8CD Fig, Tables S15 and S16).

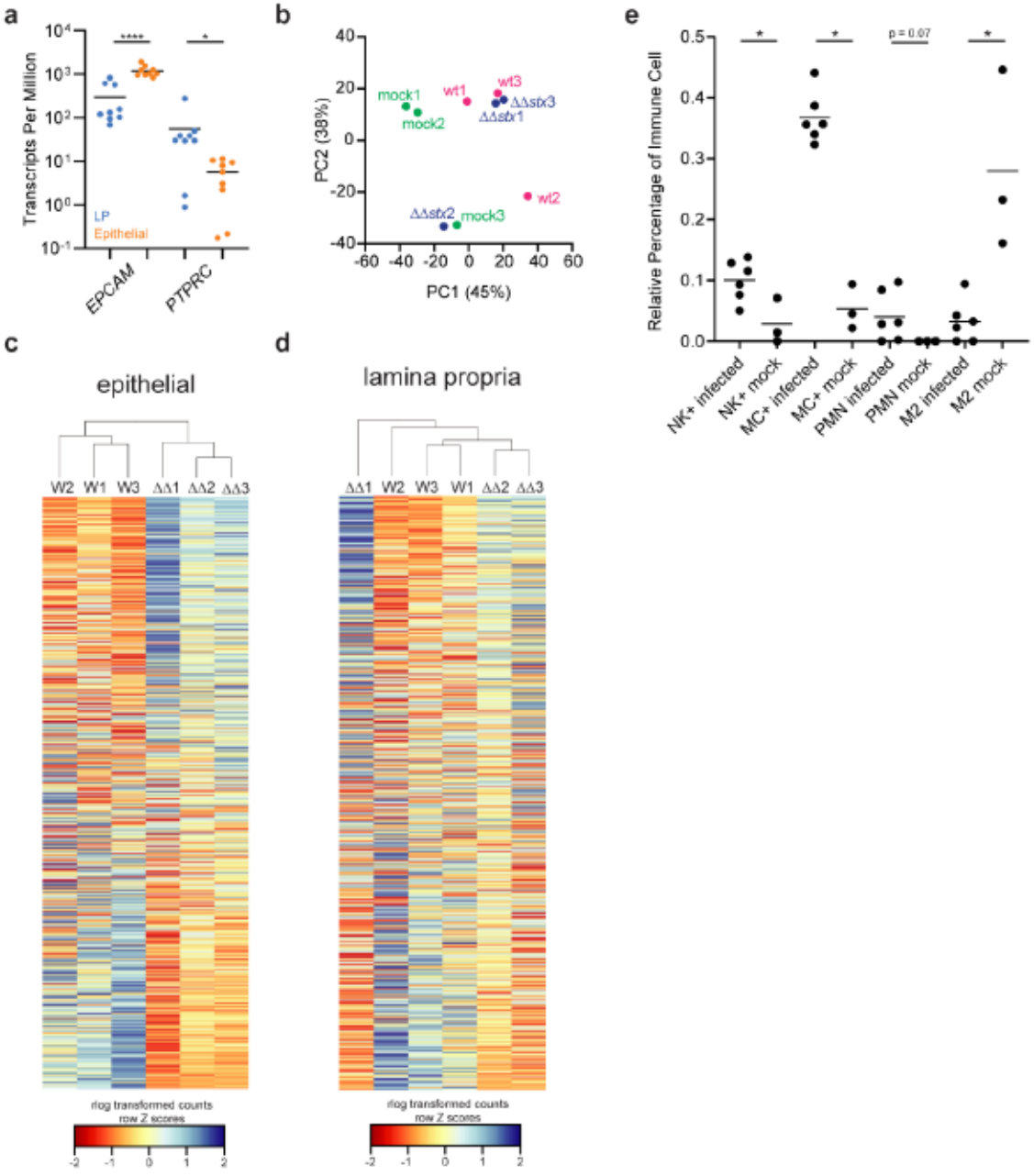


Figure 4.5: EHEC infection leads to transcriptomic changes in the colon.

a) Relative gene expression in transcripts per million for epithelial cell marker *EPCAM* and immune cell marker CD45 gene *PTPRC*. Values for individual rabbits are plotted with mean. Distributions are compared with a Mann-Whitney U test, p< 0.05(*), 0.0001 (****). b) Principal Component Analysis (PCA) of rlog-transformed expression values of lamina propria cells. c) Heat map of rlog-transformed read counts from epithelial cells for 3 animal replicates (WT or Δ*stx* EHEC infected) for all genes by rank. Hierarchical clustering performed using Euclidian sample distances. Rows are normalized by Z-score. d) Heat map of rlog-transformed read counts from lamina propria cells for 3 animal replicates (WT or Δ*stx* EHEC infected) for all genes by rank. Hierarchical clustering performed using Euclidian sample distances. Rows are normalized by Z-score. e) Relative percentages of immune cell types in infected and uninfected tissue as predicted by CIBERSORTx. Absolute distributions are compared with a Mann-Whitney U test, p< 0.05(*).

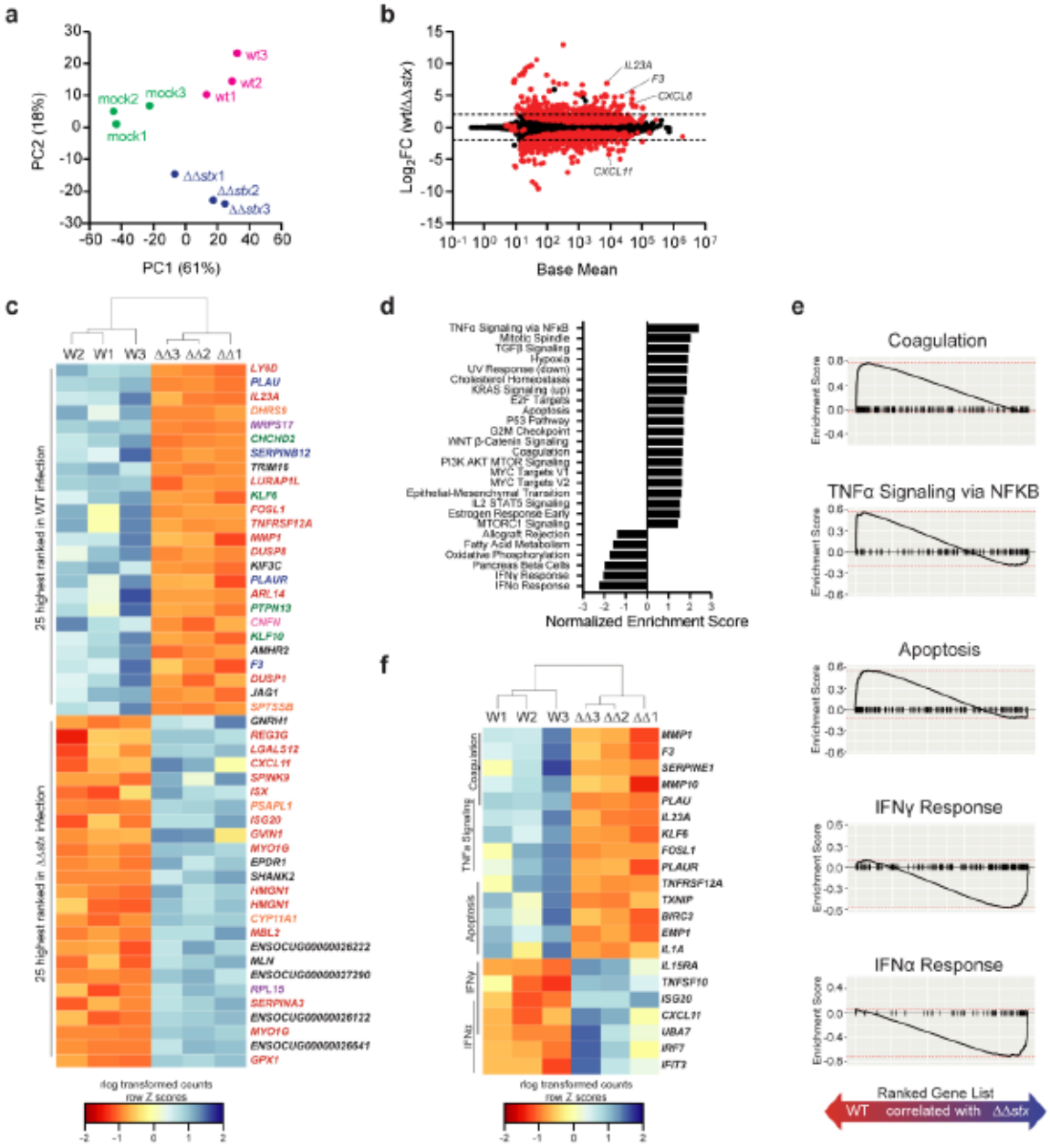


Figure 4.6: Colonic epithelial cell transcriptomes differ between animals infected with WT and ΔΔ*stx* EHEC.

a) Principal component analysis of rlog-transformed expression values b) Average expression level (base mean) and log2 fold change of transcript abundance from rabbits inoculated with WT or ΔΔ*stx* EHEC. Red dots are genes with significantly different ($p<0.05$) transcript abundance. Dashed line indicates log2 fold change >2 or <-2. c) Heat map of rlog-transformed read counts for top 25 and bottom 25 genes by rank. Rows are Z-normalized. Gene names are colored by function: red–immune, orange–metabolism, green–proliferation/apoptosis, blue–coagulation, purple–translation/protein folding, pink–barrier function/cytoskeleton, black–uncharacterized or other. d) Hallmark gene sets significantly associated with WT or ΔΔ*stx* infection. e) Gene set enrichment plot for selected pathways. Black tick marks are genes within pathway organized by rank. f) Heat map of rlog-transformed read counts for top 5 genes in the leading edge for indicated pathway organized by rank.

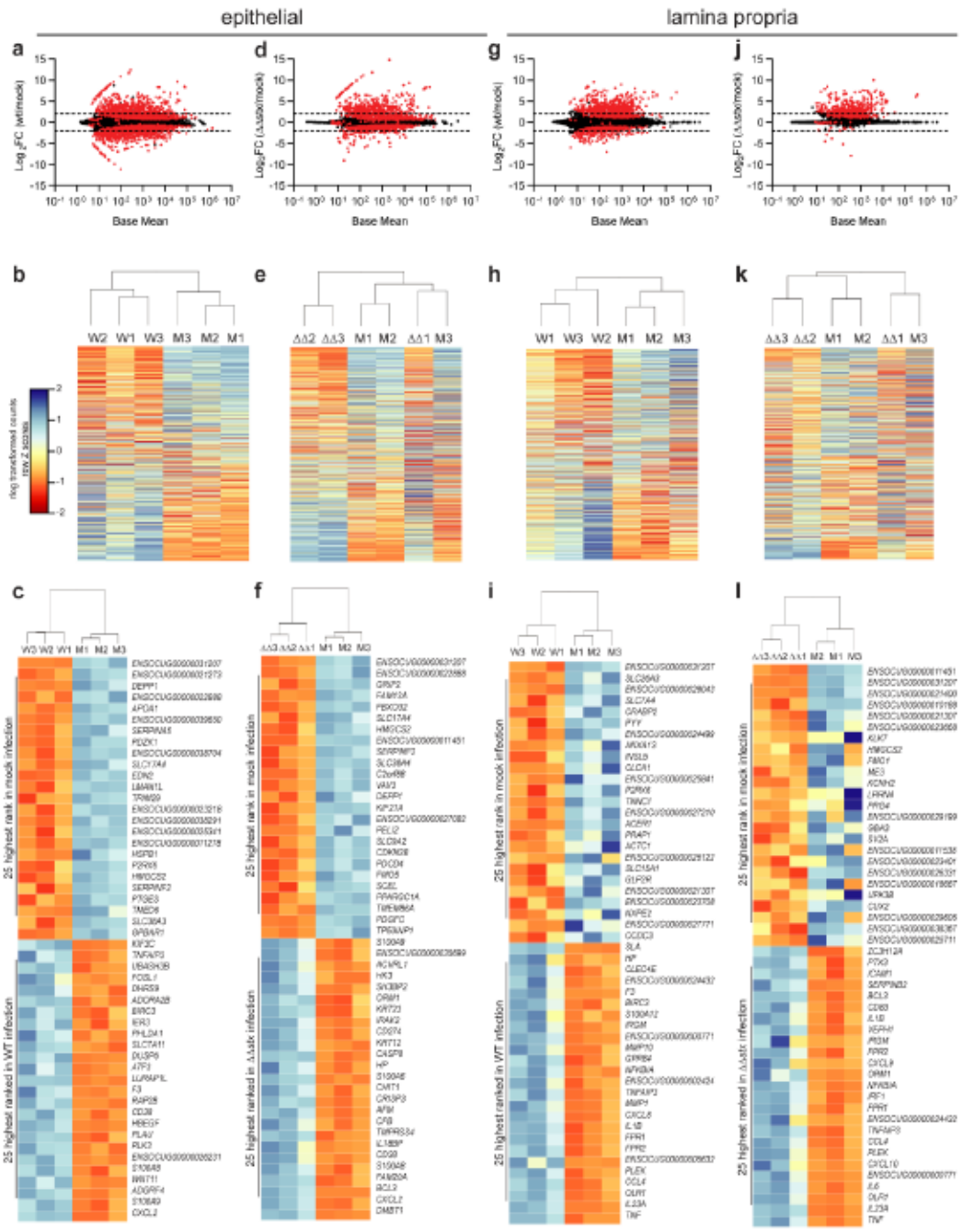


Figure 4.7: Comparison of transcriptional profiles from colonic samples of infant rabbits inoculated with PBS (mock), WT or ΔΔ*stx* EHEC

(a, d, g, j) Average expression level (base mean) and log2 fold change of transcript abundance in colonic epithelial cells (a, d) or lamina propria (g, j) from rabbits inoculated with WT EHEC (a, g) or ΔΔ*stx* EHEC (d, j) compared to PBS (mock). Genes with significantly different (adjusted p-value < 0.05) transcript abundance are highlighted in red. (b, e, h, k) Heat map of rlog-transformed read counts from epithelial cells for 3 animal replicates (WT or ΔΔ*stx* EHEC infected) for all genes by rank. Hierarchical clustering performed using Euclidian sample distances. Rows are normalized by Z-score. The 4 panels correspond to the comparisons shown in the panels immediately above (a,d,g,j); W, WT EHEC, M, mock, ΔΔ, ΔΔ*stx* EHEC. (c, f, i, l)) Heat map of rlog-transformed read counts from epithelial cells for 3 animal replicates (WT or ΔΔ*stx* EHEC infected) top 25 and bottom 25 genes by rank. The 4 panels correspond to the comparisons shown in the panels immediately above (b, e, h, k); W, WT EHEC, M, mock, ΔΔ, ΔΔ*stx* EHEC.

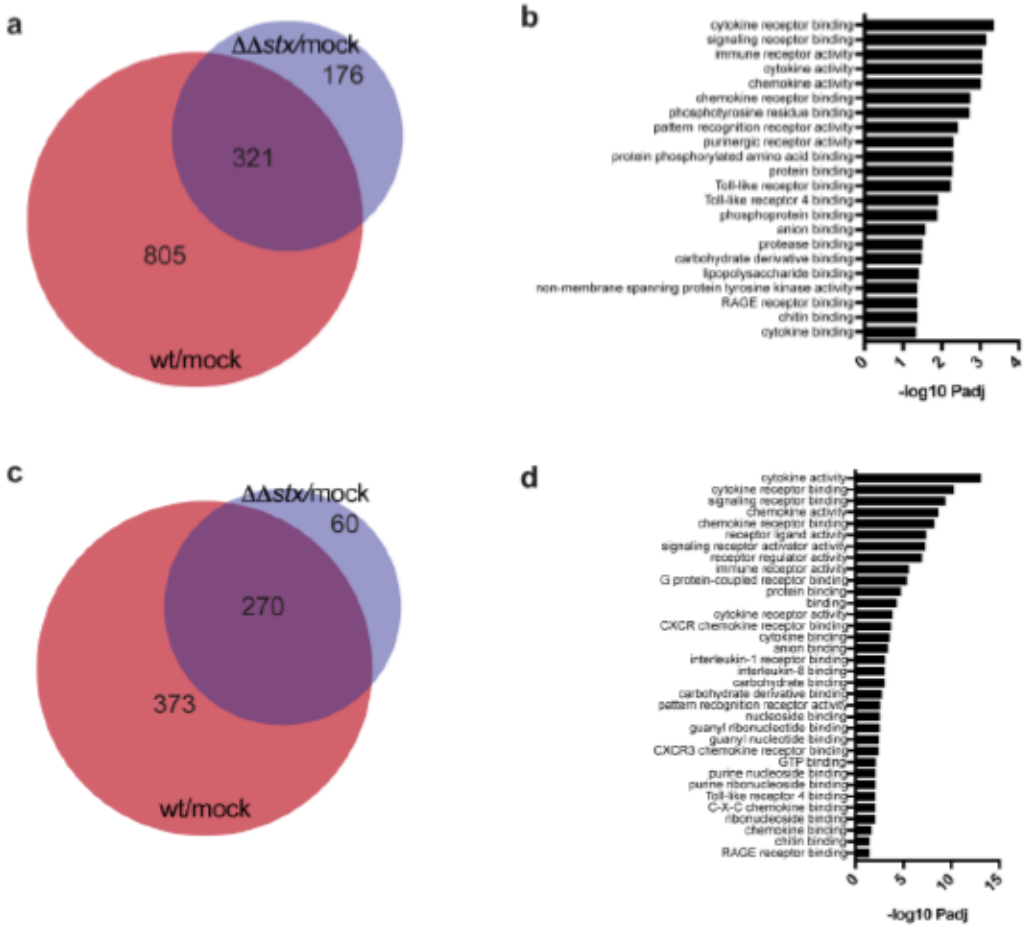


Figure 4.8: WT and ΔΔ*stx* EHEC colonic colonization both stimulate host transcriptional responses commonly associated with infection.

a,c) Venn diagrams of differentially expressed genes in wt vs mock infection and ΔΔ*stx* vs. mock infection from colonic epithelial cells (a) or lamina propria cells (c). b,d) Transcriptional changes elicited by both strains map to many pathways associated with infection by GO Molecular Function analysis in colonic epithelial cells (b) or lamina propria cells (d).

Although the WT and ΔΔ*stx* strains stimulated a shared set of transcripts, we also found robust and widespread differences in the epithelial transcriptomic responses to WT vs ΔΔ*stx* infection. 390 genes exhibited significantly different expression between these two conditions (Figure 4.6B, Table S17). Genes were ranked by adjusted p-value, and hierarchical clustering was performed on rlog transformed read counts. The clustering analysis confirmed that WT and ΔΔ*stx* infection elicit markedly distinct transcriptomic signatures in the epithelium (Figure 4.5C). Many of the top 50 genes by rank were involved in processes related to coagulation and immune signaling (Figure 4.6C). Gene set enrichment analysis (GSEA) was performed to further identify transcriptional processes associated with WT or ΔΔ*stx* infection. We identified several pathways specifically associated with WT or ΔΔ*stx* infection (Figure 4.6D, Table S17). Notably, apoptosis, coagulation, and NFKB signaling were associated with the transcriptomic response to WT infection, whereas pathways linked with IFNα and IFNγ signaling were associated with ΔΔ*stx* infection (Figure 4.6E). Using hierarchical clustering on rlog transformed counts of the top five genes from each enrichment-driving leading-edge subset, we observed dramatic differences between the transcriptional responses to WT and ΔΔ*stx* infection (Figure 4.6F). Specifically, the drivers of coagulation, including the gene coding for tissue factor *F3,* as well as pro-inflammatory cytokines *IL23A* and *IL1A* were specifically associated with WT infection (Figure 4.6F). In ΔΔ*stx* infection, many interferon-stimulated genes (ISGs) were differentially upregulated, including the T-cell chemokine *CXCL11.* Collectively, these analyses suggest that Stx skews epithelial cell innate immune response to EHEC.

In samples from the lamina propria, far fewer genes exhibited differential expression in comparisons between WT and ΔΔ*stx*-infections than in epithelial samples (91 vs 390) (Figure 4.9A, Table S18). Like the PCA analysis (S4B Fig), hierarchical clustering did not separate the transcriptional profiles of WT and ΔΔ*stx*-infected colons as clearly as observed in epithelial samples

(Figure 4.5B vs Figure 4.6A). However, clustering analysis showed that the top 50 genes by rank were distinguishable by inoculum type (Figure 4.9B). Many of these genes function in coagulation and immune signaling pathways, including several matrix metalloproteases (MMPS), immune signaling pathways including IFNγ, ISGs, and *SLAMF6* and *CD207* (Figure 4.9C). GSEA was performed (Table S18), and pathways of interest associated with WT infection included "Th17 vs Th1," "Th17 vs Naive," and "TREM1 signaling" revealing that signatures associated with type 3 immunity are present in the lamina propria (Figure 4.9C). Gene-sets associated with ΔΔ*stx* infection included "Newcastle Virus" which induces a strong interferon response (Figure 4.9C). Examining the leading-edge subset of these pathways underscored that genes associated with Th17 cells are associated with WT infection while IFNγ-stimulated genes such as *CXCL11* are associated with ΔΔ*stx* infection in the lamina propria (Figure 4.9D). Together, these analyses suggest that Stx biases lamina propria transcriptional responses toward expression of genes associated with type 3 immunity and that in the absence of this toxin, IFN-related pathways are more prominent.

Enrichment of differentially expressed genes representative of particular pathways could in part reflect infiltration of cells into a tissue in addition to alterations in expression patterns of resident cells. Both WT and ΔΔ*stx* infection led to similar and significant immune cell infiltration in the colon (Figure 4.2). We characterized these infiltrates further leveraging the bulk RNAseq data and the deconvolution algorithm CIBERSORTx[514,515], which relies on cell-type signature matrices to define cell types of interest, to reveal the relative proportion of cell types in each sample. The relative proportion of immune cell types in the WT infected animals was significantly different than

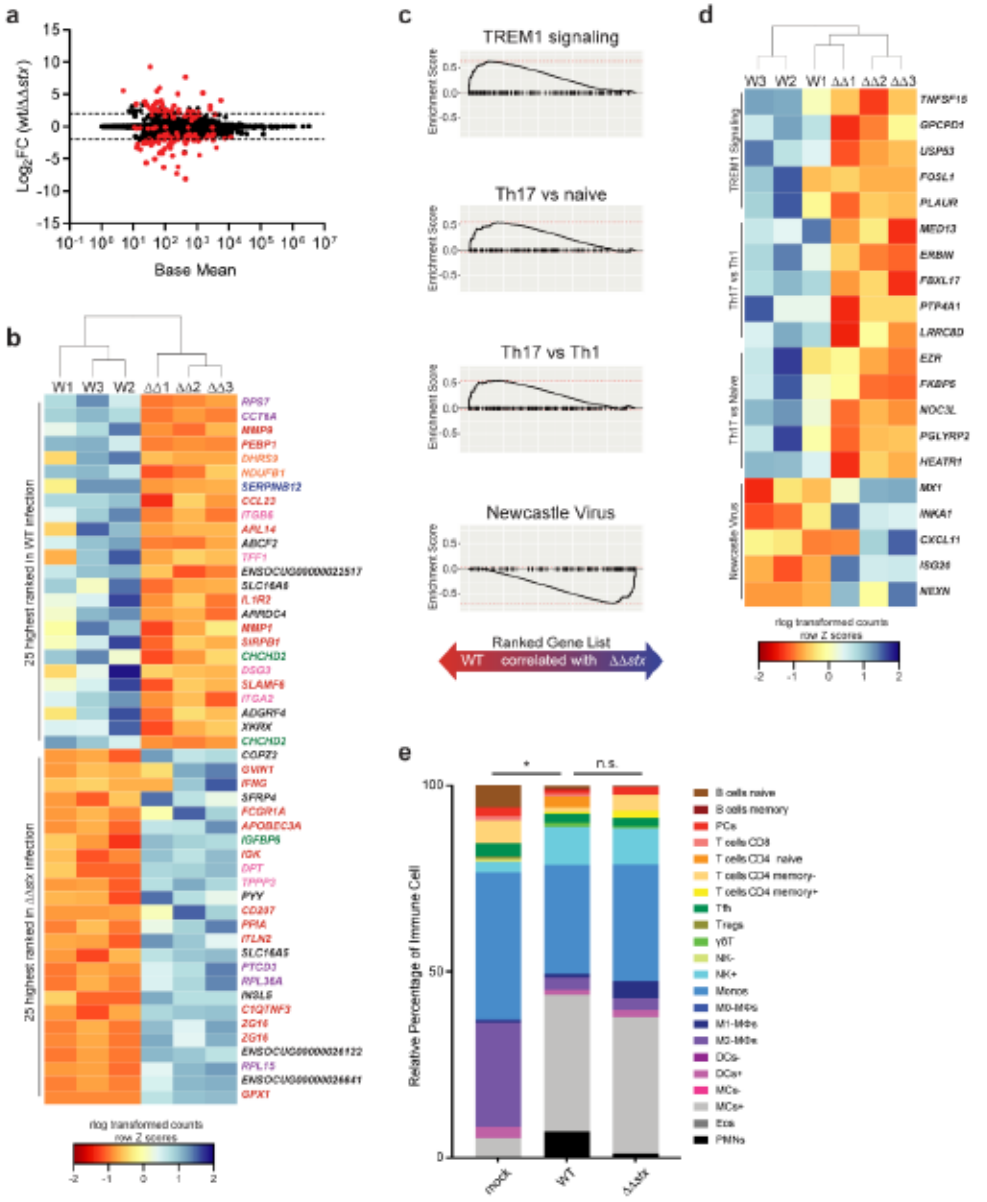


Figure 4.9: Profiles of colonic lamina propria cell transcriptional responses differ between animals infected with WT and ΔΔ*stx* EHEC.

a) Average expression level (base mean) and log2 fold change of transcript abundance from rabbits inoculated with WT or ΔΔ*stx* EHEC. Red dots are genes with significantly different ($p<0.05$) transcript abundance. Dashed line indicates log2 fold change >2 or <-2. b) Heat map of rlog-transformed read counts for top 25 and bottom 25 genes by rank. Rows are Z-normalized. Gene names are colored by function: red – immune, orange – metabolism, green – proliferation/apoptosis, blue – coagulation, purple – translation/protein folding, pink – barrier function/cytoskeleton, black – uncharacterized or other. c) Gene set enrichment plot for selected pathways. Black tick marks are genes within pathway organized by rank. d) Heat map of rlog-transformed read counts for top 5 genes in the leading edge for indicated pathway organized by rank. e) Relative cell abundance of immune cell subtypes determined by CIBERSORTx bulk RNA-seq deconvolution. Distributions compared with the Kolmogorov-Sokolov test, $p<0.05$ (*).

the uninfected (mock) samples, but not different than the distribution in the ΔΔ*stx* infected animals (Figure 4.9E). The infected tissue sections were associated with a higher relative proportion of activated NK cells, activated mast cells, and neutrophils and lower proportion of M2 macrophages compared to uninfected tissue (Figure 4.9E, 4.5E). Thus, as in the histologic sections (Figure 4.2), these analyses suggest that both WT and ΔΔ*stx* stimulate similar immune cell infiltration. Therefore, the differences in the transcriptional profiles observed in animals infected with these strains are likely attributable to Stx-induced changes in the transcriptional landscape of cells resident in the mucosa, rather than differential immune cell infiltration.

Stx stimulates expression of coagulation genes

Comparison of expression profiles from epithelial and lamina propria samples from animals infected with WT vs ΔΔ*stx* EHEC revealed differences in many genes associated with coagulation. Samples from WT infection had higher levels of transcripts for *F3,* a gene which encodes the initiator of the clotting cascade, MMPs and SERPINs, which are proteases that regulate the processing of coagulation cascade proteins, as well as urokinase and the urokinase receptor, which regulate fibrin deposition. Stx is known to cause thrombosis in vascular beds outside of the GI tract, and has been investigated in the kidney microvasculature[249,316,516], but comparatively few studies have focused on Stx-linked coagulation in the intestine.

As antibodies to detect rabbit proteins in tissue are not readily available, we used RNA FISH to investigate the localization of transcripts of interest identified in the RNAseq data. We probed for Tissue Factor (*F3*) transcripts in infected and control samples and found that there was markedly greater *F3* expression in samples from WT-infected vs ΔΔ*stx*-infected or control rabbits (Figure 4.10A). *F3* transcripts were observed in >10% of total DAPI+ colon tissue in WT infection and in only ~0.1% of tissue in ΔΔ*stx* infected animals (Figure 4.10B). Mean fluorescence intensity (MFI), a

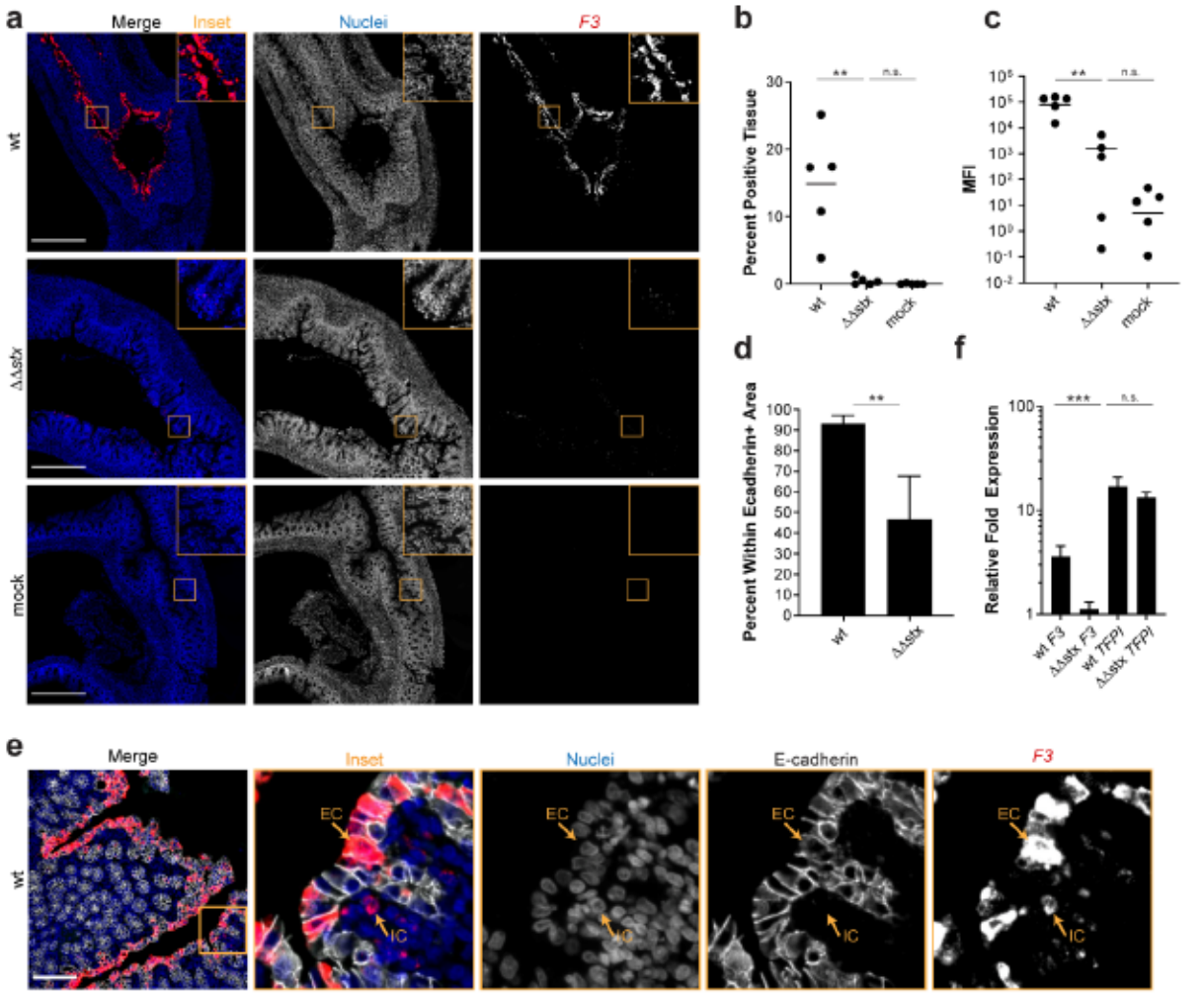


Figure 4.10: Expression of *F3* in epithelial cells is much greater in animals infected with WT vs ΔΔ*stx* EHEC.

a) Micrographs of colon sections from rabbits inoculated with WT EHEC, ΔΔ*stx* EHEC, or PBS (mock) stained with a probe to rabbit *F3* mRNA (red) and DAPI (blue). Scale bar is 500 μM. b) Percentage of tissue section with *F3* signal from individual colons. Distributions compared using Mann-Whitney U test, $p<0.01$ (**), n.s. indicates not significant. c) Mean fluorescent intensity (MFI) from individual colons plotted with mean. Distributions compared using Mann-Whitney U test, $p<0.01$ (**), n.s. indicates not significant. d) Percent *F3* signal within E-cadherin positive cells. Distributions compared using the Mann-Whitney U test, $p<0.05$ (*). e) Sections stained with a probe to rabbit *F3* mRNA (red), DAPI (blue), and anti-E-cadherin antibody (white). Scale bar is 500 μM. Example immune cell (IC) and epithelial cell (EC) are indicated. f) Normalized expression of *F3* and *TFPI* in HT29 cells infected with WT EHEC, ΔΔ*stx* EHEC, or PBS. Expression levels compared with a Students two-tailed t-test, $p<0.001$ (***), n.s. indicates not significant.

proxy for transcript abundance, was much higher in WT vs ΔΔ*stx* samples and the MFI in the latter samples did not differ from the values found in mock infected samples (Figure 4.10C), suggesting that Stx is critical for stimulating *F3* expression. Notably, almost all (>90%) of the *F3*-hybridizing signal in WT samples was detected in cells expressing E-Cadherin, an epithelial marker (Figure 4.10DE). Thus, Stx, when present, appears to primarily act on the colonic epithelium to stimulate *F3* expression. The uniformity of *F3* expression along the epithelium, even in sections where few EHEC cells were detected (Figure 4.11), suggests that Stx may diffuse along the epithelium to modify transcriptional programs in cells that do not have attached EHEC; alternatively, epithelial cells with attached EHEC may secrete factors that modify transcription in neighboring cells.

We also explored whether Stx stimulates *F3* transcription in HT29 cells, a human colonic epithelial cell line. qPCR was used to quantify *F3* transcripts in HT29 cells after infection with WT, ΔΔ*stx* EHEC, or PBS (mock). With WT infection, *F3* gene expression was 5-fold higher than in uninfected cells (Figure 4.10F); moreover, as in rabbits, induction of *F3* expression in HT29 cells was largely dependent on Stx and there was little difference in *F3* expression in uninfected cells vs those infected with ΔΔ*stx* infection (Figure 4.10F). It has been reported that Stx leads to increases in Tissue Factor pro-coagulant protein activity in renal proximal tubule cells and endothelial cells[517–520]; this effect is thought to be primarily driven by a decrease in the expression of Tissue Factor Protein Inhibitor (TFPI), and not an increase in *F3* gene expression[517,521]. We also measured *TFPI* transcript levels in the samples used to measure *F3* expression, to assess if this pathway is also active in HT29 cells. Both WT and ΔΔ*stx* infection similarly stimulated levels of *TFPI* mRNA >10-fold compared to the uninfected cells (Figure 4.10F). Thus, Stx appears to regulate Tissue Factor by different mechanisms in colonic epithelial cells vs endothelial cells.

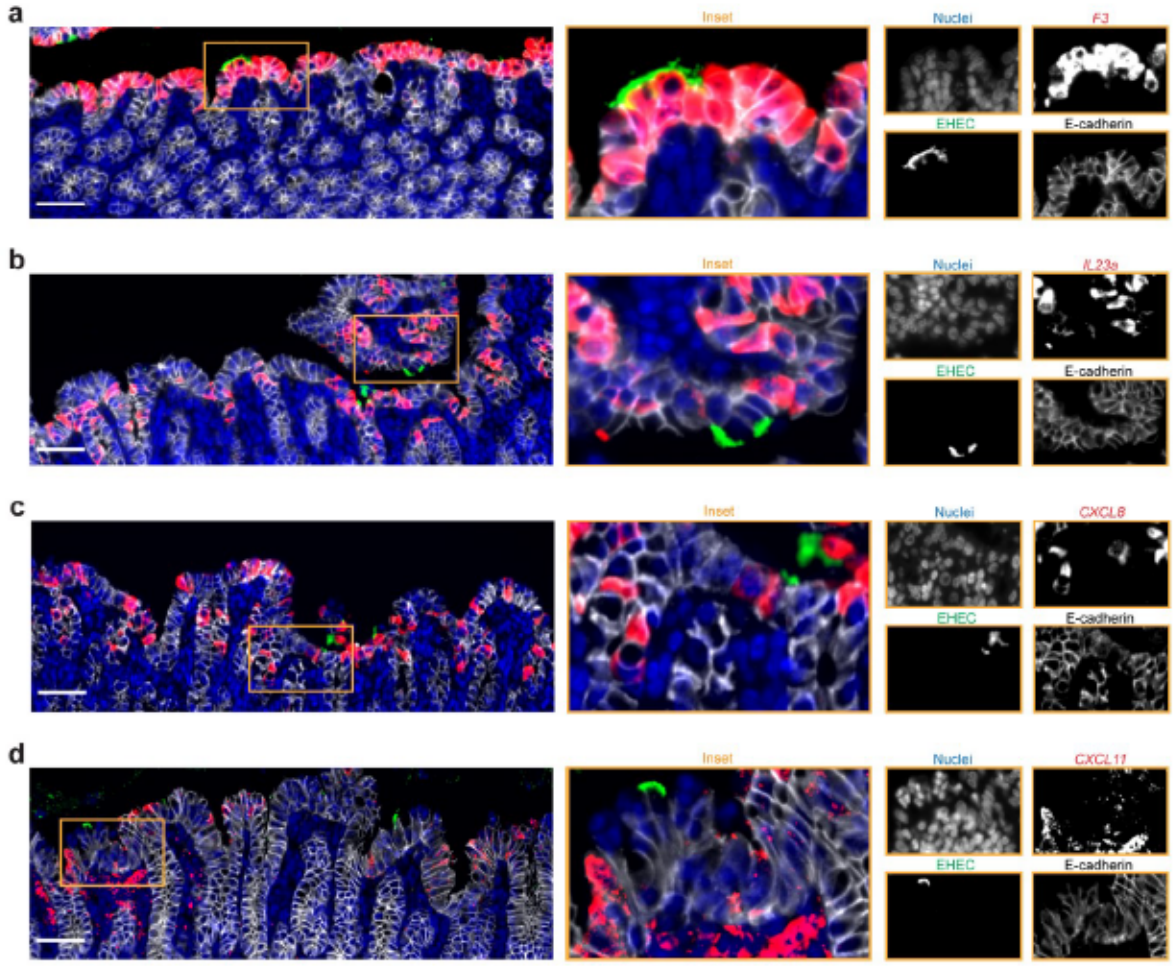


Figure 4.11: Bacteria do not colocalize with RNAscope signal.

Immunofluorescence micrographs of colon sections from rabbits inoculated with WT EHEC (a-c) or ΔΔ*stx* EHEC (d) stained with an RNAscope probe (red) for rabbit *F3* (a), *IL23A* (b), *CXCL8* (c), or *CXCL11* (d), DAPI (blue), and an anti-E-cadherin antibody (white). Scale bar is 50 μM.

Stx biases colonic immune responses toward type 3 immunity

Analyses presented above revealed that Stx skews colonic mucosal gene expression to EHEC infection, stimulating expression of several pro-inflammatory cytokines, such as IL23, relative to levels observed in ΔΔ*stx* infection (Figures 4.6 and 4.9). We used RNA FISH to compare the fluorescence intensity and distribution of *IL23A* transcripts in colons from rabbits infected with WT or ΔΔ*stx* (Figure 4.12A). The intensity and percent of tissue expressing *IL23A* signal in samples from animals infected with WT EHEC were markedly higher than those in animals infected with ΔΔ*stx* EHEC (Figure 4.12BC). These values did not differ in ΔΔ*stx* and control samples, suggesting that Stx stimulates *IL23A* expression. Unexpectedly, nearly all of the *IL23A* expression in WT samples was detected within epithelial cells (E-cadherin positive cells) compared to cells in the lamina propria (Figure 4.12D).

IL23 also promotes the expression of other cytokines such as the chemokine CXCL8, which is a neutrophil chemoattractant. Similar to *F3* and *IL23A* expression, *CXCL8* transcripts were observed at much greater intensity and in a much larger area of tissue in samples from WT vs ΔΔ*stx* infection (Figure 4.13ABC). Furthermore, these *CXCL8* transcripts were primarily present in epithelial cells (Figure 4.13DEF). Similar to *F3, CXCL8* and *IL23A* signal was uniform throughout epithelial tissue, showing no apparent correlation with EHEC foci along the epithelium (Figure 4.11). Together these observations suggest that, during EHEC infection, Stx biases the host immune response toward type 3 immunity by stimulating IL23 expression in the colonic epithelium.

In the RNAseq data, we found that animals infected with ΔΔ*stx* had statistically significant higher levels of *IFNG* transcripts than those infected with WT (Figure 4.9B), suggesting that Stx may inhibit production of IFNγ, a cytokine linked to type 1 immunity. Reduction in IFNγ-mediated STAT-1 phosphorylation has been linked to Stx activity in vitro[522]. In vivo, IFNγ is produced mainly by Th1 cells, ILC1s, NK cells and macrophages in the lamina propria. IFNγ signaling

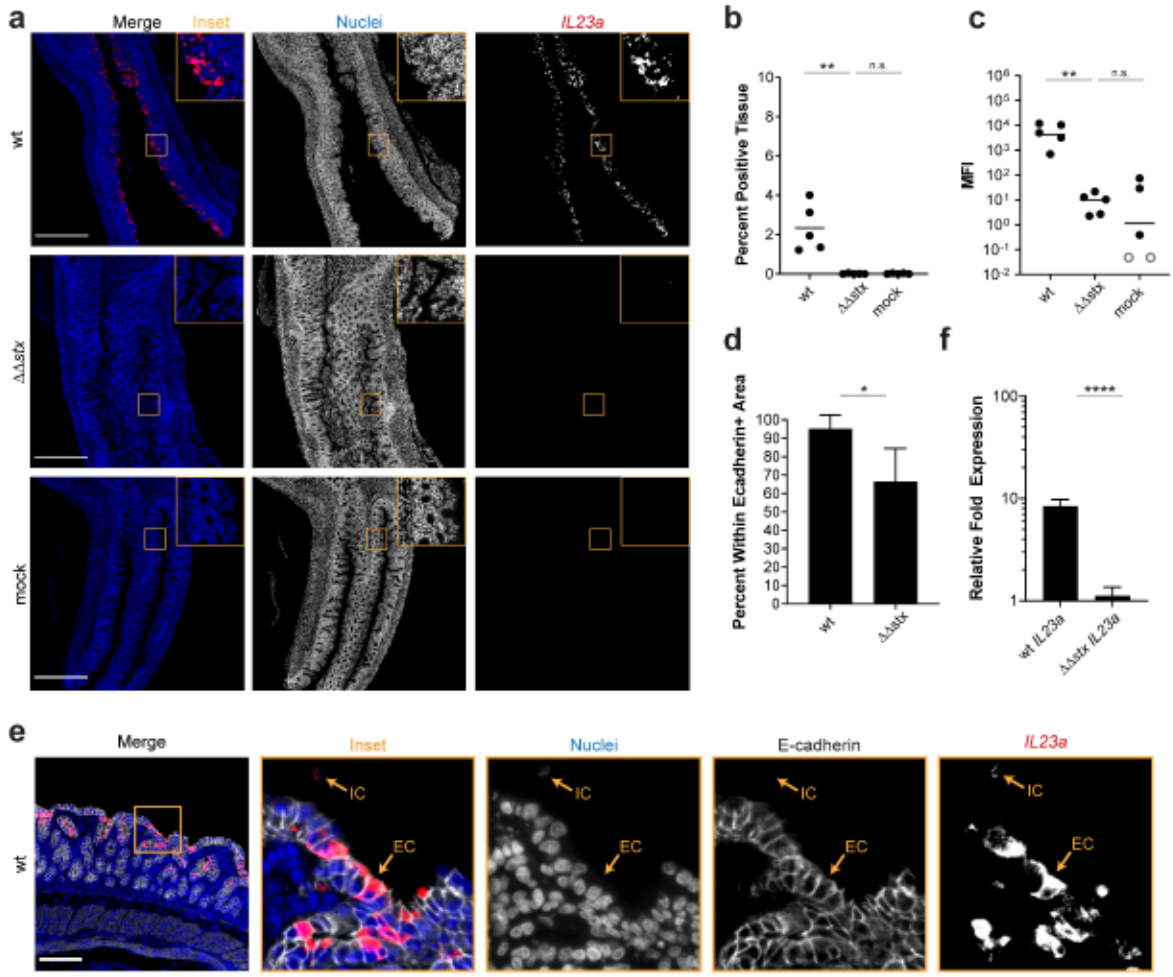


Figure 4.12: Expression of *IL23A* in epithelial cells is greater in animals infected with WT vs ΔΔ*stx* EHEC.

a) Micrographs of colon sections from rabbits inoculated with WT EHEC, ΔΔ*stx* EHEC, or PBS (mock) stained with a probe to rabbit *IL23A* mRNA (red) and DAPI (blue). Scale bar is 500 μM. b) Percentage of tissue section with *IL23A* signal from individual colons. Distributions compared using Mann-Whitney U test, $p<0.01$ (**), n.s. indicates not significant. c) Mean fluorescent intensity (MFI) from individual colons plotted with mean. Distributions compared using Mann-Whitney U test, $p<0.01$ (**), n.s. indicates not significant. d) Percent *IL23A* signal within E-cadherin positive cells. Distributions compared using the Mann-Whitney U test, $p<0.05$ (*). e) Sections stained with a probe to rabbit *IL23A* mRNA (red), DAPI (blue), and anti-E-cadherin antibody (white). Scale bar is 500 μM. Example immune cell (IC) and epithelial cell (EC) is indicated. f) Normalized expression of *IL23A* in HT29 cells infected with WT EHEC, ΔΔ*stx* EHEC, or PBS. Expression levels compared with a Students two-tailed t-test, $p<0.001$ (***), n.s. indicates not significant.

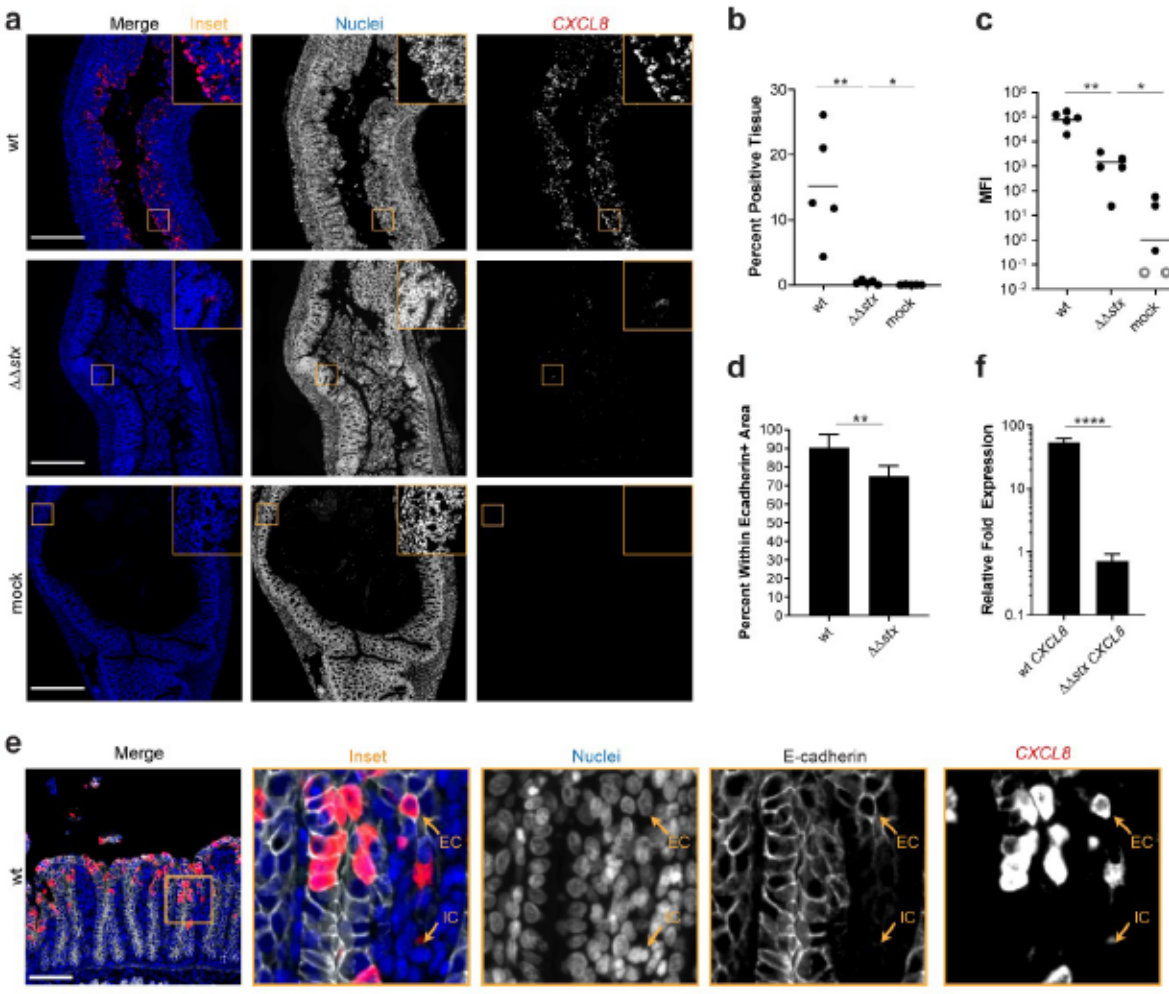


Figure 4.13: Expression of *CXCL8* in epithelial cells is much greater in animals infected with WT vs Δ*stx* EHEC.

a) Micrographs of colon sections from rabbits inoculated with WT EHEC, Δ*stx* EHEC, or PBS (mock) stained with a probe to rabbit *CXCL8* mRNA (red) and DAPI (blue). Scale bar is 500 μM. b) Percentage of tissue section with *CXCL8* signal from individual colons. Distributions compared using Mann-Whitney U test, $p<0.01$ (**), n.s. indicates not significant. c) Mean fluorescent intensity (MFI) from individual colons plotted with mean. Distributions compared using Mann-Whitney U test, $p<0.01$ (**), n.s. indicates not significant. d) Percent *CXCL8* signal within E-cadherin positive cells. Distributions compared using the Mann-Whitney U test, $p<0.05$ (*). e) Sections stained with a probe to rabbit *CXCL8* mRNA (red), DAPI (blue), and anti-E-cadherin antibody (white). Scale bar is 500 μM. Example immune cell (IC) and epithelial cell (EC) is indicated. f) Normalized expression of *CXCL8* in HT29 cells infected with WT EHEC, Δ*stx* EHEC, or PBS. Expression levels compared with a Students two-tailed t-test, $p<0.001$ (***), n.s. indicates not significant.

stimulates transcription of many other genes known as interferon-stimulated genes (ISGs)[523], many of which were found to be differentially expressed in animals infected with WT vs ΔΔ*stx* (Figure 4.6 and 4.9). One of those genes, *CXCL11,* is an IFNγ-inducible chemokine produced by macrophages that acts as a chemoattractant for CXCR3+ T cells; macrophage engagement of CXCR3+ T cells promotes Th1 cell development. RNA FISH analyses revealed that the MFI of the *CXCL11* signal was more intense and detected in more tissue area in samples from animals infected with ΔΔ*stx* vs WT (Figure 4.14ABC). Interestingly, approximately 50% of the *CXCL11* transcript signal was found in epithelial cells (Figure 4.14DE). These observations suggest that Stx inhibits IFNγ signaling in the colonic epithelium during EHEC infection, using this cellular layer to tip the host response towards type 3 immunity.

Discussion

Orogastric inoculation of infant rabbits with EHEC leads to diarrheal disease and colonic pathology that closely resembles many of the intestinal aspects of human EHEC infection[21–23]. Here, we used this model to study how Stx modifies the host response to infection in the colon by comparing the histopathology and transcriptional profiles of the colonic mucosa from animals infected with WT EHEC or an isogenic mutant lacking Stx genes (ΔΔ*stx*). We found that Stx, a potent toxin, increases apoptosis and hemorrhage in colonic tissue and dramatically remodels the colonic epithelium's transcriptional response to EHEC infection. Though the transcriptional responses elicited by WT and ΔΔ*stx* EHEC infection exhibited some overlap, particularly in pathways commonly associated with gram-negative infection, these isogenic strains provoked very distinct transcriptional profiles in the colonic epithelium. Nearly twice as many genes were differentially regulated in animals infected with WT/mock vs ΔΔ*stx*/mock, and many of the differences in the epithelial response related to

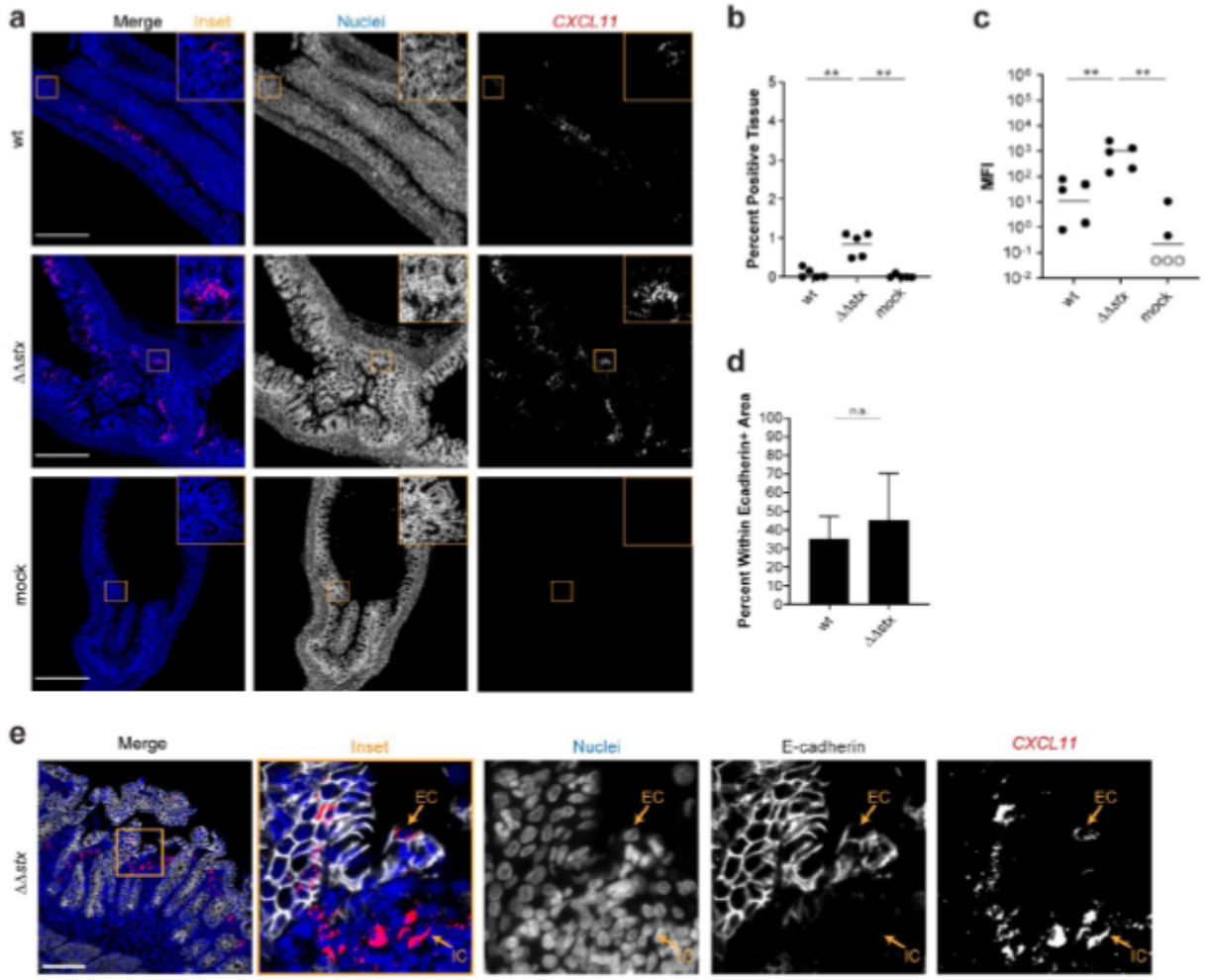


Figure 4.14: Expression of *CXCL11* in is greater in animals infected with ΔΔ*stx* vs WT EHEC.

a) Micrographs of colon sections from rabbits inoculated with WT EHEC, ΔΔ*stx* EHEC, or PBS (mock) stained with a probe to rabbit *CXCL11* mRNA (red) and DAPI (blue). Scale bar is 500 μM. b) Percentage of tissue section with *CXCL11* signal from individual colons. Distributions compared using Mann-Whitney U test, $p<0.01$ (**), n.s. indicates not significant. c) Mean fluorescent intensity (MFI) from individual colons plotted with mean. Distributions compared using Mann-Whitney U test, $p<0.01$ (**), n.s. indicates not significant. d) Percent *CXCL11* signal within E-cadherin positive cells. Distributions compared using the Mann-Whitney U test, $p<0.05$ (*). e) Sections stained with a probe to rabbit *CXCL11* mRNA (red), DAPI (blue), and anti-E-cadherin antibody (white). Scale bar is 500 μM. Example immune cell (IC) and epithelial cell (EC) is indicated.

immune signaling and coagulation. Overall, Stx appears to skew the epithelial transcriptional response toward expression of genes associated with type 3 immunity.

The cellular response to Stx has also been characterized using tissue-cultured cells. The 'ribotoxic stress response' to Stx-mediated cellular damage leads to the upregulation of a variety of genes and production of proteins which modulate the immune response[249–251]. Purified Stx promotes the production of transcription factors (including JUN and FOS) and inflammatory cytokines (CCL2, CCL3, CCL4, CCL5, CSF2, CSF3, CXCL1, CXCL2, CXCL3, CXCL5, CXCL8, IL10, IL1αβ, IL6, TNFα) in cultured epithelial and immune cells[508,255–258,254,259–262,249]. We found that several of these genes, including *JUN, FOS, CCL4, CXCL8,* and *IL1A* were differentially expressed in WT vs Δ*stx* infected colons; thus, Stx by itself, in the absence of additional EHEC-derived factors may account for a subset of the transcriptional changes we identified in the colonic epithelium of animals infected with the WT strain. However, recent transcriptomic studies of the response of human intestinal organoids did not detect differential expression of many of the coagulation-associated or immune signaling genes (including *F3, CXCL8*, and *IL23A*) that we found in infected infant rabbits[262,263]; these differences are potentially explained by organoid culture conditions, which can limit cytokine expression[524]. Overall, these discrepancies highlight the important differences in gene expression patterns observed in cultured cells and the highly complex milieu of the host intestine, and that use of purified toxin is not sufficient to capture the intricacies of the host response to a pathogen with a variety of immunomodulating signals such as LPS and T3SS effectors.

Comparisons of the epithelial transcriptional profiles induced by WT and Δ*stx* EHEC in rabbit colons suggest that Stx promotes the transcription of genes related to type 3 immunity. The expression of *IL23A,* the subunit p19 of the type 3 cytokine IL23, was markedly higher in WT than Δ*stx* infected colons. IL23 drives the immune response towards type 3 immunity, which is often activated in response to extracellular pathogens[510], and is generally thought to be produced by

dendritic cells. Unexpectedly, nearly all of the *IL23A* signal was detected in colonic epithelial cells. However, we did not detect epithelial transcripts for *IL12B*, which encodes for p40, the other subunit of the IL23 protein. Typically, the two subunits are expressed together and form the IL23 heterodimer[525], but the gene expression of these two subunits is not always temporally synchronized[526]. It is possible that IL23 p19 could have an independent function or an alternative partner in intestinal epithelial cells, as has been suggested recently[527]. In contrast to the induction of *IL23A,* Stx appears to inhibit expression of *IFNG* and many ISGs such as *CXCL11,* since these cytokines and downstream factors were downregulated in WT compared to Δ*stx* infected colons. IFNγ is the characteristic cytokine of type 1 immunity and is typically activated to protect cells against intracellular microbes[510]. In vitro studies have also shown that Stx can suppress IFNγ-mediated signaling[522], while similar attaching-and-effacing pathogens which lack Stx do not suppress IFNγ (EPEC)[528] and robustly activate type 1 responses (*Citrobacter rodentium*)[511,513]. Additional bacterial toxins with distinct mechanisms of action have also been reported to bias the innate response along this axis: *Clostridioides difficile* toxin and the heat-labile enterotoxin of *E. coli* and have also been shown to promote type 3 immunity by inducing IL23 production[529,530]. Our findings suggest that Stx shapes the early innate immune response to infection, and thus likely has implications for the development of adaptive immune responses and the host's capacity to clear the infection. Though reagents to verify changes in the abundance of proteins are not readily available for rabbits, we hope that this will be possible in the future.

Stx-mediated damage to endothelial cells, particularly in the renal microvasculature, is a well-studied hallmark of HUS. Toxin damage to endothelial cells triggers the coagulation cascade, which leads to thrombosis with fibrin deposition and hemolysis of RBCs[320]. In the colon, patient biopsies have revealed that EHEC infection can also induce microvascular thrombi and fibrin deposition[267], but few studies have investigated the patterns of gene expression which may contribute to

thrombosis in the intestine. Although we were unable to detect fibrin deposition in tissue sections, we found that *F3*, the gene encoding the initiator of the coagulation cascade, is dramatically induced in colonic epithelial cells in WT but not ΔΔ*stx* infection. Tissue Factor is typically expressed on cells which are not in contact with blood, such as epithelial cells, and can be induced in response to inflammatory stimuli[531]. In endothelial cells, Shiga toxin can promote Tissue Factor activity through a mechanism associated with a decrease in expression of Tissue Factor Protein Inhibitor (TFPI)[517–519,521]. However, in rabbit colonic epithelial cells and in the human colon cancer cell line HT-29, *TFPI* gene expression was not altered despite a marked increase in Tissue Factor expression, suggesting an alternate mechanism of Tissue Factor induction in colonic epithelial cells.

Collectively, our findings reveal that Stx powerfully shapes the host response to EHEC. Though the degree of epithelial transcriptional remodeling by Stx is striking, it is not immediately apparent what benefit this program yields for the pathogen, especially with no discernible difference in the burden of WT and ΔΔ*stx* EHEC in infected animals. It is possible that the shaping of the inflammatory response by Stx augments diarrhea and thereby enhances pathogen dissemination. Additionally, altered gene expression patterns during EHEC infection related to coagulation and inflammation are suggestive of 'thromboinflammation,' a mechanism in which thrombosis and inflammation synergize to contribute to disease pathology [532]. Our findings illustrate the potency of combining isogenic mutants with cellular compartment-specific characterization of host responses to infection, for unravelling how individual virulence factors contribute to the cell type-specific pathogen-host dialogue in disease. For EHEC infection, understanding this dialogue and the innate immune processes contributing to the early colonic phase of disease prior to HUS could offer valuable clues for developing new therapies

Materials and Methods

Ethics Statement

Animal experiments were conducted using protocols approved by Brigham and Women's Hospital Committee on Animals (Institutional Animal Care and Use Committee protocol number 2016N000334 and Animal Welfare Assurance of Compliance number A4752-01) and in accordance with recommendations in the National Institute of Health's Guide for the Care and Use of Laboratory Animals and the Animal Welfare Act of the United States Department of Agriculture.

Bacterial Strains and Growth Conditions

Bacterial strains were cultured in LB medium or on LB agar plates at 37°C. A gentamicin-resistant mutant of *E. coli* O157:H7 strain EDL933 (Δ*lacI*::*aacC1*)[24] was used in all experiments in this study and gentamicin (Gm) was used at 10 μg/mL. The ΔΔ*stx* mutant was constructed using lambda red recombineering[533] as described[534].

Infant rabbit infection and tissue processing

Two-day old litters of mixed gender New Zealand White rabbits were co-housed with a lactating dam (Charles River). Infection inocula were prepared by diluting 100 μl of overnight culture into 100 mL of LB Gm; then, following 3 hours of growth at 37°C with shaking, 30 units of culture at OD_{600} = 1 (about 8 mL) were pelleted and resuspended in 10 mL PBS. Dilutions of the inoculum were plated to enumerate CFU. Each infant rabbit was orogastrically inoculated with 500 μl of the inoculum (~$1x10^9$ CFU), using a size 4 French catheter. Following inoculation, the infant rabbits were monitored at least 2x/day for signs of illness and euthanized 2 days (36-40 hours) post infection, when the entire intestinal tract was removed.

One cm sections of the medial and distal colon were removed post necropsy and the tissue pieces were homogenized in 1 mL of sterile PBS using a minibeadbeater-16 (BioSpec Products, Inc.). Dilution series of the homogenates were plated on LB Gm plates, which were incubated overnight at 37°C, to determine CFU/g bacterial burdens in tissue sections.

Tissue preservation and histopathology

Two cm sections of medial and distal colon were fixed in 2 mL 10% neutral-buffered formalin overnight (~16 hours) at room temperature. The next day, tissue sections were transferred to 2 mL 70% ethanol. Formalin-fixed, paraffin embedded 5 uM sections were stained with hematoxylin and eosin (H&E) by the Rodent Histopathology Core at Dana Farber Cancer Institute. Slides were blindly evaluated by a histopathologist and scored semi-quantitatively. Sections were evaluated for inflammation using the following criteria: 0, none; 1, mild infiltration of immune cells into lamina propria; 2, moderate infiltration; 3, extensive infiltration; 4, severe and extensive infiltration. Apoptosis was evaluated using the following criteria: 0, none; 1, few cells observed with fragmented nuclei; 2, many cells with fragmented nuclei; 3, significant apoptotic nuclei and penetration to crypts; 4, transmural apoptosis. Edema, congestion and hemorrhage were evaluated using the following criteria: 0, none; 1, mild vascular congestion and/or mild edema; 2, moderate congestion and/or edema; 3, congestion with hemorrhage +/ edema; 4, congestion with severe multifocal hemorrhage +/- edema. Heterophil infiltration was evaluated using the following criteria: 0, none; 1, scattered individual heterophils or small clusters in the lamina propria; 2, multifocal aggregates in mucosa with few cells in lumen; 3, multifocal aggregates in mucosa with abundant cell extrusion into lumen; 4, multifocal aggregates in mucosa with large heterophilic intraluminal rafts. Sloughing was evaluated using the following criteria: 0, none; 1, few epithelial cells sloughed from luminal surface; 2, moderate number of epithelial cells sloughed from luminal surface; 3, epithelial surface is severely

disrupted; 4, extensive and severe sloughing (epithelial layer is absent). Scores were compared between infection types using a two-tailed Mann-Whitney U statistical test. The Bejmamini-Hochberg Procedure was used to control for the false discovery rate with multiple comparisons at 20%. P-values were considered significant at less than 0.05 (*), 0.01 (**), and 0.001 (***).

Tissue preparation for RNA-sequencing

Five cm sections from between the medial and distal colon were harvested and processed immediately post necropsy for RNA sequencing from 3 rabbits inoculated with PBS (mock), WT or ΔΔ*stx* EHEC. Epithelial cell and lamina propria cell fractions were isolated from tissue using a method similar to that described previously [535]. First, fat was trimmed from the tissue, and luminal contents were gently pressed out. The tissue was then cut longitudinally and rinsed in 1 mL of wash solution (RPMI 1640, 2% Fetal Bovine Serum (FBS), 10mM HEPES, and 100 μg/mL penicillin-streptomycin). Next, the tissue was rinsed in 40 mL ice-cold Ca/Mg-free HBSS before being transferred to 10 mL of epithelial dissociation solution (HBSS, 100 μg/mL penicillin-streptomycin, 10 mM HEPES, 2% FBS, 10mM EDTA) freshly supplemented with an additional 100 μL of 0.5M EDTA. To remove dying and dead epithelial cells, the tissue was incubated in epithelial dissociation solution 37°C at 125 rpm or 5 minutes, then incubated on ice for 5 minutes, then shaken vigorously 10 times and vortexed for 2 seconds. Supernatants were discarded and the tissue piece was transferred into a new tube of 10 mL of epithelial dissociation solution freshly supplemented with an additional 100 μL of 0.5M EDTA. The solution was brought to room temperature quickly by briefly warming in a 37°C bath, then incubated for 20 minutes at 37°C at 125 rpm centrifugal rotation, then incubated on ice for 5 minutes, shaken vigorously 15 times, and vortexed vigorously for 10 seconds. The supernatant was transferred to a fresh tube and centrifuged at 300xg for five minutes. The cell pellet was resuspended in 2 mL Trizol and the solution was stored at -80°C until RNA extraction.

After epithelial cell dissociation, the remaining tissue piece was transferred to a tube containing 5 mL of enzymatic digestion solution (RPMI 1640, 2% Fetal Bovine Serum (FBS), 10mM HEPES, and 100 μg/mL penicillin-streptomycin, fresh 100 μg/mL Liberase TM, fresh 100 μg/mL DNaseI) and incubated at 37°C with centrifugal rotation at 125 rpm for 30 minutes. The digestion was quenched by adding 80 μL of 0.5M EDTA. The solution was filtered through a 40 uM cell strainer and rinsed with HBSS to a final volume of 30 mL. This tube was spun down at 400xg for 10 minutes, and the cell pellet was resuspended in 2 mL Trizol. Samples were stored at -80°C until RNA extraction.

RNA extraction and mRNA seq library preparation and sequencing

RNA was extracted from Trizol using the Direct-Zol RNA MiniPrep Plus kit from Zymo with some modifications. First, Trizol samples were incubated at 65°C until just thawed (5-10 minutes). One mL samples were added to RNase-free microcentrifuge tubes and 200 μL of chloroform was added to each tube. The tubes were inverted 10x for mixing, and incubated at room temperature for three minutes. The samples were spun at 12,000xg for 15 minutes at 4°C, to separate the aqueous and organic layers. The clear aqueous phase was removed, an equal volume of 100% ethanol was added, and the sample was mixed by inversion ten times before incubating at room temperature for five minutes. The entire volume was transferred to a Direct-Zol spin column and spun for one minute. The samples were washed with pre-wash buffer twice, with wash buffer once, spun empty to remove residual buffer twice, and eluted with 50 μL RNase-free water. Total RNA was assessed for quality and integrity (RINe) using a High Sensitivity RNA ScreenTape (Agilent) at the HMS Biopolymers Facility. RNA of high quality (RINe>8) was prepared for mRNA-sequencing using the KAPA mRNA HyperPrep kit (Roche). Libraries were quantified using a High Sensitivity D1000 ScreenTape (Agilent) and High Sensitivity Qubit. Libraries were sequenced on a NextSeq 550.

Differential expression analysis

First, resources for the rabbit genome were compiled. We concatenated FASTA files for each rabbit chromosome, mitochondrial DNA, and unplaced scaffolds from OryCun2.0 (Assembly GCA_000003625.1) to create a reference genome FASTA file. The Ensembl annotation (May 2019) was used. Sequencing reads and genome resources were uploaded to the Galaxy web platform [536], and the public server usegalaxy.org was used to process and map reads. First, reads were trimmed using Trim Galore! (Galaxy Tool version 0.4.3.1) with automatic adapter sequence detection. Then, trimmed reads were mapped to the rabbit reference genome and annotation using RNA STAR (Galaxy Tool version 2.6.0b-1). featureCounts (Galaxy Tool version 1.6.4+galaxy1) was used to build a count matrix from mapped reads using the Ensembl annotation as a guide. Count matrices were exported from Galaxy and imported into R (version 3.5.3) [537]. The read counts for each gene were normalized to Transcripts Per Kilobase Million (TPM) to compare expression of marker genes across samples. TPM was calculated by first dividing the number of read counts by the length of the gene in kilobases to yield reads per kilobase (RPK). Gene RPKs were summed for each sample and this number was divided by 1,000,000 to yield a sample-specific scaling factor. The RPK value for each gene was divided by the scaling factor to yield TPM. The non-normalized count matrix was also analyzed using DESeq2 (version 1.22.2) [538] to compare the abundance of transcripts between different inoculum types to identify differentially expressed genes. Parametric dispersion was used and shrinkage of effect size was performed using the package apeglm [539]. Genes with an adjusted p-value of less than 0.05 were considered to be differentially expressed. Normalized read counts were generated using a regularized log (r-log) transformation and used to perform principal component analysis (PCA) of each sample. Raw reads and count matrices have been deposited into the Gene Expression Omnibus (GEO) repository (GSE156056).

Hierarchical clustering

For each gene, the parameter 'rank' was calculated by multiplying the sign of the fold-change by the log10 transformed adjusted p-value from DESeq2. Genes were ordered by rank, and hierarchical clustering of samples was performed using Euclidian sample distances using the gplot (version 3.0.1) function Heatmap.2 using r-log transformed read counts, constructing a dendrogram by column (samples). Heatmaps were scaled by row and colors assigned using ColorBrewer palette RdYlBu (version 1.1.2) [540]. Similar clustering analysis was performed for the top 50 genes by rank, and genes in the leading-edge subset from GSEA.

Gene set enrichment

Gene set enrichment was performed using fast GSEA (fGSEA) in R (version 1.8.0) [541]. Only genes with annotation were considered. Genes were ranked with the parameter 'rank'. The hallmark gene sets [542] and immunological signatures gene sets [543] from MSigDB [544] were used. 1000 permutations were completed, and categories with less than 10 genes and greater than 500 genes were excluded. Pathways were considered to be significantly enriched if the adjusted p-value was less than the false-discovery rate of 5%. Enrichment scores were plotted using the plotEnrichment function in R.

Comparing differentially expressed genes in infected vs uninfected tissue

Genes with an adjusted p-value of <0.05 and log2 fold-change of >2 or <-2 from DESeq2 (WT vs mock, Δ*stx* vs mock) were considered differentially expressed. These lists were compared using the online tool BioVenn [545]. The list of commonly differentially expressed genes from these comparisons was analyzed using the webtool g:Profiler to perform functional enrichment analysis (g:GOst) [546].

CIBERSORTx

CIBERSORTx [514,515] was performed using the webtool (https://cibersortx.stanford.edu/) to impute cell fractions. TPM values for all annotated genes were imported as a mixture matrix, and LM22[514,547] was used as a signature matrix file. Distributions of the absolute abundance of immune cells were compared using a Kolmogorov-Sokolov test. Relative abundance of individual cell types were compared using a Mann-Whitney U test.

Fluorescent *in situ* RNA hybridization

Fluorescent in-situ hybridization (FISH) was performed using RNAscope (ACDBio) Multiplex Fluorescent V2 Assay in 5 different rabbit colon tissue sections per inoculum type (WT, ΔΔ*stx* and mock). Manufacturer's protocols were followed and the RNAscope HybEZ oven was used for all incubations. Freshly sectioned formalin-fixed, parrafin embedded (FFPE) tissue sections (5 uM) were processed following manufacturer's protocol. Briefly, sections were treated with boiling target retrieval buffer for 15 minutes and digested with Protease Plus for 28 minutes. Custom RNAscope C1 probes for rabbit mRNA *IL23A, CXCL8, CXCL11, F3,* or *dapB* (negative control) were hybridized to the tissue. The C1 probe was detected with Opal 570 dye (Akoya Biosciences) diluted 1:1000 in Multiplex TSA buffer (ACDBio). Following completion of the RNAscope assay, sections were processed further for immunofluorescence. Sections were permeabilized with 0.1% TritonX100 for 10 minutes at room temperature, washed, and treated with 5% BSA for 1 hour to block non-specific signal. Sections were washed and stained with primary antibody to E-cadherin (1:100 anti E-cadherin mouse monoclonal antibody, BD Biosciences 610181) and/or primary antibody to O157-antigen (1:2000 anti *E. coli* O157:H7 goat polyclonal antibody, Abcam ab30521) overnight at 4°C. The next day, sections were washed and stained with an anti-mouse secondary antibody conjugated to Alexa 647 and/or anti-goal secondary antibody conjugated to Alexa 488

diluted 1:500 in PBS (Invitrogen A-21235). Sections were then stained with DAPI (2 μg/mL) for 5 minutes before mounting with ProLong Diamond antifade mountant (Thermo Fisher).

Quantitative image analysis

Samples were imaged with a Nikon Ti Eclipse microscope equipped with a widefield Andor NeoZyla camera and a 20x objective. For H&E stained sections, Kohler alignment and white-balance was performed, and then RGB images were captured. For sections stained with fluorescent antibodies and FISH probes, 4x6 stitched images (10% overlap) were captured for each tissue section using multi-channel acquisition (blue, red, far-red) using 16-bit imaging. In ImageJ/FIJI[492], threshold values for each channel were determined using the sections stained with the negative control probe (*dapB*), which should have no signal in the red channel above background. To analyze mean fluorescence intensity (MFI), the threshold value for the blue channel (DAPI) was set using FIJI to create a binary mask. This mask was applied to the red channel (RNAscope probe), and a histogram of intensity values for pixels within this mask was recorded. The average value above background was recorded for each sample. To analyze the percentage of tissue expressing a transcript of interest, thresholds were applied to both the blue (DAPI) and red (RNAscope) channels to create binary masks. The area of the DAPI mask was recorded. FIJI 'create selection' tool was used to draw a selection around the DAPI area. This selection was transferred to the RNAscope channel, and the area of RNAscope signal within this area was recorded. We calculated area of tissue expressing signal by dividing area of RNAscope signal (RNAscope area within DAPI selection) by total tissue area (area of DAPI mask). To calculate the percent of the RNAscope signal derived from epithelial cells, we first applied a threshold to the far-red (E-cadherin) and red (RNAscope) channel to create a binary mask. The area of the RNAscope mask was recorded. Then, the 'create selection' tool in FIJI was used to draw a selection around the binary mask of the E-

cadherin area. The section was enlarged by 2μM to accommodate signal at the edge of the cells. This section was transferred to the binary masked RNAscope channel. The area of RNAscope binary signal within the E-cadherin selection was recorded. We divided the area of RNAscope signal within E-cadherin selection by the total RNAscope signal to determine percentage of signal within epithelial cells. These three quantifications (mean fluorescent intensity, percent of tissue expressing signal, and percent of signal within epithelial cells) were compared for WT infected vs mock and ΔΔ*stx* infected vs mock for each probe using a two-tailed Mann-Whitney U statistical test. The Bejmamini-Hochberg Procedure was used to control for the false discovery rate with multiple comparisons at 20%. P-values were considered significant at less than 0.05 (*), 0.01 (**), and 0.001 (***).

Tissue culture infection and RT-qPCR

Human colon colorectal adenocarcinoma cells (HT-29, ATCC HTB-38) were purchased from ATCC and cultured in McCoy's 5A Medium supplemented with 10% fetal bovine serum. Cells were grown at 37°C with 5% CO2. Two days before infection, 500,000 cells were seeded in 6-well plates so that infections occurred at approximately 75% confluency. One hour before infection, the media was changed to Dulbecco's modified Eagle's medium (DMEM) (4.5 g/mL glucose). The bacterial inoculum was prepared by first growing EHEC strains statically in LB overnight at 37°C to OD_{600} of 0.6. Bacteria were resuspended in high-glucose DMEM to OD 0.5 and 45 uL of each inoculum was added to 5 wells of HT29 (MOI 10:1); DMEM alone was added to 5 wells as a mock infection (uninfected). The infections were carried out for a total of 6 hours, but the cells were washed once with DPBS and the media was replaced after 3 hours.

After the 6-hour infection, each well was washed twice with DPBS to remove serum-containing media. RNA was extracted using the RNeasy Plus Mini Kit (Qiagen) and cDNA was generated from

2ug RNA using a High-Capacity cDNA Reverse Transcription kit (Thermo Fisher). Quantitative real-time PCR was performed using a Step One Plus Real-Time PCR machine using Taqman 2x master mix and Taqman probes for *GAPDH* (Hs02758991_g1), *IL23A* (Hs00372324_m1), *F3* (Hs00175225_m1) and *CXCL8* (Hs00174103_m1). Undiluted cDNA was used in the qPCR reactions. Expression levels were calculated using the delta-delta CT method normalized to *GAPDH*. Expression was normalized to the average expression of the 5 uninfected wells. Expression levels were compared using a two-tailed Student's t-test. P-values were considered significant at less than 0.05 (*), 0.01 (**), 0.001 (***), and 0.0001 (****).

5

Conclusions and Outlook

Overview

EHEC is a formidable human pathogen. A critical barrier in the field limiting deeper understanding of the host pathogen interactions that occur during EHEC infection has been lack of a suitable animal model[17,18]. To this day, no animal model completely recapitulates both intestinal and systemic manifestations of EHEC infection. We used infant rabbits, a relatively underused but physiologically relevant model of EHEC intestinal disease[19–23], to conduct functional genomic studies directed at both pathogen and host. We characterized mechanisms of EHEC pathogenicity and the host response to infection in the colonic mucosa, expanding knowledge in the field. This body of work raises many interesting questions for future study.

Perspective

We used transposon-insertion sequencing (TIS) screens to define EHEC genes required for both growth *in vitro* and colonization of the colon[24] (Chapter 2). We generated a library of over 100,000 unique EHEC transposon-insertion mutants using the Himar1 transposon and analyzed their abundance in different conditions as a proxy for fitness. Loci with lower frequency of transposon-insertion than predicted for random mutagenesis were classified as 'underrepresented'. We consider two hypotheses in Chapter 2 for why these regions cannot sustain insertions: either the gene is required for growth and any disruption is lethal, or physical occlusion by DNA binding proteins or chromosome structure limits successful insertion. It is difficult to precisely define what constitutes an 'essential gene', as gene requirements differ across growth conditions, but targeted knock-out libraries in laboratory strains of *E. coli* have estimated there are about 300 loci required for growth[378–380]. Surprisingly, of the 6,032 loci in EHEC, we classified 895 as underrepresented after *in vitro* growth. It is unlikely that EHEC requires an extra 595 genes for growth. We hypothesize that the

abundance of underrepresented loci in EHEC reflects the difficulty of transposon insertion in horizontally acquired regions. These regions usually have a lower GC content than the surrounding genome[548], which could inhibit insertion of transposons due to higher likelihood of association with nucleoid binding proteins like H-NS[368]. H-NS has been shown to selectively bind to and silence horizontally acquired DNA[549,550]. Consistent with this, EHEC loci that were not conserved in *E. coli* K-12 were more likely to be underrepresented in our analysis. However, TIS analysis in *E. coli* using a different transposon, Tn5, only classified 358 underrepresented loci[382]. This implies that the abundance of underrepresented mutants in our analysis is more likely to be a technical issue resulting from limitations on Himar1 insertion rather than underlying biology. Himar1 does have a sequence bias[369], which may be more pronounced in *E. coli* than other species. Analogous TIS analysis using Himar1 in *E. coli* K-12 also classified a high proportion of genes as underrepresented (786 out of 4,483), but in *Vibrio cholerae* and *Vibrio parahaemolyticus* the number was much lower (343/3,456 and 405/4,463, respectively)[551,376]. More work is needed to understand what mechanisms restrict Himar1 insertion in *E. coli,* particularly in horizontally acquired regions. These regions often contain virulence and colonization factors, and limits on inserting transposons into these genes result in diminished statistical power for robust for analysis.

Additional research is also necessary to expand on our observation in Chapter 2 regarding isolate-specific differences in gene essentiality. 131 genes were classified as underrepresented in EHEC that did not have a fitness phenotype in *E. coli* K-12, and hypothesized that these genes could represent pathogen-specific antimicrobial targets. Specifically targeting bacterial pathogens during infection is challenging but critical because broad spectrum antibiotics also remove helpful commensal species and can cause microbiome dysbiosis[552]. Anti-virulence therapeutics, which target genes unique to pathogens, have been considered as a potential strategy to maintain the microbiota during treatment of infections[553]. Another option is to target proteins that are conserved in both

pathogens and commensals, but disruption only produces a fitness defect in pathogens. To robustly identify such candidates, TIS analysis could be performed across many *E. coli* isolates of both pathogenic and non-pathogenic types. Ideally these libraries would be generated using a transposon with minimal bias in *E. coli,* like Tn5, for maximum coverage and successful identification of differentially underrepresented regions.

We also passaged the EHEC mutant library through an infant rabbit to identify mutants that could not colonize the colon (Chapter 2). Theoretically, loss of mutants with *in vivo* fitness defects leaves only those with robust colonization capacity in the population. We recovered 23-38% of EHEC mutants originally present in the inoculum from the colon. It is unlikely that 77% of disrupted loci are actually required for colonization; some of these mutants were eliminated due to random populations constrictions (bottlenecks) that happen *in vivo*[385]. In an infection, bottlenecks represent any pressure that non-specifically and dramatically culls a population, such as passage through the acidic stomach. As discussed in Chapter 2, discrimination between reductions in genotypes attributable to genuine negative selection and stochastic reductions caused by bottlenecks is difficult. We used two computational approaches to model the severity of the bottleneck observed in our data, and used these estimates to identify mutants that had larger than expected reductions in fitness than would be predicted to occur from random loss alone. Surprisingly, EHEC seems to experience a tighter bottleneck in infant rabbits than for other enteric pathogens we have tested, like *V. parahaemolyticus* (61% recovered), despite its higher infectivity in humans. We recently observed that *Shigella flexneri,* a subspecies of *E. coli,* also experiences a dramatic bottleneck in vivo in infant rabbits (~20% recovered)[554]. Factors such as milk-derived antimicrobial peptides in the stomach or members of the rabbit microbiota could disproportionally affect *E. coli* during infection. The microbiota of rabbit species used in our studies (*Oryctolagus cuniculus*) has not been characterized. It

will be helpful to define the infant rabbit microbiota for future studies evaluating the pathogen-commensal interaction during EHEC infection.

Our TIS analysis revealed that approximately 300 genes are required for optimal colonic colonization. These genes included most of the structural elements of the T3SS, but almost none of the effector proteins, except for Tir, NleA, and EspM1. Additional studies are required to investigate the role of NleA and EspM1 further *in vivo*. NleA is thought to affect inflammasome activity[397], and EspM1 can perturb the actin cytoskeleton[399,400]. Interestingly, no other adhesin proteins thought to be required for EHEC attachment to the epithelial surface, including fimbriae or Curli fibers, were required for colonization in our screen. This could reflect species specific differences (rabbit vs human), or redundancy or that these adhesins are not as critical as thought when a functional T3SS is present. Many of the genes that we found to be required for colonization are functionally related to amino acid metabolism. In Chapter 2, we hypothesized that this could reflect nutrient scarcity *in vivo*, and highlights the reliance of EHEC on its own metabolic processes for survival. Another hypothesis is that inactivation of these genes may limit the efficiency of amino-acid decarboxylase systems used by EHEC for acid resistance. For example, *gadE,* encoding a transcription factor that regulates glutamate decarboxylation, and *ilvE,* encoding an amino acid aminotransferase have been linked to acid resistance[104] and were both required for optimal colonic colonization in our screen. Additional studies should aim to investigate the relationship between other amino acid metabolism genes and colonic colonization.

One of the most interesting discoveries from the colonization screen was the gene *cvpA*. *cvpA* mutants were considerably depleted *in vivo*. We found that Δ*cvpA* mutants are extremely sensitive to the antimicrobial activity of bile salts *in vitro*, and in particular the bile salt DOC. We realized that *cvpA* has been identified previously in screens for colonization factors in other enteric pathogens[366,376,427], but had never been investigated mechanistically. The absence of studies regarding

CvpA function is probably attributable to the protein's annotation. A screen 30 years ago to identify proteins required for *E. coli* to secrete the toxic peptide Colicin V identified and named CvpA (Colicin V production protein A)[428]. Later, the genes involved in this process were fully characterized, but a role for CvpA was not found[463]. Despite this, CvpA is still linked to GO terms such as 'toxin biosynthetic processes'. We used a genetic approach to investigate the function of CvpA, identifying suppressor mutations that restored DOC resistance (Chapter 3). These suppressors revealed that CvpA is linked genetically to the σ^E extracytoplasmic stress response. Several independent mutations that enhance expression or activity of σ^E permitted Δ*cvpA* growth on DOC. We do not how CvpA interacts with σ^E. Future work should examine if CvpA is required after DOC exposure for transcription of *rpoE*, or if it regulates activation of σ^E through direct interaction or interaction with any of the known upstream σ^E regulatory proteins. As CvpA is widely conserved across bacteria phyla, including in non-enteric species, we suggested in Chapter 3 that the function of CvpA is not limited to bile resistance and hypothesized that CvpA is involved in maintaining ion homeostasis. Additional work is needed to test if Δ*cvpA* has growth defects in antimicrobial reagents that disrupt ion transport, such as protonophores or electrophiles. We found that CvpA was linked genetically to genes involved with potassium transport, so measuring efflux of K+ ions in the WT and Δ*cvpA* mutant after exposure to DOC could be informative. Biochemical investigations will also be useful to determine molecular mechanisms. CvpA has an unstructured C-terminal tail that is predicted to be periplasmic. We were unable to fuse GFP to this tail and maintain resistance to DOC, which may indicate that an uncharacterized periplasmic binding partner interacts with CvpA's C-terminus. As CvpA is a membrane protein, partner binding assays such as co-immunoprecipitation are difficult. Expression of just the periplasmic portion could simplify future work for binding partners. Another open question is why Δ*cvpA* mutants are specifically sensitive to DOC and not the structurally similar bile salt CHO. As discussed in Chapter 3, DOC is

more hydrophobic than CHO, which may facilitate passage through lipid bilayers[443,487,555]. DOC and DOC-derivatives are also known to trigger ion efflux from eukaryotic cells[556,557] and alter mitochondrial membrane permeability[486], which supports our hypothesis that CvpA responds to ion homeostasis perturbations created by DOC. Future work examining Δ*cvpA* sensitivity to additional reagents will be helpful to clarify mechanism.

We also investigated the host intestine's transcriptional response to EHEC infection (Chapter 4). We inoculated infant rabbits with WT EHEC and an isogenic mutant lacking Stx (ΔΔ*stx*) to determine the effect of Stx on histopathology and gene expression in epithelial and lamina propria cells. To facilitate transcriptional analysis of different cellular compartments in the colon, we performed tissue digestion to enrich for epithelial cells and lamina propria cells. The calcium-chelating agent EDTA disrupts epithelial barrier integrity, releasing sheets of cells from the underlying tissue. Isolation of deeper tissue immune cells relies on enzymatic digestion. We used a mixture of collagenase I, II and the protease thermolysin to degrade the collagen structure of the lamina propria, releasing the non-adherent cells, which are enriched in leukocytes. Nearly complete tissue digestion is required to liberate stromal cells, as these adherent cells are embedded in the ECM[558]. Ideally, we would use flow cytometry to precisely quantify the yield of each cell type, but reagents are unavailable for comprehensive cell profiling in rabbits. We were able to detect enrichment of transcripts associated with epithelial and immune cells in the appropriate fractions, providing confidence in our enrichment strategy. Future work to improve our protocol could include evaluation of murine or human antibodies for detection of rabbit cell markers. This would allow for true cell type isolation by FACS rather than enrichment.

Transcriptionally, both WT and ΔΔ*stx* infection induced expression of genes typically associated with bacterial infection, including pathways that detect and respond to PAMPs such as LPS. Comparing the transcriptional profiles of WT and ΔΔ*stx* infected colons revealed that Stx

dramatically shapes the response to infection. 390 genes were differentially expressed between the two infections in epithelial cells, and 91 in lamina propria cells. Many differentially expressed genes of interest were related to coagulation and immune signaling processes.

Stx contributes to significant hemorrhage in rabbit tissue, so we were initially not surprised to see a number of coagulation-related genes were highly expressed in an Stx-dependent fashion. We first hypothesized that blood in the tissue triggered this gene expression. We measured the distribution of *F3,* which encodes Tissue Factor, in rabbit tissue infected with WT EHEC and found the transcript was evenly distributed across the epithelium, and not localized to regions of hemorrhage. We also detected expression of *F3* in cultured epithelial cells *in vitro* after exposure to WT EHEC but not ΔΔ*stx,* which suggests that Stx interaction with enterocytes alone is sufficient to induce these changes without immune or red blood cells. Future work should focus on protein-level changes. Previous work in renal and endothelial cells *in vitro* described a mechanism for increase in Tissue Factor activity after Stx exposure that relies on decreased expression of Tissue Factor Protein Inhibitor (TFPI)[518–520,517,521]. We found that TFPI expression was unchanged in epithelial cells after EHEC infection, suggesting an alternate mechanism of induction. Future work should focus on characterizing mechanisms of *F3* gene expression in epithelial cells. Screening a CRISPR/Cas9 mutant epithelial cell library for *F3* expression after EHEC infection could identify if genes associated with the canonical ribotoxic stress response or other pathways are required. *F3* expression is also induced by *Yersinia pseudotuberculosis* infection in SI Peyer's patches[559], but not during *Vibrio cholerae* infection in infant rabbits (unpublished). Future work should examine the characteristics of pathogens that can induce coagulation-related genes *in vivo* and the contributions of these protein to tissue pathology or resolution of infection.

The expression of *IFNG* and a number of ISGs such as *CXCL11* was suppressed during WT infection, and induced by the ΔΔ*stx* strain. Others have observed Stx can suppress the IFNγ-

inflammatory pathway *in vitro* at the protein level by blocking phosphorylation of the upstream protein STAT-1[560]. In contrast, A/E pathogens that lack Stx, like *Citrobacter rodentium* and EPEC, activate expression of IFNγ during infection[511–513,528]. It is unclear how Stx blocks STAT-1 phosphorylation and *IFNG* gene expression. T3SS effector protein activity is considered to be the primary mechanism to suppress eukaryotic signaling pathways[561]. Recent studies have shown that T3SS and Stx rely on similar host factors for activity[534], which implies there could be unrecognized synergy between T3SS effectors and Stx which allows for successful modulation of the host response.

We also showed that Stx induces robust expression of the cytokine subunit *IL23A* in epithelial cells, the gene encoding the p19 subunit of the cytokine IL23. In Chapter 4 we discuss that IL23A is typically produced by immune cells, particularly dendritic cells, in response to PAMPs[562], and speculate that Stx could be promoting a shift towards type 3 immunity and away from IFNγ-mediated type 1 immunity during infection. This is not unprecedented, as other bacterial toxins can induce type 3 immunity and IL23 production[529,530]. We did not, however, measure the abundance of specific immune cell types or evaluate their cytokine expression profiles, and this information is needed to classify the type of cell-mediated effector immunity[510]. Analysis of specific immune cell populations during infection has been completed for *Citrobacter rodentium* infections in mice[512,513,563], but the lack of immunogenic flagella and Stx make direct comparisons to EHEC difficult. Future work should evaluate the abundance of Th17, Th1, and ILC subtypes in the colonic mucosa infected with EHEC with or without Stx in order to test this hypothesis. Interestingly, we did not detect transcripts for *IL12B,* the gene encoding the p40 subunit of IL23. *IL23A* has been observed to be expressed independently of *IL12B* in epithelial cells, but the function of IL23A in the absence of p40 is unknown[527]. We also did not detect *IL17A* transcripts in the lamina propria cells. This cytokine is induced in ILC3s and Th17s by IL23 expression. These gene expression patterns suggest

that canonical IL23 signaling is not occurring at the time point of our transcriptional screen. Future work should investigate if p40 protein is produced by epithelial cells, and if IL23A can act independently to trigger transcriptional chances in immune cells during EHEC infection, which may represent an unrecognized aspect of epithelial immunity.

An open question in the field is how colonic disease relates to systemic manifestations of infection, including progression to HUS. Immune cell activation in the colon by Stx is hypothesized to be critical to initiate pathology in the kidneys[564]. Anti-cytokine and anti-inflammatory therapeutics have been suggested as a way to minimize immune cell activation that can lead to renal damage[326]. Our work demonstrates that Stx does have a powerful role in shaping immune responses to infection in the colon. Further work, such as single-cell-RNAseq and proteomics, is needed to more specifically characterize what cell types are present and their cytokine profiles. Although mice are not a useful model for EHEC colonic disease, they might be useful to investigate if elimination or manipulation of gene expression in colon epithelial cells alters the progression to renal disease. This information will be relevant to developing novel therapeutics that mitigate Stx toxicity in extraintestinal tissues.

References

1. Vos, T. *et al.* Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet* **396**, 1204–1222 (2020).
2. GBD 2017 Diarrhoeal Disease Collaborators. Quantifying risks and interventions that have affected the burden of diarrhoea among children younger than 5 years: an analysis of the Global Burden of Disease Study 2017. *Lancet Infect Dis* **20**, 37–59 (2020).
3. Guerrant, R. L., DeBoer, M. D., Moore, S. R., Scharf, R. J. & Lima, A. A. M. The impoverished gut--a triple burden of diarrhoea, stunting and chronic disease. *Nat Rev Gastroenterol Hepatol* **10**, 220–229 (2013).
4. Santosham, M. *et al.* Progress and barriers for the control of diarrhoeal disease. *The Lancet* **376**, 63–67 (2010).
5. GBD 2016 Diarrhoeal Disease Collaborators. Estimates of the global, regional, and national morbidity, mortality, and aetiologies of diarrhoea in 195 countries: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Infect Dis* **18**, 1211–1228 (2018).
6. International Vaccine Access Center. *Pneumonia and Diarrhea Progress Report 2018*.
7. Jonesteller, C. L., Burnett, E., Yen, C., Tate, J. E. & Parashar, U. D. Effectiveness of Rotavirus Vaccination: A Systematic Review of the First Decade of Global Postlicensure Data, 2006–2016. *Clin Infect Dis* **65**, 840–850 (2017).
8. Carvalho, M. F. & Gill, D. Rotavirus vaccine efficacy: current status and areas for improvement. *Hum Vaccin Immunother* **15**, 1237–1250 (2018).
9. Venkatesan, M. M. & Van de Verg, L. L. Combination vaccines against diarrheal diseases. *Hum Vaccin Immunother* **11**, 1434–1448 (2015).
10. Devarajan, P. Acute kidney injury in children: Clinical features, etiology, evaluation, and diagnosis. in *UpToDate* (2020).
11. Gould, L. H. *et al.* Hemolytic Uremic Syndrome and Death in Persons with Escherichia coli O157:H7 Infection, Foodborne Diseases Active Surveillance Network Sites, 2000–2006. *Clin Infect Dis* **49**, 1480–1485 (2009).
12. Wong, C. S., Jelacic, S., Habeeb, R. L., Watkins, S. L. & Tarr, P. I. The risk of the hemolytic-uremic syndrome after antibiotic treatment of Escherichia coli O157:H7 infections. *N. Engl. J. Med.* **342**, 1930–1936 (2000).
13. Mølbak, K., Mead, P. S. & Griffin, P. M. Antimicrobial Therapy in Patients With Escherichia coli O157:H7 Infection. *JAMA* **288**, 1014–1016 (2002).
14. Holtz, L. R. & Tarr, P. I. Shiga toxin-producing Escherichia coli: Clinical manifestations, diagnosis, and treatment. in *Up To Date* (2020).
15. Murano, E. A., Cross, H. R. & Riggs, P. K. The outbreak that changed meat and poultry inspection systems worldwide. *Anim Front* **8**, 4–8 (2018).
16. Majowicz, S. E. *et al.* Global Incidence of Human Shiga Toxin–Producing Escherichia coli Infections and Deaths: A Systematic Review and Knowledge Synthesis. *Foodborne Pathog Dis* **11**, 447–455 (2014).
17. Ritchie, J. M. Animal Models of Enterohemorrhagic Escherichia coli Infection. *Microbiol Spectr* **2**, EHEC-0022-2013 (2014).
18. Mohawk, K. L. & O'Brien, A. D. Mouse Models of Escherichia coli O157:H7 Infection and Shiga Toxin Injection. *J Biomed Biotechnol* **2011**, (2011).
19. Li, Z. *et al.* The effect of enterohemorrhagic Escherichia coli O157:H7 on intestinal structure and solute transport in rabbits. *Gastroenterology* **104**, 467–474 (1993).
20. Elliott, E. *et al.* Modulation of host response to Escherichia coli 0157:H7 infection by anti-CD18 antibody in rabbits. *Gastroenterology* **106**, 1554–1561 (1994).
21. Ritchie, J. M., Thorpe, C. M., Rogers, A. B. & Waldor, M. K. Critical Roles for stx2, eae, and tir in Enterohemorrhagic Escherichia coli-Induced Diarrhea and Intestinal Inflammation in Infant Rabbits. *Infection and Immunity* **71**, 7129–7139 (2003).
22. Ritchie, J. M. & Waldor, M. K. The Locus of Enterocyte Effacement-Encoded Effector Proteins All Promote Enterohemorrhagic Escherichia coli Pathogenicity in Infant Rabbits. *Infection and Immunity* **73**, 1466–1474 (2005).
23. Stone, S. *et al.* Shiga toxin 2-induced intestinal pathology in infant rabbits is A-subunit dependent and responsive to the tyrosine kinase and potential ZAK inhibitor imatinib. *Front. Cell. Infect. Microbiol.* **2**, (2012).
24. Warr, A. R. *et al.* Transposon-insertion sequencing screens unveil requirements for EHEC growth and intestinal colonization. *PLOS Pathogens* **15**, e1007652 (2019).

25. Warr, A. R., Giorgio, R. T. & Waldor, M. K. Genetic analyses link the conserved inner membrane protein CvpA to the σEextracytoplasmic stress response. *bioRxiv* 2020.09.25.314492 (2020)
26. Warr, A. R., Kuehl, C. J. & Waldor, M. K. Shiga toxin remodels the intestinal epithelial transcriptional response to Enterohemorrhagic Escherichia coli. *bioRxiv* 2020.08.11.245555 (2020)
27. Escherich, T. Th. Eseherich. die Darmbakterien des Neugeboreiien und Säuglings. *Dtsch Med Wochenschr* **11**, 740–741 (1885).
28. Blount, Z. D. The unexhausted potential of E. coli. *eLife* **4**, e05826 (2015).
29. Sender, R., Fuchs, S. & Milo, R. Are We Really Vastly Outnumbered? Revisiting the Ratio of Bacterial to Host Cells in Humans. *Cell* **164**, 337–340 (2016).
30. Sender, R., Fuchs, S. & Milo, R. Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLoS Biol* **14**, (2016).
31. Bettelheim, K. A. & Lennox-King, S. M. The acquisition of Escherichia coli by new-born babies. *Infection* **4**, 174–179 (1976).
32. Nowrouzian, F. *et al.* Escherichia coli in Infants' Intestinal Microflora: Colonization Rate, Strain Turnover, and Virulence Gene Carriage. *Pediatric Research* **54**, 8–14 (2003).
33. Palmer, C., Bik, E. M., DiGiulio, D. B., Relman, D. A. & Brown, P. O. Development of the Human Infant Intestinal Microbiota. *PLOS Biology* **5**, e177 (2007).
34. Secher, T., Brehin, C. & Oswald, E. Early settlers: which *E. coli* strains do you not want at birth? *American Journal of Physiology-Gastrointestinal and Liver Physiology* **311**, G123–G129 (2016).
35. Milani, C. *et al.* The First Microbial Colonizers of the Human Gut: Composition, Activities, and Health Implications of the Infant Gut Microbiota. *Microbiol. Mol. Biol. Rev.* **81**, (2017).
36. Forster, S. C. *et al.* A human gut bacterial genome and culture collection for improved metagenomic analyses. *Nature Biotechnology* **37**, 186–192 (2019).
37. Eckburg, P. B. *et al.* Diversity of the Human Intestinal Microbial Flora. *Science* **308**, 1635–1638 (2005).
38. Bentley, R. & Meganathan, R. Biosynthesis of vitamin K (menaquinone) in bacteria. *Microbiol Rev* **46**, 241–280 (1982).
39. Buffie, C. G. & Pamer, E. G. Microbiota-mediated colonization resistance against intestinal pathogens. *Nat Rev Immunol* **13**, 790–801 (2013).
40. Ducarmon, Q. R. *et al.* Gut Microbiota and Colonization Resistance against Bacterial Enteric Infection. *Microbiol. Mol. Biol. Rev.* **83**, (2019).
41. Wu, H.-J. & Wu, E. The role of gut microbiota in immune homeostasis and autoimmunity. *Gut Microbes* **3**, 4–14 (2012).
42. Lazar, V. *et al.* Aspects of Gut Microbiota and Immune System Interactions in Infectious Diseases, Immunopathology, and Cancer. *Front Immunol* **9**, (2018).
43. Negi, S., Das, D. K., Pahari, S., Nadeem, S. & Agrewala, J. N. Potential Role of Gut Microbiota in Induction and Regulation of Innate Immune Memory. *Front. Immunol.* **10**, (2019).
44. Nataro, J. P. & Kaper, J. B. Diarrheagenic Escherichia coli. *Clin. Microbiol. Rev.* **11**, 142–201 (1998).
45. Clements, A., Young, J. C., Constantinou, N. & Frankel, G. Infection strategies of enteric pathogenic Escherichia coli. *Gut Microbes* **3**, 71–87 (2012).
46. Leimbach, A., Hacker, J. & Dobrindt, U. E. coli as an All-Rounder: The Thin Line Between Commensalism and Pathogenicity. in *Between Pathogenicity and Commensalism* (eds. Dobrindt, U., Hacker, J. H. & Svanborg, C.) 3–32 (Springer, 2013).
47. Croxen, M. A. *et al.* Recent Advances in Understanding Enteric Pathogenic Escherichia coli. *Clin Microbiol Rev* **26**, 822–880 (2013).
48. Robins-Browne, R. M. *et al.* Are Escherichia coli Pathotypes Still Relevant in the Era of Whole-Genome Sequencing? *Front Cell Infect Microbiol* **6**, (2016).
49. Pitout, J. Extraintestinal Pathogenic Escherichia coli: A Combination of Virulence with Antibiotic Resistance. *Front. Microbiol.* **3**, (2012).
50. Trabulsi, L. R., Keller, R. & Tardelli Gomes, T. A. Typical and atypical enteropathogenic Escherichia coli. *Emerging Infect. Dis.* **8**, 508–513 (2002).
51. Abram, K. *et al.* What can we learn from over 100,000 Escherichia coli genomes? *bioRxiv* 708131 (2020)
52. Rasko, D. A. *et al.* The Pangenome Structure of Escherichia coli: Comparative Genomic Analysis of E. coli Commensal and Pathogenic Isolates. *Journal of Bacteriology* **190**, 6881–6893 (2008).
53. Touchon, M. *et al.* Organised Genome Dynamics in the Escherichia coli Species Results in Highly Diverse Adaptive Paths. *PLOS Genetics* **5**, e1000344 (2009).

54. Lukjancenko, O., Wassenaar, T. M. & Ussery, D. W. Comparison of 61 Sequenced Escherichia coli Genomes. *Microb Ecol* **60**, 708–720 (2010).
55. Chattopadhyay, S. & Sokurenko, E. V. Chapter 3 - Evolution of pathogenic Escherichia coli. in *Escherichia coli (Second Edition)* (ed. Donnenberg, M. S.) 45–71 (Academic Press, 2013).
56. Sahl, J. W., Morris, C. R. & Rasko, D. A. Chapter 2 - Comparative genomics of pathogenic Escherichia coli. in *Escherichia coli (Second Edition)* (ed. Donnenberg, M. S.) 21–43 (Academic Press, 2013).
57. Maddamsetti, R. *et al.* Core Genes Evolve Rapidly in the Long-Term Evolution Experiment with Escherichia coli. *Genome Biol Evol* **9**, 1072–1083 (2017).
58. Bielaszewska, M. *et al.* Aspects of genome plasticity in pathogenic Escherichia coli. *International Journal of Medical Microbiology* **297**, 625–639 (2007).
59. Brzuszkiewicz, E., Gottschalk, G., Ron, E., Hacker, J. & Dobrindt, U. Adaptation of Pathogenic E. coli to Various Niches: Genome Flexibility is the Key. *Microbial Pathogenomics* **6**, 110–125 (2009).
60. Porse, A. *et al.* Genome Dynamics of Escherichia coli during Antibiotic Treatment: Transfer, Loss, and Persistence of Genetic Elements In situ of the Infant Gut. *Front. Cell. Infect. Microbiol.* **7**, (2017).
61. Delannoy, S., Beutin, L. & Fach, P. Discrimination of Enterohemorrhagic Escherichia coli (EHEC) from Non-EHEC Strains Based on Detection of Various Combinations of Type III Effector Genes. *Journal of Clinical Microbiology* **51**, 3257–3262 (2013).
62. Valilis, E., Ramsey, A., Sidiq, S. & DuPont, H. L. Non-O157 Shiga toxin-producing Escherichia coli—A poorly appreciated enteric pathogen: Systematic review. *International Journal of Infectious Diseases* **76**, 82–87 (2018).
63. Perna, N. T. *et al.* Genome sequence of enterohaemorrhagic Escherichia coli O157:H7. *Nature* **409**, 529–533 (2001).
64. Sheppard, N. MCDONALD'S LINKED TO 47 STOMACH DISORDER CASES - The New York Times. (1982).
65. Johnson, W. M., Lior, H. & Bezanson, G. S. Cytotoxic Escherichia coli O157:H7 associated with haemorrhagic colitis in Canada. *Lancet* **1**, 76 (1983).
66. Karmali, M. A., Petric, M., Steele, B. T. & Lim, C. Sporadic Cases of Haemolytic-Uraemic Syndrome Associated with Faecal Cytotoxin-producing Escherichia coli in stools. *The Lancet* **321**, 619–620 (1983).
67. Riley, L. *et al.* Hemorrhagic colitis associated with a rare Eshcerichia coli serotype. *New England Journal of Medicine* **308**, 681–685 (1983).
68. O'Brien, A. O., Lively, T. A., Chen, M. E., Rothman, S. W. & Formal, S. B. Escherichia coli O157:H7 strains associated with haemorrhagic colitis in the United States produce a Shigella dysenteriae 1 (SHIGA) like cytotoxin. *Lancet* **1**, 702 (1983).
69. Bell, B. P. *et al.* A Multistate Outbreak of Escherichia coli O157:H7—Associated Bloody Diarrhea and Hemolytic Uremic Syndrome From Hamburgers: The Washington Experience. *JAMA* **272**, 1349–1353 (1994).
70. Tuttle, J. *et al.* Lessons from a large outbreak of Escherichia coli O157:H7 infections: insights into the infectious dose and method of widespread contamination of hamburger patties. *Epidemiol Infect* **122**, 185–192 (1999).
71. Centers for Disease Control and Prevention (CDC). Update: multistate outbreak of Escherichia coli O157:H7 infections from hamburgers--western United States, 1992-1993. *MMWR Morb. Mortal. Wkly. Rep.* **42**, 258–263 (1993).
72. Glantz, P. J., Gyles, C. L., Greenfield, J. & Øtskov, F. Isolation of Escherichia coli O157 from Pigs with Colibacillosis in Canada and the United States. *Can J Comp Med* **37**, 200–202 (1973).
73. Furowicz, A. J. & Orskov, F. Two new Escherichia coli O antigens, O150 and O157, and one new K antigen, K92, in strains isolated from veterinary diseases. *Acta Pathol Microbiol Scand B Microbiol Immunol* **80**, 441–444 (1972).
74. McLean, M. M., Jones, C. H. & Sutherland, D. A. Haemolytic-uraemic syndrome. A report of an outbreak. *Arch Dis Child* **41**, 76–81 (1966).
75. Rogers, M. F. *et al.* Hemolytic-uremic syndrome--an outbreak in Sacramento, California. *West. J. Med.* **144**, 169–173 (1986).
76. Steele, B. T., Murphy, N., Arbus, G. S. & Rance, C. P. An outbreak of hemolytic uremic syndrome associated with ingestion of fresh apple juice. *The Journal of Pediatrics* **101**, 963–965 (1982).
77. Griffin, P. M. & Tauxe, R. V. The Epidemiology of Infections Caused by Escherichia coli O157: H7, Other Enterohemorrhagic E. coli , and the Associated Hemolytic Uremic Syndrome. *Epidemiol Rev* **13**, 60–98 (1991).
78. Armstrong, G. L., Hollingsworth, J. & Morris, J. G. Emerging foodborne pathogens: Escherichia coli O157:H7 as a model of entry of a new pathogen into the food supply of the developed world. *Epidemiol Rev* **18**, 29–51 (1996).
79. Encyclopedia of Food Safety - 1st Edition.
80. Persad, A. K. & LeJeune, J. T. Animal Reservoirs of Shiga Toxin-Producing Escherichia coli. *Microbiology Spectrum* **2**, (2014).

81. Woerner, D. R. *et al.* Determining the prevalence of Escherichia coli O157 in cattle and beef from the feedlot to the cooler. *J. Food Prot.* **69**, 2824–2827 (2006).
82. Arthur, T. M. *et al.* Super shedding of Escherichia coli O157:H7 by cattle and the impact on beef carcass contamination. *Meat Sci.* **86**, 32–37 (2010).
83. Besser, T. E., Schmidt, C. E., Shah, D. H. & Shringi, S. "Preharvest" Food Safety for Escherichia coli O157 and Other Pathogenic Shiga Toxin-Producing Strains. *Microbiology Spectrum* 2, (2014).
84. Moxley, R. A. & Acuff, G. R. Peri- and Postharvest Factors in the Control of Shiga Toxin-Producing Escherichia coli in Beef. *Microbiology Spectrum* 2, (2014).
85. Golan, E. *et al.* Food Safety Innovation in the United States: Evidence from the Meat Industry.
86. Doyle, M. E., Archer, J., Kaspar, C. W. & Weiss, R. Human Illness Caused by E. coli O157:H7 from Food and Non-food Sources. 37 (2006).
87. Centers for Disease Control. Reports of Selected E. coli Outbreak Investigations. https://www.cdc.gov/ecoli/outbreaks.html (2020).
88. Zhou, Z. *et al.* Derivation of Escherichia coli O157:H7 from Its O55:H7 Precursor. *PLOS ONE* **5**, e8700 (2010).
89. Feng, P. C. H. *et al.* Genetic diversity among clonal lineages within Escherichia coli O157:H7 stepwise evolutionary model. *Emerging Infect. Dis.* **13**, 1701–1706 (2007).
90. Leopold, S. R. *et al.* A precise reconstruction of the emergence and constrained radiations of Escherichia coli O157 portrayed by backbone concatenomic analysis. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 8713–8718 (2009).
91. Jenke, C. *et al.* Identification of intermediate in evolutionary model of enterohemorrhagic Escherichia coli O157. *Emerging Infectious Diseases* **18**, 582–588 (2012).
92. Jenke, C. *et al.* Phylogenetic analysis of enterohemorrhagic Escherichia coli O157, Germany, 1987-2008. *Emerging Infectious Diseases* **16**, 610–616 (2010).
93. Strockbine, N. A. *et al.* Two toxin-converting phages from Escherichia coli O157:H7 strain 933 encode antigenically distinct toxins with similar biologic activities. *Infect. Immun.* **53**, 135–140 (1986).
94. Scotland, S. M., Smith, H. R., Willshaw, G. A. & Rowe, B. Vero cytotoxin production in strain of Escherichia coli is determined by genes carried on bacteriophage. *Lancet* 2, 216 (1983).
95. O'Brien, A. D. *et al.* Shiga-like toxin-converting phages from Escherichia coli strains that cause hemorrhagic colitis or infantile diarrhea. *Science* 226, 694–696 (1984).
96. O'Brien, A. D., Marques, L. R., Kerry, C. F., Newland, J. W. & Holmes, R. K. Shiga-like toxin converting phage of enterohemorrhagic Escherichia coli strain 933. *Microb. Pathog.* 6, 381–390 (1989).
97. Karch, H. *et al.* The enemy within us: lessons from the 2011 European Escherichia coli O104:H4 outbreak. *EMBO Mol Med* 4, 841–848 (2012).
98. Denamur, E. The 2011 Shiga toxin-producing Escherichia coli O104:H4 German outbreak: a lesson in genomic plasticity. *Clin. Microbiol. Infect.* **17**, 1124–1125 (2011).
99. Muniesa, M., Hammerl, J. A., Hertwig, S., Appel, B. & Brüssow, H. Shiga Toxin-Producing Escherichia coli O104:H4: a New Challenge for Microbiology. *Appl Environ Microbiol* 78, 4065–4073 (2012).
100. Navarro-Garcia, F. Escherichia coli O104:H4 Pathogenesis: an Enteroaggregative E. coli/Shiga Toxin-Producing E. coli Explosive Cocktail of High Virulence. *Microbiology Spectrum* 2, (2014).
101. Rasko, D. A. *et al.* Origins of the E. coli strain causing an outbreak of hemolytic-uremic syndrome in Germany. *N Engl J Med* **365**, 709–717 (2011).
102. Tennant, S. M. *et al.* Influence of Gastric Acid on Susceptibility to Infection with Ingested Bacterial Pathogens. *Infection and Immunity* **76**, 639–645 (2008).
103. Sun, F. J. *et al.* Decreased gastric bacterial killing and up-regulation of protective genes in small intestine in gastrin-deficient mouse. *Dig. Dis. Sci.* **48**, 976–985 (2003).
104. Lund, P., Tramonti, A. & De Biase, D. Coping with low pH: molecular strategies in neutralophilic bacteria. *FEMS Microbiol Rev* **38**, 1091–1125 (2014).
105. Wi, S. M. & Yoon, J. W. Acid resistance mechanisms in enterohemorrhagic Escherichia coli O157:H7. *J. Prev. Vet. Med* **42**, 124–132 (2018).
106. Richard, H. & Foster, J. W. Escherichia coli Glutamate- and Arginine-Dependent Acid Resistance Systems Increase Internal pH and Reverse Transmembrane Potential. *J Bacteriol* **186**, 6032–6041 (2004).
107. U, K. & Wa, H. Mechanisms of acid resistance in Escherichia coli. *Annual review of microbiology* vol. 67 https://pubmed.ncbi.nlm.nih.gov/23701194/ (2013).
108. Pennacchietti, E., D'Alonzo, C., Freddi, L., Occhialini, A. & De Biase, D. The Glutaminase-Dependent Acid Resistance System: Qualitative and Quantitative Assays and Analysis of Its Distribution in Enteric Bacteria. *Front Microbiol* **9**, (2018).

109. Teunis, P. F. M., Ogden, I. D. & Strachan, N. J. C. Hierarchical dose response of E. coli O157:H7 from human outbreaks incorporating heterogeneity in exposure. *Epidemiology & Infection* **136**, 761–770 (2008).
110. Teunis, P., Takumi, K. & Shinagawa, K. Dose Response for Infection by Escherichia coli O157:H7 from Outbreak Data. *Risk Analysis* 24, 401–407 (2004).
111. Bhagwat, A. A. *et al.* Characterization of enterohemorrhagic Escherichia coli strains based on acid resistance phenotypes. *Infect Immun* **73**, 4993–5003 (2005).
112. Fang, F. C., Frawley, E. R., Tapscott, T. & Vazquez-Torres, A. Bacterial Stress Responses during Host Infection. *Cell Host Microbe* **20**, 133–143 (2016).
113. Woodward, S. E., Krekhno, Z. & Finlay, B. B. Here, there, and everywhere: How pathogenic Escherichia coli sense and respond to gastrointestinal biogeography. *Cellular Microbiology* **21**, e13107 (2019).
114. Begley, M., Gahan, C. G. M. & Hill, C. The interaction between bacteria and bile. *FEMS Microbiology Reviews* **29**, 625–651 (2005).
115. Urdaneta, V. & Casadesús, J. Interactions between Bacteria and Bile Salts in the Gastrointestinal and Hepatobiliary Tracts. *Frontiers in Medicine* **4**, (2017).
116. Pifer, R. & Sperandio, V. The Interplay between the Microbiota and Enterohemorrhagic Escherichia coli. *Microbiology Spectrum* 2, (2014).
117. Pacheco, A. R. & Sperandio, V. Enteric Pathogens Exploit the Microbiota-generated Nutritional Environment of the Gut. in *Metabolism and Bacterial Pathogenesis* 279–296 (John Wiley & Sons, Ltd, 2015).
118. Njoroge, J. W., Nguyen, Y., Curtis, M. M., Moreira, C. G. & Sperandio, V. Virulence meets metabolism: Cra and KdpE gene regulation in enterohemorrhagic Escherichia coli. *mBio* **3**, e00280-00212 (2012).
119. Wong, J. M. W. & Jenkins, D. J. A. Carbohydrate Digestibility and Metabolic Effects. *J Nutr* **137**, 2539S-2546S (2007).
120. Holst, J. J., Gribble, F., Horowitz, M. & Rayner, C. K. Roles of the Gut in Glucose Homeostasis. *Diabetes Care* **39**, 884–892 (2016).
121. Miranda, R. L. *et al.* Glycolytic and gluconeogenic growth of Escherichia coli O157:H7 (EDL933) and E. coli K-12 (MG1655) in the mouse intestine. *Infection and Immunity* 72, 1666–1676 (2004).
122. Jimenez, A. G., Ellermann, M., Abbott, W. & Sperandio, V. Diet-derived galacturonic acid regulates virulence and intestinal colonization in enterohaemorrhagic Escherichia coli and Citrobacter rodentium. *Nat Microbiol* **5**, 368–378 (2020).
123. Kumar, A. & Sperandio, V. Indole Signaling at the Host-Microbiota-Pathogen Interface. *mBio* **10**, (2019).
124. Hirakawa, H., Kodama, T., Takumi-Kobayashi, A., Honda, T. & Yamaguchi, A. Secreted indole serves as a signal for expression of type III secretion system translocators in enterohaemorrhagic Escherichia coli O157 : H7. *Microbiology* **155**, 541–550 (2009).
125. Tobe, T., Nakanishi, N. & Sugimoto, N. Activation of motility by sensing short-chain fatty acids via two steps in a flagellar gene regulatory cascade in enterohemorrhagic Escherichia coli. *Infection and Immunity* **79**, 1016–1024 (2011).
126. Nakanishi, N. *et al.* Regulation of virulence by butyrate sensing in enterohaemorrhagic Escherichia coli. *Microbiology (Reading, England)* **155**, 521–530 (2009).
127. Fukuda, S. *et al.* Bifidobacteria can protect from enteropathogenic infection through production of acetate. *Nature* **469**, 543–547 (2011).
128. Johansson, M. E. V. *et al.* The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria. *PNAS* **105**, 15064–15069 (2008).
129. Pelaseyed, T. *et al.* The mucus and mucins of the goblet cells and enterocytes provide the first defense line of the gastrointestinal tract and interact with the immune system. *Immunol Rev* **260**, 8–20 (2014).
130. Pacheco, A. R. *et al.* Fucose sensing regulates bacterial intestinal colonization. *Nature* **492**, 113–117 (2012).
131. Martens, E. C., Chiang, H. C. & Gordon, J. I. Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont. *Cell Host & Microbe* **4**, 447–457 (2008).
132. Hews, C. L. *et al.* The StcE metalloprotease of enterohaemorrhagic Escherichia coli reduces the inner mucus layer and promotes adherence to human colonic epithelium ex vivo. *Cell Microbiol* **19**, (2017).
133. Günther, J. & Seyfert, H.-M. The first line of defence: insights into mechanisms and relevance of phagocytosis in epithelial cells. *Semin Immunopathol* **40**, 555–565 (2018).
134. Garcia, M. A., Nelson, W. J. & Chavez, N. Cell–Cell Junctions Organize Structural and Signaling Networks. *Cold Spring Harb Perspect Biol* **10**, (2018).
135. Madara, J. L. Intestinal absorptive cell tight junctions are linked to cytoskeleton. *Am J Physiol* **253**, C171-175 (1987).

136. Buckley, A. & Turner, J. R. Cell Biology of Tight Junction Barrier Regulation and Mucosal Disease. *Cold Spring Harb Perspect Biol* **10**, a029314 (2018).
137. Cheng, H. & Leblond, C. P. Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine. V. Unitarian Theory of the origin of the four epithelial cell types. *Am J Anat* **141**, 537–561 (1974).
138. Winton, D. J., Blount, M. A. & Ponder, B. A. A clonal marker induced by mutation in mouse intestinal epithelium. *Nature* **333**, 463–466 (1988).
139. Bjerknes, M. & Cheng, H. Clonal analysis of mouse intestinal epithelial progenitors. *Gastroenterology* **116**, 7–14 (1999).
140. van de Wetering, M. *et al.* The beta-catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. *Cell* **111**, 241–250 (2002).
141. Barker, N. *et al.* Identification of stem cells in small intestine and colon by marker gene Lgr5. *Nature* **449**, 1003–1007 (2007).
142. Gehart, H. & Clevers, H. Tales from the crypt: new insights into intestinal stem cells. *Nature Reviews Gastroenterology & Hepatology* **16**, 19–34 (2019).
143. Krndija, D. *et al.* Active cell migration is critical for steady-state epithelial turnover in the gut. *Science* **365**, 705–710 (2019).
144. Darwich, A. S., Aslam, U., Ashcroft, D. M. & Rostami-Hodjegan, A. Meta-Analysis of the Turnover of Intestinal Epithelia in Preclinical Animal Species and Humans. *Drug Metab Dispos* **42**, 2016–2022 (2014).
145. Barker, N. Adult intestinal stem cells: critical drivers of epithelial homeostasis and regeneration. *Nature Reviews Molecular Cell Biology* **15**, 19–33 (2014).
146. Arike, L. *et al.* Protein Turnover in Epithelial Cells and Mucus along the Gastrointestinal Tract Is Coordinated by the Spatial Location and Microbiota. *Cell Reports* **30**, 1077-1087.e3 (2020).
147. Sasaki, N. *et al.* Reg4+ deep crypt secretory cells function as epithelial niche for Lgr5+ stem cells in colon. *PNAS* **113**, E5399–E5407 (2016).
148. Lei, J. *et al.* The antimicrobial peptides and their potential clinical applications. *Am J Transl Res* **11**, 3919–3931 (2019).
149. Worthington, J. J., Reimann, F. & Gribble, F. M. Enteroendocrine cells-sensory sentinels of the intestinal environment and orchestrators of mucosal immunity. *Mucosal Immunology* **11**, 3–20 (2018).
150. von Moltke, J., Ji, M., Liang, H.-E. & Locksley, R. M. Tuft-cell-derived IL-25 regulates an intestinal ILC2–epithelial response circuit. *Nature* **529**, 221–225 (2016).
151. Gerbe, F. *et al.* Intestinal epithelial tuft cells initiate type 2 mucosal immunity to helminth parasites. *Nature* **529**, 226–230 (2016).
152. Wang, Y. *et al.* Single-cell transcriptome analysis reveals differential nutrient absorption functions in human intestine. *J Exp Med* **217**, (2020).
153. Billing, L. J. *et al.* Single cell transcriptomic profiling of large intestinal enteroendocrine cells in mice – Identification of selective stimuli for insulin-like peptide-5 and glucagon-like peptide-1 co-expressing cells. *Molecular Metabolism* **29**, 158–169 (2019).
154. Gao, S. *et al.* Tracing the temporal-spatial transcriptome landscapes of the human fetal digestive tract using single-cell RNA-sequencing. *Nat Cell Biol* **20**, 721–734 (2018).
155. Haber, A. L. *et al.* A single-cell survey of the small intestinal epithelium. *Nature* **551**, 333–339 (2017).
156. Owens, B. M. J. & Simmons, A. Intestinal stromal cells in mucosal immunity and homeostasis. *Mucosal Immunology* **6**, 224–234 (2013).
157. Powell, D. w., Pinchuk, I. v., Saada, J. i., Chen, X. & Mifflin, R. c. Mesenchymal Cells of the Intestinal Lamina Propria. *Annu. Rev. Physiol.* **73**, 213–237 (2011).
158. Pinchuk, I. V., Mifflin, R. C., Saada, J. I. & Powell, D. W. Intestinal mesenchymal cells. *Curr Gastroenterol Rep* **12**, 310–318 (2010).
159. Mowat, A. M. & Agace, W. W. Regional specialization within the intestinal immune system. *Nat Rev Immunol* **14**, 667–685 (2014).
160. The Mucosal Immune System. in *Janeway's Immunobiology*.
161. OKUMURA, R. & TAKEDA, K. Maintenance of gut homeostasis by the mucosal immune system. *Proc Jpn Acad Ser B Phys Biol Sci* **92**, 423–435 (2016).
162. Brandtzaeg, P. & Johansen, F.-E. Mucosal B cells: phenotypic characteristics, transcriptional regulation, and homing properties. *Immunol Rev* **206**, 32–63 (2005).
163. Spits, H. *et al.* Innate lymphoid cells — a proposal for uniform nomenclature. *Nature Reviews Immunology* **13**, 145–149 (2013).

164. Buettner, M. & Lochner, M. Development and Function of Secondary and Tertiary Lymphoid Organs in the Small Intestine and the Colon. *Front. Immunol.* **7**, (2016).
165. McWilliams, B. D. & Torres, A. G. EHEC Adhesins. *Microbiol Spectr* **2**, EHEC-0003-2013 (2014).
166. Olsén, A., Jonsson, A. & Normark, S. Fibronectin binding mediated by a novel class of surface organelles on Escherichia coli. *Nature* **338**, 652–655 (1989).
167. Olsén, A., Arnqvist, A., Hammar, M., Sukupolvi, S. & Normark, S. The RpoS sigma factor relieves H-NS-mediated transcriptional repression of csgA, the subunit gene of fibronectin-binding curli in Escherichia coli. *Mol Microbiol* **7**, 523–536 (1993).
168. Farfan, M. J., Cantero, L., Vidal, R., Botkin, D. J. & Torres, A. G. Long Polar Fimbriae of Enterohemorrhagic Escherichia coli O157:H7 Bind to Extracellular Matrix Proteins ▿ . *Infect Immun* **79**, 3744–3750 (2011).
169. Tzipori, S. *et al.* The pathogenesis of hemorrhagic colitis caused by Escherichia coli O157:H7 in gnotobiotic piglets. *J. Infect. Dis.* **154**, 712–716 (1986).
170. Knutton, S., Baldwin, T., Williams, P. H. & McNeish, A. S. Actin accumulation at sites of bacterial adhesion to tissue culture cells: basis of a new diagnostic test for enteropathogenic and enterohemorrhagic Escherichia coli. *Infect. Immun.* **57**, 1290–1298 (1989).
171. Wong, A. R. C. *et al.* Enteropathogenic and enterohaemorrhagic Escherichia coli: even more subversive elements. *Molecular Microbiology* **80**, 1420–1438 (2011).
172. Stevens, M. P. & Frankel, G. M. The Locus of Enterocyte Effacement and Associated Virulence Factors of Enterohemorrhagic Escherichia coli. *Microbiol Spectr* **2**, EHEC-0007-2013 (2014).
173. Vingadassalom, D. *et al.* Insulin receptor tyrosine kinase substrate links the E. coli O157:H7 actin assembly effectors Tir and EspFU during pedestal formation. *Proc Natl Acad Sci U S A* **106**, 6754–6759 (2009).
174. de Groot, J. C. *et al.* Structural basis for complex formation between human IRSp53 and the translocated intimin receptor Tir of enterohemorrhagic E. coli. *Structure* **19**, 1294–1306 (2011).
175. Campellone, K. G., Robbins, D. & Leong, J. M. EspFU is a translocated EHEC effector that interacts with Tir and N-WASP and promotes Nck-independent actin assembly. *Dev. Cell* **7**, 217–228 (2004).
176. Garmendia, J. *et al.* TccP is an enterohaemorrhagic Escherichia coli O157:H7 type III effector protein that couples Tir to the actin-cytoskeleton. *Cell. Microbiol.* **6**, 1167–1183 (2004).
177. Tzipori, S. *et al.* The Role of the eaeA Gene in Diarrhea and Neurological Complications in a Gnotobiotic Piglet Model of Enterohemorrhagic Escherichia coli Infection. *Infection and Immunity* **63**, 7 (1995).
178. Beebakhee, G., Louie, M., De Azavedo, J. & Brunton, J. Cloning and nucleotide sequence of the eae gene homologue from enterohemorrhagic Escherichia coli serotype O157:H7. *FEMS Microbiol. Lett.* **70**, 63–68 (1992).
179. Yu, J. & Kaper, J. B. Cloning and characterization of the eae gene of enterohaemorrhagic Escherichia coli O157:H7. *Mol. Microbiol.* **6**, 411–417 (1992).
180. Donnenberg, M. S. *et al.* The role of the eae gene of enterohemorrhagic Escherichia coli in intimate attachment in vitro and in a porcine model. *Journal of Clinical Investigation* **92**, 1418–1424 (1993).
181. McKee, M. L., Melton-Celsa, A. R., Moxley, R. A., Francis, D. H. & O'Brien, A. D. Enterohemorrhagic Escherichia coli O157:H7 requires intimin to colonize the gnotobiotic pig intestine and to adhere to HEp-2 cells. *Infect. Immun.* **63**, 3739–3744 (1995).
182. Dean-Nystrom, E. A., Bosworth, B. T., Moon, H. W. & O'Brien, A. D. Escherichia coli O157:H7 Requires Intimin for Enteropathogenicity in Calves. *Infection and Immunity* **66**, 4560–4563 (1998).
183. Ritchie, J. M. *et al.* EspFU, a type III-translocated effector of actin assembly, fosters epithelial association and late-stage intestinal colonization by E. coli O157:H7. *Cell Microbiol* **10**, 836–847 (2008).
184. Nguyen, Y. & Sperandio, V. Enterohemorrhagic E. coli (EHEC) pathogenesis. *Frontiers in Cellular and Infection Microbiology* **2**, 1–7 (2012).
185. Connolly, J. P. R., Finlay, B. B. & Roe, A. J. From ingestion to colonization: the influence of the host environment on regulation of the LEE encoded type III secretion system in enterohaemorrhagic Escherichia coli. *Frontiers in Microbiology* **6**, (2015).
186. Mckee, M. L. & O'Brien, A. D. Investigation of Enterohemorrhagic Escherichia coli O157:H7 Adherence Characteristics and Invasion Potential Reveals a New Attachment Pattern Shared by Intestinal E. coli. *Infection and Immunity* **63**, 2070–2074 (1995).
187. McDaniel, T. K., Jarvis, K. G., Donnenberg, M. S. & Kaper, J. B. A genetic locus of enterocyte effacement conserved among diverse enterobacterial pathogens. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 1664–1668 (1995).
188. Furniss, R. C. D. & Clements, A. Regulation of the Locus of Enterocyte Effacement in Attaching and Effacing Pathogens. *J Bacteriol* **200**, (2017).
189. Barnett Foster, D. Modulation of the enterohemorrhagic *E. coli* virulence program through the human gastrointestinal tract. *Virulence* **4**, 315–323 (2013).

190. Sistrunk, J. R., Nickerson, K. P., Chanin, R. B., Rasko, D. A. & Faherty, C. S. Survival of the Fittest: How Bacterial Pathogens Utilize Bile To Enhance Infection. *Clin Microbiol Rev* **29**, 819–836 (2016).
191. Sengupta, C., Ray, S. & Chowdhury, R. Fine tuning of virulence regulatory pathways in enteric bacteria in response to varying bile and oxygen concentrations in the gastrointestinal tract. *Gut Pathog* **6**, 38 (2014).
192. Sperandio, V., Mellies, J. L., Nguyen, W., Shin, S. & Kaper, J. B. Quorum sensing controls expression of the type III secretion gene transcription and protein secretion in enterohemorrhagic and enteropathogenic Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 15196–15201 (1999).
193. Moreira, C. G. *et al.* Bacterial Adrenergic Sensors Regulate Virulence of Enteric Pathogens in the Gut. *mBio* **7**, (2016).
194. Hughes, D. T., Clarke, M. B., Yamamoto, K., Rasko, D. A. & Sperandio, V. The QseC Adrenergic Signaling Cascade in Enterohemorrhagic E. coli (EHEC). *PLoS Pathogens* **5**, e1000553 (2009).
195. Kendall, M. M., Rasko, D. A. & Sperandio, V. Global Effects of the Cell-to-Cell Signaling Molecules Autoinducer-2, Autoinducer-3, and Epinephrine in a luxS Mutant of Enterohemorrhagic Escherichia coli. *Infection and Immunity* **75**, 4875–4884 (2007).
196. Reading, N. C., Rasko, D. A., Torres, A. G. & Sperandio, V. The two-component system QseEF and the membrane protein QseG link adrenergic and stress sensing to bacterial pathogenesis. *PNAS* **106**, 5889–5894 (2009).
197. Kendall, M. Interkingdom Chemical Signaling in Enterohemorrhagic Escherichia coli O157:H7. in *Microbial Endocrinology: Interkingdom SIgnaling in Infectious Disease and Health*.
198. Moreira, C. G. & Sperandio, V. The Epinephrine/Norepinephrine/Autoinducer-3 Interkingdom Signaling System in Escherichia coli O157:H7. *Adv Exp Med Biol* **874**, 247–261 (2016).
199. Ellermann, M. *et al.* Endocannabinoids Inhibit the Induction of Virulence in Enteric Pathogens. *Cell* (2020) doi:10.1016/j.cell.2020.09.022.
200. Kumar, A. *et al.* The Serotonin Neurotransmitter Modulates Virulence of Enteric Pathogens. *Cell Host Microbe* **28**, 41-53.e8 (2020).
201. Kumar, A., Ellermann, M. & Sperandio, V. Taming the Beast: Interplay between Gut Small Molecules and Enteric Pathogens. *Infect Immun* **87**, (2019).
202. S), A. M. R. I. for I. D. (U & Dembek, Z. F. *Usamriid's Medical Management of Biological Casualties Handbook*. (Government Printing Office, 2016).
203. Trofa, A. F., Ueno-Olsen, H., Oiwa, R. & Yoshikawa, M. Dr. Kiyoshi Shiga: Discoverer of the Dysentery Bacillus. *Clin Infect Dis* **29**, 1303–1306 (1999).
204. Shiga, K. Ueber den Erreger der Dysenterie in Japan, Vorläufige Mitteilung. *Zentralbl Bakteriol Microbiol Hyg* **23**, 599–600 (1898).
205. Neisser, M. & Shiga, K. Ueber freie Receptoren von Typhusund Dysenterie-Bazillen und ueber das Dysenterie Toxin. *Dtsch Med Wochenschr* **29**, 61–62 (1903).
206. Conradi, H. Uber Iosliche, durch asptische Autolyse erhaltene Giftstoffe vonRuhr- und Typhus-Bazillen. *tsch MedWochenschr* **29**, 26–28 (1903).
207. Keusch, G. T. The rediscovery of Shiga toxin and its role in clinical disease. *Jpn. J. Med. Sci. Biol.* **51 Suppl**, S5-22 (1998).
208. Van Heyningen, W. E. & Gladstone, G. P. The neurotoxin of Shigella shigae. I. Production, purification and properties of the toxin. *Br J Exp Pathol* **34**, 202–216 (1953).
209. Howard, J. G. Observations on the Intoxication Produced in Mice and Rabbits by the Neurotoxin of Shigella shigae. *Br J Exp Pathol* **36**, 439–446 (1955).
210. Bridgwater, F. A. J., Morgan, R. S., Rowson, K. E. K. & Payling Wright, G. The Neurotoxin of Shigella shigae. *Br J Exp Pathol* **36**, 447–453 (1955).
211. Keusch, G. T., Grady, G. F., Mata, L. J. & McIver, J. The pathogenesis of Shigella diarrhea. I. Enterotoxin production by Shigella dysenteriae I. *J. Clin. Invest.* **51**, 1212–1218 (1972).
212. Keusch, G. T., Grady, G. F., Takeuchi, A. & Sprinz, H. The Pathogenesis of Shigella Diarrhea. II. Enterotoxin-induced Acute Enteritis in the Rabbit Ileum. *J Infect Dis* **126**, 92–95 (1972).
213. Vicari, G., Olitzki, A. L. & Olitzki, Z. The action of the thermolabile toxin of Shigella dysenteriae on cells cultivated in vitro. *Br J Exp Pathol* **41**, 179–189 (1960).
214. Fontaine, A., Arondel, J. & Sansonetti, P. J. Role of Shiga toxin in the pathogenesis of bacillary dysentery, studied by using a Tox- mutant of Shigella dysenteriae 1. *Infect Immun* **56**, 3099–3109 (1988).
215. Levine, M. M. *et al.* Pathogenesis of Shigella dysenteriae 1 (Shiga) dysentery. *J. Infect. Dis.* **127**, 261–270 (1973).

216. Keusch, G. T., Jacewicz, M., Levine, M. M., Hornick, R. B. & Kochwa, S. Pathogenesis of shigella diarrhea. Serum anticytotoxin antibody response produced by toxigenic and nontoxigenic Shigella dysenteriae 1. *J Clin Invest* **57**, 194–202 (1976).
217. Gemski, P., Takeuchi, A., Washington, O. & Formal, S. B. Shigellosis Due to Shigella dysenteriae 1: Relative Importance of Mucosal Invasion versus Toxin Production in Pathogenesis. *The Journal of Infectious Diseases* **126**, 523–530 (1972).
218. Konowalchuk, J., Speirs, J. I. & Stavric, S. Vero response to a cytotoxin of Escherichia coli. *Infect. Immun.* **18**, 775–779 (1977).
219. Jackson, M. P., Newland, J. W., Holmes, R. K. & O'Brien, A. D. Nucleotide sequence analysis of the structural genes for Shiga-like toxin I encoded by bacteriophage 933J from Escherichia coli. *Microb. Pathog.* **2**, 147–153 (1987).
220. Calderwood, S. B., Auclair, F., Donohue-Rolfe, A., Keusch, G. T. & Mekalanos, J. J. Nucleotide sequence of the Shiga-like toxin genes of Escherichia coli. *PNAS* **84**, 4364–4368 (1987).
221. Takao, T. *et al.* Identity of molecular structure of Shiga-like toxin I (VT1) from Escherichia coli O157:H7 with that of Shiga toxin. *Microb. Pathog.* **5**, 57–69 (1988).
222. Strockbine, N. A., Jackson, M. P., Sung, L. M., Holmes, R. K. & O'Brien, A. D. Cloning and sequencing of the genes for Shiga toxin from Shigella dysenteriae type 1. *J. Bacteriol.* **170**, 1116–1122 (1988).
223. Calderwood, S. Proposed new nomenclature for SLT(VT) family. *ASM News* **62**, 118–119 (1996).
224. Luna-Gierke, R. E. *et al.* Outbreaks of non-O157 Shiga toxin-producing Escherichia coli infection: USA. *Epidemiol. Infect.* **142**, 2270–2280 (2014).
225. Friedrich, A. W. *et al.* Escherichia coli harboring Shiga toxin 2 gene variants: frequency and association with clinical symptoms. *J. Infect. Dis.* **185**, 74–84 (2002).
226. Melton-Celsa, A. R. Shiga Toxin (Stx) Classification, Structure, and Function. *Microbiology Spectrum* **2**, (2014).
227. Karmali, M. A. *et al.* The association between idiopathic hemolytic uremic syndrome and infection by verotoxin-producing Escherichia coli. *J. Infect. Dis.* **151**, 775–782 (1985).
228. Blaser, M. J. Bacteria and diseases of unknown cause: hemolytic-uremic syndrome. *J. Infect. Dis.* **189**, 552–555 (2004).
229. Wagner, P. L. *et al.* Role for a Phage Promoter in Shiga Toxin 2 Expression from a Pathogenic Escherichia coliStrain. *Journal of Bacteriology* **183**, 2081–2085 (2001).
230. Wagner, P. L. *et al.* Bacteriophage control of Shiga toxin 1 production and release by Escherichia coli. *Molecular Microbiology* **44**, 957–970 (2002).
231. Łoś, J. M., Łoś, M., Węgrzyn, A. & Węgrzyn, G. Altruism of Shiga toxin-producing Escherichia coli: recent hypothesis versus experimental results. *Front Cell Infect Microbiol* **2**, (2013).
232. Beddoe, T., Paton, A. W., Le Nours, J., Rossjohn, J. & Paton, J. C. Structure, Biological Functions and Applications of the AB5 Toxins. *Trends Biochem Sci* **35**, 411–418 (2010).
233. Jacewicz, M. S. *et al.* Pathogenesis of Shigella diarrhea: XVII. A mammalian cell membrane glycolipid, Gb3, is required but not sufficient to confer sensitivity to Shiga toxin. *J. Infect. Dis.* **169**, 538–546 (1994).
234. Jacewicz, M., Feldman, H. A., Donohue-Rolfe, A., Balasubramanian, K. A. & Keusch, G. T. Pathogenesis of Shigella diarrhea. XIV. Analysis of Shiga toxin receptors on cloned HeLa cells. *J. Infect. Dis.* **159**, 881–889 (1989).
235. Lindberg, A. A. *et al.* Identification of the carbohydrate receptor for Shiga toxin produced by Shigella dysenteriae type 1. *J. Biol. Chem.* **262**, 1779–1785 (1987).
236. Jacewicz, M., Clausen, H., Nudelman, E., Donohue-Rolfe, A. & Keusch, G. T. Pathogenesis of shigella diarrhea. XI. Isolation of a shigella toxin-binding glycolipid from rabbit jejunum and HeLa cells and its identification as globotriaosylceramide. *J. Exp. Med.* **163**, 1391–1404 (1986).
237. Nyholm, P.-G. *et al.* Two distinct binding sites for globotriaosyl ceramide on verotoxins: identification by molecular modelling and confirmation using deoxy analogues and a new glycolipid receptor for all verotoxins. *Chemistry & Biology* **3**, 263–275 (1996).
238. Stein, P. E., Boodhoo, A., Tyrrell, G. J., Brunton, J. L. & Read, R. J. Crystal structure of the cell-binding B oligomer of verotoxin-1 from E. coli. *Nature* **355**, 748–750 (1992).
239. Fraser, M. E., Chernaia, M. M., Kozlov, Y. V. & James, M. N. Crystal structure of the holotoxin from Shigella dysenteriae at 2.5 A resolution. *Nat. Struct. Biol.* **1**, 59–64 (1994).
240. Richardson, J. M., Evans, P. D., Homans, S. W. & Donohue-Rolfe, A. Solution structure of the carbohydrate-binding B-subunit homopentamer of verotoxin VT-1 from E. coli. *Nature Structural Biology* **4**, 190–193 (1997).
241. Shimizu, H., Field, R. A., Homans, S. W. & Donohue-Rolfe, A. Solution Structure of the Complex between the B-Subunit Homopentamer of Verotoxin VT-1 from Escherichia coli and the Trisaccharide Moiety of Globotriaosylceramide. *Biochemistry* **37**, 11078–11082 (1998).

242. Ling, H. *et al.* Structure of the shiga-like toxin I B-pentamer complexed with an analogue of its receptor Gb3. *Biochemistry* **37**, 1777–1788 (1998).
243. Lencer, W. I. & Saslowsky, D. Raft trafficking of AB5 subunit bacterial toxins. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1746**, 314–321 (2005).
244. Sandvig, K., Bergan, J., Dyve, A.-B., Skotland, T. & Torgersen, M. L. Endocytosis and retrograde transport of Shiga toxin. *Toxicon* **56**, 1181–1185 (2010).
245. Cho, J. A. *et al.* Insights on the trafficking and retro-translocation of glycosphingolipid-binding bacterial toxins. *Front. Cell. Infect. Microbiol.* 2, (2012).
246. Lord, J. M., Roberts, L. M. & Lencer, W. I. Entry of protein toxins into mammalian cells by crossing the endoplasmic reticulum membrane: co-opting basic mechanisms of endoplasmic reticulum-associated degradation. *Curr Top Microbiol Immunol* **300**, 149–168 (2005).
247. Endo, Y., Mitsui, K., Motizuki, M. & Tsurugi, K. The mechanism of action of ricin and related toxic lectins on eukaryotic ribosomes. The site and the characteristics of the modification in 28 S ribosomal RNA caused by the toxins. *J. Biol. Chem.* 262, 5908–5912 (1987).
248. Endo, Y. *et al.* Site of action of a Vero toxin (VT2) from Escherichia coli O157:H7 and of Shiga toxin on eukaryotic ribosomes. RNA N-glycosidase activity of the toxins. *Eur. J. Biochem.* **171**, 45–50 (1988).
249. Menge, C. Molecular Biology of Escherichia coli Shiga Toxins' Effects on Mammalian Cells. *Toxins* 12, 345 (2020).
250. Tesh, V. L. Activation of cell stress response pathways by Shiga toxins. *Cellular Microbiology* **14**, 1–9 (2012).
251. Jandhyala, D. M. *et al.* Activation of the Classical Mitogen-Activated Protein Kinases Is Part of the Shiga Toxin-Induced Ribotoxic Stress Response and May Contribute to Shiga Toxin-Induced Inflammation. *Infection and Immunity* 84, 138–148 (2016).
252. Iordanov, M. S. *et al.* Ribotoxic stress response: activation of the stress-activated protein kinase JNK1 by inhibitors of the peptidyl transferase reaction and by sequence-specific RNA damage to the alpha-sarcin/ricin loop in the 28S rRNA. *Molecular and Cellular Biology* **17**, 3373–3381 (1997).
253. Thorpe, C. M., Smith, W. E., Hurley, B. P. & Acheson, D. W. Shiga toxins induce, superinduce, and stabilize a variety of C-X-C chemokine mRNAs in intestinal epithelial cells, resulting in increased chemokine expression. *Infect. Immun.* 69, 6140–6147 (2001).
254. Schüller, S., Frankel, G. & Phillips, A. D. Interaction of Shiga toxin from Escherichia coli with human intestinal epithelial cell lines and explants: Stx2 induces epithelial damage in organ culture. *Cellular Microbiology* 6, 289–301 (2004).
255. Ramegowda, B. & Tesh, V. L. Differentiation-associated toxin receptor modulation, cytokine production, and sensitivity to Shiga-like toxins in human monocytes and monocytic cell lines. *Infection and Immunity* 64, 8 (1996).
256. Thorpe, C. M. *et al.* Shiga Toxins Stimulate Secretion of Interleukin-8 from Intestinal Epithelial Cells. *Infect Immun* 67, 5985–5993 (1999).
257. Smith, W. E. *et al.* Shiga Toxin 1 Triggers a Ribotoxic Stress Response Leading to p38 and JNK Activation and Induction of Apoptosis in Intestinal Epithelial Cells. *Infect Immun* **71**, 1497–1504 (2003).
258. Cameron, P., Smith, S. J., Giembycz, M. A., Rotondo, D. & Plevin, R. Verotoxin activates mitogen-activated protein kinase in human peripheral blood monocytes: role in apoptosis and proinflammatory cytokine release. *Br. J. Pharmacol.* **140**, 1320–1330 (2003).
259. Harrison, L. M., Hoogen, C. van den, Haaften, W. C. E. van & Tesh, V. L. Chemokine Expression in the Monocytic Cell Line THP-1 in Response to Purified Shiga Toxin 1 and/or Lipopolysaccharides. *Infection and Immunity* 73, 403–412 (2005).
260. Lee, M.-S., Kim, M. H. & Tesh, V. L. Shiga toxins expressed by human pathogenic bacteria induce immune responses in host cells. *J. Microbiol.* **51**, 724–730 (2013).
261. Brigotti, M. *et al.* Human monocytes stimulated by Shiga toxin 1a via globotriaosylceramide release proinflammatory molecules associated with hemolytic uremic syndrome. *Int. J. Med. Microbiol.* 308, 940–946 (2018).
262. Pradhan, S. *et al.* Tissue Responses to Shiga Toxin in Human Intestinal Organoids. *Cellular and Molecular Gastroenterology and Hepatology* **10**, 171–190 (2020).
263. Karve, S. S., Pradhan, S., Ward, D. V. & Weiss, A. A. Intestinal organoids model human responses to infection by commensal and Shiga toxin producing Escherichia coli. *PLOS ONE* 12, e0178966 (2017).
264. Sandvig, K. & van Deurs, B. Toxin-induced cell lysis: Protection by 3-methyladenine and cycloheximide. *Experimental Cell Research* 200, 253–262 (1992).
265. Obrig, T. G. *et al.* Direct cytotoxic action of Shiga toxin on human vascular endothelial cells. *Infection and Immunity* 56, 2373–2378 (1988).
266. Griffin, P. M., Olmstead, L. C. & Petras, R. E. Escherichia coli O157:H7-associated colitis. A clinical and histological study of 11 cases. *Gastroenterology* 99, 142–149 (1990).

267. Kelly, J., Oryshak, A., Wenetsek, M., Grabiec, J. & Handy, S. The colonic pathology of Escherichia coli O157:H7 infection. *Am. J. Surg. Pathol.* **14**, 87–92 (1990).
268. Holgersson, J., Strömberg, N. & Breimer, M. E. Glycolipids of human large intestine: difference in glycolipid expression related to anatomical localization, epithelial/non-epithelial tissue and the ABO, Le and Se phenotypes of the donors. *Biochimie* **70**, 1565–1574 (1988).
269. Holgersson, J., Jovall, P.-Å. & Breimer, M. E. Glycosphingolipids of Human Large Intestine: Detailed Structural Characterization with Special Reference to Blood Group Compounds and Bacterial Receptor Structures1. *The Journal of Biochemistry* **110**, 120–131 (1991).
270. Zumbrun, S. D. *et al.* Human Intestinal Tissue and Cultured Colonic Cells Contain Globotriaosylceramide Synthase mRNA and the Alternate Shiga Toxin Receptor Globotetraosylceramide. *Infection and Immunity* **78**, 4488–4499 (2010).
271. Karve, S. S. & Weiss, A. A. Glycolipid Binding Preferences of Shiga Toxin Variants. *PLoS One* **9**, e101173 (2014).
272. Malyukova, I. *et al.* Macropinocytosis in Shiga toxin 1 uptake by human intestinal epithelial cells and transcellular transcytosis. *Am J Physiol Gastrointest Liver Physiol* **296**, G78–G92 (2009).
273. Acheson, D. W. *et al.* Translocation of Shiga toxin across polarized intestinal cells in tissue culture. *Infect Immun* **64**, 3294–3300 (1996).
274. Brigotti, M. *et al.* Identification of TLR4 as the receptor that recognizes Shiga toxins in human neutrophils. *J. Immunol.* **191**, 4748–4758 (2013).
275. Ståhl, A. *et al.* A novel mechanism of bacterial toxin transfer within host blood cell-derived microvesicles. *PLoS Pathog.* **11**, e1004619 (2015).
276. Karpman, D. & Ståhl, A.-L. Enterohemorrhagic Escherichia coli Pathogenesis and the Host Response. *Microbiology Spectrum* **2**, (2014).
277. Cordonnier, C. *et al.* Enterohemorrhagic Escherichia coli pathogenesis: role of Long polar fimbriae in Peyer's patches interactions. *Scientific Reports* **7**, 44655 (2017).
278. Phillips, A. *et al.* Enterohaemorrhagic Escherichia coli O157:H7 target Peyer's patches in humans and cause attaching/effacing lesions in both human and bovine intestine. *Gut* **47**, 377–381 (2000).
279. Pearson, J. S. & Hartland, E. L. The Inflammatory Response during Enterohemorrhagic Escherichia coli Infection. *Microbiology Spectrum* **2**, (2014).
280. Yamasaki, C. *et al.* Induction of cytokines in a human colon epithelial cell line by Shiga toxin 1 (Stx1) and Stx2 but not by non-toxic mutant Stx1 which lacks N-glycosidase activity. *FEBS Letters* **442**, 231–234 (1999).
281. Berin, M. C., Darfeuille-Michaud, A., Egan, L. J., Miyamoto, Y. & Kagnoff, M. F. Role of EHEC O157:H7 virulence factors in the activation of intestinal epithelial cell NF-kappaB and MAP kinase pathways and the upregulated expression of interleukin 8. *Cell. Microbiol.* **4**, 635–648 (2002).
282. Dahan, S. *et al.* Enterohemorrhagic Escherichia coli infection induces interleukin-8 production via activation of mitogen-activated protein kinases and the transcription factors NF-kappaB and AP-1 in T84 cells. *Infect. Immun.* **70**, 2304–2310 (2002).
283. Hurley, B. P., Thorpe, C. M. & Acheson, D. W. Shiga toxin translocation across intestinal epithelial cells is enhanced by neutrophil transmigration. *Infect. Immun.* **69**, 6148–6155 (2001).
284. Stekel, D. J. *et al.* Analysis of host response to bacterial infection using error model based gene expression microarray experiments. *Nucleic Acids Res* **33**, e53–e53 (2005).
285. Gobert, A. P. *et al.* Shiga Toxin Produced by Enterohemorrhagic Escherichia coli Inhibits PI3K/NF-κB Signaling Pathway in Globotriaosylceramide-3-Negative Human Intestinal Epithelial Cells. *The Journal of Immunology* **178**, 8168–8174 (2007).
286. Kim, Y., Oh, S., Park, S. & Kim, S. H. Interactive transcriptome analysis of enterohemorrhagic Escherichia coli (EHEC) O157:H7 and intestinal epithelial HT-29 cells after bacterial attachment. *Int. J. Food Microbiol.* **131**, 224–232 (2009).
287. Bruballa, A. C. *et al.* Role of Shiga Toxins in Cytotoxicity and Immunomodulatory Effects of Escherichia coli O157:H7 during Host-Bacterial Interactions in vitro. *Toxins* **12**, 48 (2020).
288. Okumura, R. & Takeda, K. Roles of intestinal epithelial cells in the maintenance of gut homeostasis. *Experimental & Molecular Medicine* **49**, e338–e338 (2017).
289. Kelly, J. K. *et al.* The histopathology of rectosigmoid biopsies from adults with bloody diarrhea due to verotoxin-producing Escherichia coli. *Am. J. Clin. Pathol.* **88**, 78–82 (1987).
290. Shigeno, T., Akamatsu, T., Fujimori, K., Nakatsuji, Y. & Nagata, A. The clinical significance of colonoscopy in hemorrhagic colitis due to enterohemorrhagic Escherichia coli O157:H7 infection. *Endoscopy* **34**, 311–314 (2002).
291. Shigeno, T., Akamatsu, T., Fujimori, K., Nakatsuji, Y. & Nakamura, Y. Evaluation of Colonoscopic Findings in Patients with Diarrheagenic Escherichia Coli-Induced Hemorrhagic Colitis. *Digestive Endoscopy* **20**, 123–129 (2008).

292. Land, W. G. The Role of Damage-Associated Molecular Patterns (DAMPs) in Human Diseases. *Sultan Qaboos Univ Med J* **15**, e157–e170 (2015).
293. van Setten, P. A., Monnens, L. A., Verstraten, R. G., van den Heuvel, L. P. & van Hinsbergh, V. W. Effects of verocytotoxin-1 on nonadherent human monocytes: binding characteristics, protein synthesis, and induction of cytokine release. *Blood* **88**, 174–183 (1996).
294. Sakiri, R., Ramegowda, B. & Tesh, V. L. Shiga toxin type 1 activates tumor necrosis factor-alpha gene transcription and nuclear translocation of the transcriptional activators nuclear factor-kappaB and activator protein-1. *Blood* **92**, 558–566 (1998).
295. Cherla, R. P., Lee, S.-Y., Mees, P. L. & Tesh, V. L. Shiga toxin 1-induced cytokine production is mediated by MAP kinase pathways and translation initiation factor eIF4E in the macrophage-like THP-1 cell line. *J. Leukoc. Biol.* **79**, 397–407 (2006).
296. Foster, G. H. & Tesh, V. L. Shiga toxin 1-induced activation of c-Jun NH(2)-terminal kinase and p38 in the human monocytic cell line THP-1: possible involvement in the production of TNF-alpha. *J. Leukoc. Biol.* **71**, 107–114 (2002).
297. Brandelli, J. R., Griener, T. P., Laing, A., Mulvey, G. & Armstrong, G. D. The Effects of Shiga Toxin 1, 2 and Their Subunits on Cytokine and Chemokine Expression by Human Macrophage-Like THP-1 Cells. *Toxins (Basel)* **7**, 4054–4066 (2015).
298. Leyva-Illades, D., Cherla, R. P., Galindo, C. L., Chopra, A. K. & Tesh, V. L. Global transcriptional response of macrophage-like THP-1 cells to Shiga toxin type 1. *Infect. Immun.* **78**, 2454–2465 (2010).
299. Lee, M.-S. & Tesh, V. L. Roles of Shiga Toxins in Immunopathology. *Toxins (Basel)* **11**, (2019).
300. Arfilli, V. *et al.* Shiga toxin 1 and ricin A chain bind to human polymorphonuclear leucocytes through a common receptor. *Biochem. J.* **432**, 173–180 (2010).
301. Tazzari, P. L. *et al.* Flow cytometry detection of Shiga toxins in the blood from children with hemolytic uremic syndrome. *Cytometry B Clin Cytom* **61**, 40–44 (2004).
302. Te Loo, D. M., van Hinsbergh, V. W., van den Heuvel, L. P. & Monnens, L. A. Detection of verocytotoxin bound to circulating polymorphonuclear leukocytes of patients with hemolytic uremic syndrome. *J. Am. Soc. Nephrol.* **12**, 800–806 (2001).
303. Ståhl, A., Sartz, L., Nelsson, A., Békássy, Z. D. & Karpman, D. Shiga toxin and lipopolysaccharide induce platelet-leukocyte aggregates and tissue factor release, a thrombotic mechanism in hemolytic uremic syndrome. *PLoS ONE* **4**, e6990 (2009).
304. Robson, W. L., Fick, G. H. & Wilson, P. C. Prognostic factors in typical postdiarrhea hemolytic-uremic syndrome. *Child Nephrol Urol* **9**, 203–207 (1988).
305. Walters, M. D., Matthei, I. U., Kay, R., Dillon, M. J. & Barratt, T. M. The polymorphonuclear leucocyte count in childhood haemolytic uraemic syndrome. *Pediatr Nephrol* **3**, 130–134 (1989).
306. Li, Y., Frey, E., Mackenzie, A. M. R. & Finlay, B. B. Human Response to Escherichia coli O157:H7 Infection: Antibodies to Secreted Virulence Factors. *Infect Immun* **68**, 5090–5095 (2000).
307. Karpman, D. *et al.* Antibodies to intimin and Escherichia coli secreted proteins A and B in patients with enterohemorrhagic Escherichia coli infections. *Pediatr Nephrol* **17**, 201–211 (2002).
308. Ludwig, K. *et al.* Antibody Response to Shiga Toxins Stx2 and Stx1 in Children with Enteropathic Hemolytic-Uremic Syndrome. *J Clin Microbiol* **39**, 2272–2279 (2001).
309. Geelen, J. M., van der Velden, T. J. A. M., van den Heuvel, L. P. W. J. & Monnens, L. A. H. Interactions of Shiga-like toxin with human peripheral blood monocytes. *Pediatr. Nephrol.* **22**, 1181–1187 (2007).
310. Bitzan, M. *et al.* Evidence that verotoxins (Shiga-like toxins) from Escherichia coli bind to P blood group antigens of human erythrocytes in vitro. *Infect. Immun.* **62**, 3337–3347 (1994).
311. Betz, J. *et al.* Shiga toxin of enterohaemorrhagic Escherichia coli directly injures developing human erythrocytes. *Cell. Microbiol.* **18**, 1339–1348 (2016).
312. Ghosh, S. A., Polanowska-Grabowska, R. K., Fujii, J., Obrig, T. & Gear, A. R. L. Shiga toxin binds to activated platelets. *J. Thromb. Haemost.* **2**, 499–506 (2004).
313. Cooling, L. L., Walker, K. E., Gille, T. & Koerner, T. A. Shiga toxin binds human platelets via globotriaosylceramide (Pk antigen) and a novel platelet glycosphingolipid. *Infect. Immun.* **66**, 4355–4366 (1998).
314. Brigotti, M. *et al.* Particulate Shiga Toxin 2 in Blood is Associated to the Development of Hemolytic Uremic Syndrome in Children. *Thrombosis and haemostasis* **120**, 107–120 (2020).
315. Williams, J. M. *et al.* A comparison of the effects of verocytotoxin-1 on primary human renal cell cultures. *Toxicol. Lett.* **105**, 47–57 (1999).
316. Obrig, T. G. & Karpman, D. Shiga Toxin Pathogenesis: Kidney Complications and Renal Failure. *Curr Top Microbiol Immunol* **357**, 105–136 (2012).

317. Zoja, C., Buelli, S. & Morigi, M. Shiga toxin-associated hemolytic uremic syndrome: pathophysiology of endothelial dysfunction. *Pediatr Nephrol* **25**, 2231–2240 (2010).
318. Obrig, T. G. Escherichia coli Shiga Toxin Mechanisms of Action in Renal Disease. *Toxins* 2, 2769–2794 (2010).
319. Bauwens, A. *et al.* Facing glycosphingolipid–Shiga toxin interaction: dire straits for endothelial cells of the human vasculature. *Cell. Mol. Life Sci.* **70**, 425–457 (2013).
320. Exeni, R. A. *et al.* Pathogenic role of inflammatory response during Shiga toxin-associated hemolytic uremic syndrome (HUS). *Pediatr. Nephrol.* **33**, 2057–2071 (2018).
321. Louise, C. B. & Obrig, T. G. Shiga toxin-associated hemolytic uremic syndrome: combined cytotoxic effects of shiga toxin and lipopolysaccharide (endotoxin) on human vascular endothelial cells in vitro. *Infection and Immunity* **60**, 1536–1543 (1992).
322. van de Kar, N. C., Monnens, L. A., Karmali, M. A. & van Hinsbergh, V. W. Tumor necrosis factor and interleukin-1 induce expression of the verocytotoxin receptor globotriaosylceramide on human endothelial cells: implications for the pathogenesis of the hemolytic uremic syndrome. *Blood* **80**, 2755–2764 (1992).
323. Stricklett, P. K., Hughes, A. K., Ergonul, Z. & Kohan, D. E. Molecular Basis for Up-Regulation by Inflammatory Cytokines of Shiga Toxin 1 Cytotoxicity and Globotriaosylceramide Expression. *J Infect Dis* **186**, 976–982 (2002).
324. Villysson, A., Tontanahal, A. & Karpman, D. Microvesicle Involvement in Shiga Toxin-Associated Infection. *Toxins (Basel)* **9**, (2017).
325. Karpman, D., Sartz, L. & Johnson, S. Pathophysiology of typical hemolytic uremic syndrome. *Semin Thromb Hemost* **36**, 575–585 (2010).
326. Hall, G., Kurosawa, S. & Stearns-Kurosawa, D. J. Shiga Toxin Therapeutics: Beyond Neutralization. *Toxins (Basel)* **9**, (2017).
327. Melton-Celsa, A. R. & O'Brien, A. D. New Therapeutic Developments against Shiga Toxin-Producing Escherichia coli. *Microbiol Spectr* 2, (2014).
328. Calderwood, S. B. Clinical manifestations, diagnosis and treatment of enterohemorrhagic Escherichia coli (EHEC) infection. in *UpToDate* (2017).
329. Trachtman, H. *et al.* Effect of an oral Shiga toxin-binding agent on diarrhea-associated hemolytic uremic syndrome in children: a randomized controlled trial. *JAMA* **290**, 1337–1344 (2003).
330. Stechmann, B. *et al.* Inhibition of retrograde transport protects mice from lethal ricin challenge. *Cell* **141**, 231–242 (2010).
331. Secher, T. *et al.* Retrograde Trafficking Inhibitor of Shiga Toxins Reduces Morbidity and Mortality of Mice Infected with Enterohemorrhagic Escherichia coli. *Antimicrob Agents Chemother* **59**, 5010–5013 (2015).
332. Mukhopadhyay, S. & Linstedt, A. D. Manganese blocks intracellular trafficking of Shiga toxin and protects against Shiga toxicosis. *Science* **335**, 332–335 (2012).
333. Mukhopadhyay, S., Redler, B. & Linstedt, A. D. Shiga toxin-binding site for host cell receptor GPP130 reveals unexpected divergence in toxin-trafficking mechanisms. *Mol Biol Cell* **24**, 2311–2318 (2013).
334. Havers, F. P. Use of Tetanus Toxoid, Reduced Diphtheria Toxoid, and Acellular Pertussis Vaccines: Updated Recommendations of the Advisory Committee on Immunization Practices — United States, 2019. *MMWR Morb Mortal Wkly Rep* **69**, (2020).
335. Dowling, T. C. *et al.* Phase 1 safety and pharmacokinetic study of chimeric murine-human monoclonal antibody c alpha Stx2 administered intravenously to healthy adult volunteers. *Antimicrob. Agents Chemother.* **49**, 1808–1812 (2005).
336. López, E. L. *et al.* Safety and pharmacokinetics of urtoxazumab, a humanized monoclonal antibody, against Shiga-like toxin 2 in healthy adults and in pediatric patients infected with Shiga-like toxin-producing Escherichia coli. *Antimicrob. Agents Chemother.* **54**, 239–243 (2010).
337. Szu, S. C. & Ahmed, A. Clinical Studies of Escherichia coli O157:H7 Conjugate Vaccines in Adults and Young Children. *Microbiology Spectrum* 2, (2014).
338. Ahmed, A., Li, J., Shiloach, Y., Robbins, J. B. & Szu, S. C. Safety and immunogenicity of Escherichia coli O157 O-specific polysaccharide conjugate vaccine in 2-5-year-old children. *J Infect Dis* **193**, 515–521 (2006).
339. Konadu, E. Y. *et al.* Investigational vaccine for Escherichia coli O157: phase 1 study of O157 O-specific polysaccharide-Pseudomonas aeruginosa recombinant exoprotein A conjugates in adults. *J Infect Dis* **177**, 383–387 (1998).
340. Konadu, E. *et al.* Syntheses and immunologic properties of Escherichia coli O157 O-specific polysaccharide and Shiga toxin 1 B subunit conjugates in mice. *Infect Immun* **67**, 6191–6193 (1999).
341. Marcato, P., Griener, T. P., Mulvey, G. L. & Armstrong, G. D. Recombinant Shiga toxin B-subunit-keyhole limpet hemocyanin conjugate vaccine protects mice from Shigatoxemia. *Infect Immun* **73**, 6523–6529 (2005).

342. World Health Organization. *Shiga toxin-producing Escherichia coli (STEC) and food: attribution, characterization, and monitoring*. (2018).
343. Hartland, E. L. & Leong, J. M. Enteropathogenic and enterohemorrhagic E. coli: ecology, pathogenesis, and evolution. *Frontiers in Cellular and Infection Microbiology* **3**, 1–3 (2015).
344. Davis, K. T., Van De Kar, N. C. & Tarr, P. I. Shiga Toxin/Verocytotoxin-Producing Escherichia coli Infections: Practical Clinical Perspectives. *Enterohemorrhagic Escherichia coli and Other Shiga Toxin-Producing E. coli* 321–339 (2014) doi:doi:10.1128/microbiolspec.EHEC-0025-2014.
345. Boyce, T. G., Swerdlow, D. L. & Griffin, P. M. Escherichia coli O157:H7 and the Hemolytic–Uremic Syndrome. *New England Journal of Medicine* **333**, 364–368 (1995).
346. Latif, H., Li, H. J., Charusanti, P., Palsson, B. O. & Aziz, R. K. A Gapless, Unambiguous Genome Sequence of the Enterohemorrhagic Escherichia coli O157:H7 Strain EDL933. *Genome Announcements* 2, (2014).
347. Chattopadhyay, S. & Sokurenko, E. V. Chapter 3 - Evolution of pathogenic Escherichia coli. in *Escherichia coli (Second Edition)* (ed. Donnenberg, M. S.) 45–71 (Academic Press, 2013). doi:10.1016/B978-0-12-397048-0.00003-6.
348. Rasko, D. A. *et al.* The pangenome structure of Escherichia coli: comparative genomic analysis of E. coli commensal and pathogenic isolates. *J. Bacteriol.* **190**, 6881–6893 (2008).
349. Cornick, N., Booher, S. & Moon, H. Intimin Facilitates Colonization by Escherichia coli O157:H7 in Adult Ruminants. *Infection and Immunity* **70**, 2704–2707 (2002).
350. Sperandio, V., Torres, A. G., Jarvis, B., Nataro, J. P. & Kaper, J. B. Bacteria-host communication: The language of hormones. *Proceedings of the National Academy of Sciences* **100**, 8951–8956 (2003).
351. Eckert, S. E. *et al.* Retrospective Application of Transposon-Directed Insertion Site Sequencing to a Library of Signature-Tagged Mini-Tn5Km2 Mutants of Escherichia coli O157:H7 Screened in Cattle. *Journal of Bacteriology* **193**, 1771–1776 (2011).
352. van Opijnen, T., Bodi, K. L. & Camilli, A. Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. *Nature Methods* **6**, 767–772 (2009).
353. Goodman, A. L. *et al.* Identifying Genetic Determinants Needed to Establish a Human Gut Symbiont in Its Habitat. *Cell Host & Microbe* **6**, 279–289 (2009).
354. Gawronski, J. D., Wong, S. M. S., Giannoukos, G., Ward, D. V. & Akerley, B. J. Tracking insertion mutants within libraries by deep sequencing and a genome-wide screen for Haemophilus genes required in the lung. *Proceedings of the National Academy of Sciences* **106**, 16422–16427 (2009).
355. Langridge, G. C. *et al.* Simultaneous assay of every Salmonella Typhi gene using one million transposon mutants. *Genome Research* **19**, 2308–2316 (2009).
356. Gao, B. *et al.* Metabolic and fitness determinants for in vitro growth and intestinal colonization of the bacterial pathogen Campylobacter jejuni. *PLOS Biology* **15**, e2001390 (2017).
357. Cole, B. J. *et al.* Genome-wide identification of bacterial plant colonization genes. *PLOS Biology* **15**, e2002860 (2017).
358. McCarthy, A. J., Stabler, R. A. & Taylor, P. W. Genome-Wide Identification by Transposon Insertion Sequencing of *Escherichia coli* K1 Genes Essential for *In Vitro* Growth, Gastrointestinal Colonizing Capacity, and Survival in Serum. *Journal of Bacteriology* **200**, (2018).
359. de Moraes, M. H. *et al.* Salmonella Persistence in Tomatoes Requires a Distinct Set of Metabolic Functions Identified by Transposon Insertion Sequencing. *Applied and Environmental Microbiology* **83**, (2017).
360. Armbruster, C. E. *et al.* Genome-wide transposon mutagenesis of Proteus mirabilis: Essential genes, fitness factors for catheter-associated urinary tract infection, and the impact of polymicrobial infection on fitness requirements. *PLOS Pathogens* **13**, e1006434 (2017).
361. Capel, E. *et al.* Peripheral blood vessels are a niche for blood-borne meningococci. *Virulence* **8**, 1808–1819 (2017).
362. Zhu, L. *et al.* Novel Genes Required for the Fitness of Streptococcus pyogenes in Human Saliva. 2, 13 (2017).
363. George, A. S. *et al.* Interactions of *Salmonella enterica* Serovar Typhimurium and *Pectobacterium carotovorum* within a Tomato Soft Rot. *Applied and Environmental Microbiology* **84**, (2017).
364. White, K. M. *et al.* A Metagenome-Wide Association Study and Arrayed Mutant Library Confirm *Acetobacter* Lipopolysaccharide Genes Are Necessary for Association with *Drosophila melanogaster*. *G3: Genes|Genomes|Genetics* g3.300530.2017 (2018)
365. Price, M. N. *et al.* Mutant phenotypes for thousands of bacterial genes of unknown function. *Nature* **557**, 503–509 (2018).
366. Pritchard, J. R. *et al.* ARTIST: High-Resolution Genome-Wide Assessment of Fitness Using Transposon-Insertion Sequencing. *PLoS Genetics* **10**, e1004782 (2014).
367. Chao, M. C., Abel, S., Davis, B. M. & Waldor, M. K. The design and analysis of transposon insertion sequencing experiments. *Nature Reviews Microbiology* **14**, 119–128 (2016).

368. Kimura, S., Hubbard, T. P., Davis, B. M. & Waldor, M. K. The Nucleoid Binding Protein H-NS Biases Genome-Wide Transposon Insertion Landscapes. *mBio* **7**, (2016).
369. DeJesus, M. A. *et al.* Comprehensive Essentiality Analysis of the *Mycobacterium tuberculosis* Genome via Saturating Transposon Mutagenesis. *mBio* **8**, (2017).
370. Munoz-Lopez, M. & Garcia-Perez, J. DNA Transposons: Nature and Applications in Genomics. *Current Genomics* **11**, 115–128 (2010).
371. Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Research* **44**, D457–D462 (2016).
372. Tatusov, R. L. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Research* **28**, 33–36 (2000).
373. Galperin, M. Y., Makarova, K. S., Wolf, Y. I. & Koonin, E. V. Expanded microbial genome coverage and improved protein family annotation in the COG database. *Nucleic Acids Research* **43**, D261–D269 (2015).
374. Shields, R. C., Zeng, L., Culp, D. J. & Burne, R. A. Genomewide Identification of Essential Genes and Fitness Determinants of *Streptococcus mutans* UA159. *mSphere* **3**, e00031-18 (2018).
375. Luo, H., Lin, Y., Gao, F., Zhang, C.-T. & Zhang, R. DEG 10, an update of the database of essential genes that includes both protein-coding genes and noncoding genomic elements: Table 1. *Nucleic Acids Research* **42**, D574–D580 (2014).
376. Hubbard, T. P. *et al.* Genetic analysis of *Vibrio parahaemolyticus* intestinal colonization. *Proceedings of the National Academy of Sciences* **113**, 6283–6288 (2016).
377. Blattner, F. *et al.* The Complete Genome Sequence of Escherichia coli K-12. *Science* **277**, 1453–1462 (1997).
378. Baba, T. *et al.* Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular Systems Biology* **2**, (2006).
379. Kato, J. & Hashimoto, M. Construction of consecutive deletions of the Escherichia coli chromosome. *Molecular Systems Biology* **3**, (2007).
380. Yamamoto, N. *et al.* Update on the Keio collection of Escherichia coli single-gene deletion mutants. *Molecular Systems Biology* **5**, (2009).
381. Sperandeo, P., Pozzi, C., Dehò, G. & Polissi, A. Non-essential KDO biosynthesis and new essential cell envelope biogenesis genes in the Escherichia coli yrbG–yhbG locus. *Research in Microbiology* **157**, 547–558 (2006).
382. Goodall, E. C. A. *et al.* The Essential Genome of Escherichia coli K-12. **9**, 18 (2018).
383. Abel, S. & Waldor, M. K. Infant Rabbit Model for Diarrheal Diseases: Infant Rabbit Model for Diarrheal Diseases. *Current Protocols in Microbiology* 6A.6.1-6A.6.15 (2015).
384. Ritchie, J. M., Rui, H., Bronson, R. T. & Waldor, M. K. Back to the Future: Studying Cholera Pathogenesis Using Infant Rabbits. *mBio* **1**, (2010).
385. Abel, S., Abel zur Wiesch, P., Davis, B. M. & Waldor, M. K. Analysis of Bottlenecks in Experimental Models of Infection. *PLOS Pathogens* **11**, e1004823 (2015).
386. Hubbard, T. P., D'Gama, J. D., Billings, G., Davis, B. M. & Waldor, M. K. Unsupervised Learning Approach for Comparing Multiple Transposon Insertion Sequencing Studies. **4**, 13 (2019).
387. Franzin, F. M. & Sircili, M. P. Locus of Enterocyte Effacement: A Pathogenicity Island Involved in the Virulence of Enteropathogenic and Enterohemorragic *Escherichia coli* Subjected to a Complex Network of Gene Regulation. *BioMed Research International* **2015**, 1–10 (2015).
388. Fu, Y., Waldor, M. K. & Mekalanos, J. J. Tn-Seq Analysis of Vibrio cholerae Intestinal Colonization Reveals a Role for T6SS-Mediated Antibacterial Activity in the Host. *Cell Host & Microbe* **14**, 652–663 (2013).
389. Abu Kwaik, Y. & Bumann, D. Host Delivery of Favorite Meals for Intracellular Pathogens. *PLOS Pathogens* **11**, e1004866 (2015).
390. Meylan, S. *et al.* Carbon Sources Tune Antibiotic Susceptibility in Pseudomonas aeruginosa via Tricarboxylic Acid Cycle Control. *Cell Chemical Biology* **24**, 195–206 (2017).
391. Murima, P., McKinney, J. D. & Pethe, K. Targeting Bacterial Central Metabolism for Drug Development. *Chemistry & Biology* **21**, 1423–1432 (2014).
392. Knutton, S. A novel EspA-associated surface organelle of enteropathogenic Escherichia coli involved in protein translocation into epithelial cells. *The EMBO Journal* **17**, 2166–2176 (1998).
393. Kresse, A. U. & Rohde, M. The EspD Protein of Enterohemorrhagic Escherichia coli Is Required for the Formation of Bacterial Surface Appendages and Is Incorporated in the Cytoplasmic Membranes of Target Cells. *Infect. Immun.* **67**, 9 (1999).
394. Kresse, A. U. *et al.* Pas, a Novel Protein Required for Protein Secretion and Attaching and Effacing Activities of Enterohemorrhagic Escherichia coli. *J. Bacteriol.* **180**, 10 (1998).

395. Tatsuno, I. *et al.* Isolation and Characterization of Mini-Tn5Km2 Insertion Mutants of Enterohemorrhagic Escherichia coli O157:H7 Deficient in Adherence to Caco-2 Cells. *Infection and Immunity* **68**, 5943–5952 (2000).
396. Sal-Man, N. *et al.* EscE and EscG Are Cochaperones for the Type III Needle Protein EscF of Enteropathogenic Escherichia coli. *Journal of Bacteriology* **195**, 2481–2489 (2013).
397. Gruenheid, S. *et al.* Identification and characterization of NleA, a non-LEE-encoded type III translocated virulence factor of enterohaemorrhagic Escherichia coli O157:H7: Identification and characterization of NleA. *Molecular Microbiology* **51**, 1233–1249 (2004).
398. Yen, H., Sugimoto, N. & Tobe, T. Enteropathogenic Escherichia coli Uses NleA to Inhibit NLRP3 Inflammasome Activation. *PLOS Pathogens* **11**, e1005121 (2015).
399. Arbeloa, A. *et al.* Subversion of actin dynamics by EspM effectors of attaching and effacing bacterial pathogens. *Cellular Microbiology* **10**, 1429–1441 (2008).
400. Simovitch, M. *et al.* EspM inhibits pedestal formation by enterohaemorrhagic *Escherichia coli* and enteropathogenic *E. coli* and disrupts the architecture of a polarized epithelial monolayer. *Cellular Microbiology* **12**, 489–505 (2010).
401. Eckert, S. E. *et al.* Retrospective Application of Transposon-Directed Insertion Site Sequencing to a Library of Signature-Tagged Mini-Tn5Km2 Mutants of Escherichia coli O157:H7 Screened in Cattle. *Journal of Bacteriology* **193**, 1771–1776 (2011).
402. Munera, D., Crepin, V. F., Marches, O. & Frankel, G. N-Terminal Type III Secretion Signal of Enteropathogenic Escherichia coli Translocator Proteins. *Journal of Bacteriology* **192**, 3534–3539 (2010).
403. Iyoda, S. & Watanabe, H. ClpXP Protease Controls Expression of the Type III Protein Secretion System through Regulation of RpoS and GrlR Levels in Enterohemorrhagic Escherichia coli. *Journal of Bacteriology* **187**, 4086–4094 (2005).
404. Tomoyasu, T., Takaya, A., Handa, Y., Karata, K. & Yamamoto, T. ClpXP controls the expression of LEE genes in enterohaemorrhagic Escherichia coli. *FEMS Microbiology Letters* **253**, 59–66 (2005).
405. MacRitchie, D. M., Ward, J. D., Nevesinjac, A. Z. & Raivio, T. L. Activation of the Cpx Envelope Stress Response Down-Regulates Expression of Several Locus of Enterocyte Effacement-Encoded Genes in Enteropathogenic Escherichia coli. *Infection and Immunity* **76**, 1465–1475 (2008).
406. MacRitchie, D. M., Acosta, N. & Raivio, T. L. DegP Is Involved in Cpx-Mediated Posttranscriptional Regulation of the Type III Secretion Apparatus in Enteropathogenic Escherichia coli. *Infection and Immunity* **80**, 1766–1772 (2012).
407. Hara, H., Yamamoto, Y., Higashitani, A., Suzuki, H. & Nishimura, Y. Cloning, mapping, and characterization of the Escherichia coli prc gene, which is involved in C-terminal processing of penicillin-binding protein 3. *Journal of Bacteriology* **173**, 4799–4813 (1991).
408. Kerr, C. H., Culham, D. E., Marom, D. & Wood, J. M. Salinity-Dependent Impacts of ProQ, Prc, and Spr Deficiencies on Escherichia coli Cell Structure. *Journal of Bacteriology* **196**, 1286–1296 (2014).
409. Pos, K. M. Drug transport mechanism of the AcrB efflux pump. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1794**, 782–793 (2009).
410. Christman, M. F., Storz, G. & Ames, B. N. OxyR, a positive regulator of hydrogen peroxide-inducible genes in Escherichia coli and Salmonella typhimurium, is homologous to a family of bacterial regulatory proteins. *Proceedings of the National Academy of Sciences* **86**, 3484–3488 (1989).
411. Daugherty, A., Suvarnapunya, A. E. & Runyen-Janecky, L. The role of OxyR and SoxRS in oxidative stress survival in Shigella flexneri. *Microbiological Research* **167**, 238–245 (2012).
412. De Paepe, M. *et al.* Trade-Off between Bile Resistance and Nutritional Competence Drives Escherichia coli Diversification in the Mouse Gut. *PLoS Genetics* **7**, e1002107 (2011).
413. Berks, B. C. The Twin-Arginine Protein Translocation Pathway. *Annual Review of Biochemistry* **84**, 843–864 (2015).
414. Green, E. R. & Mecsas, J. Bacterial Secretion Systems: An Overview. in *Virulence Mechanisms of Bacterial Pathogens, Fifth Edition* (eds. Kudva, I. T. et al.) 215–239 (American Society of Microbiology, 2016).
415. Mickael, C. S. *et al.* Salmonella enterica Serovar Enteritidis tatB and tatC Mutants Are Impaired in Caco-2 Cell Invasion In Vitro and Show Reduced Systemic Spread in Chickens. *Infection and Immunity* **78**, 3493–3505 (2010).
416. Craig, M., Sadik, A. Y., Golubeva, Y. A., Tidhar, A. & Slauch, J. M. Twin-arginine translocation system (*tat*) mutants of *Salmonella* are attenuated due to envelope defects, not respiratory defects: Role of Tat in *Salmonella* virulence. *Molecular Microbiology* **89**, 887–902 (2013).
417. Fujimoto, M. *et al.* Tat-exported peptidoglycan amidase-dependent cell division contributes to Salmonella Typhimurium fitness in the inflamed gut. *PLOS Pathogens* **14**, e1007391 (2018).
418. Avican, U., Doruk, T., Östberg, Y., Fahlgren, A. & Forsberg, Å. The Tat Substrate SufI Is Critical for the Ability of Yersinia pseudotuberculosis To Cause Systemic Infection. *Infection and Immunity* **85**, (2017).

419. Lavander, M., Ericsson, S. K., Broms, J. E. & Forsberg, A. The Twin Arginine Translocation System Is Essential for Virulence of Yersinia pseudotuberculosis. *Infection and Immunity* **74**, 1768–1776 (2006).
420. Rajashekara, G. *et al.* Functional Characterization of the Twin-Arginine Translocation System in *Campylobacter jejuni*. *Foodborne Pathogens and Disease* **6**, 935–945 (2009).
421. Zhang, L. *et al.* Pleiotropic effects of the twin-arginine translocation system on biofilm formation, colonization, and virulence in Vibrio cholerae. *BMC Microbiology* **9**, 114 (2009).
422. Pradel, N. *et al.* Contribution of the Twin Arginine Translocation System to the Virulence of Enterohemorrhagic Escherichia coli O157:H7. *Infection and Immunity* **71**, 4908–4916 (2003).
423. Samaluru, H., SaiSree, L. & Reddy, M. Role of SufI (FtsP) in Cell Division of Escherichia coli: Evidence for Its Involvement in Stabilizing the Assembly of the Divisome. *Journal of Bacteriology* **189**, 8044–8052 (2007).
424. Tarry, M. *et al.* The Escherichia coli Cell Division Protein and Model Tat Substrate SufI (FtsP) Localizes to the Septal Ring and Has a Multicopper Oxidase-Like Structure. *Journal of Molecular Biology* **386**, 504–519 (2009).
425. Stanley, N. R., Findlay, K., Berks, B. C. & Palmer, T. Escherichia coli Strains Blocked in Tat-Dependent Protein Export Exhibit Pleiotropic Defects in the Cell Envelope. *Journal of Bacteriology* **183**, 139–144 (2001).
426. Bernhardt, T. G. & De Boer, P. A. J. Screening for synthetic lethal mutants in Escherichia coli and identification of EnvC (YibP) as a periplasmic septal ring factor with murein hydrolase activity: E. coli synthetic lethal screen. *Molecular Microbiology* **52**, 1255–1269 (2004).
427. Chaudhuri, R. R. *et al.* Comprehensive Assignment of Roles for Salmonella Typhimurium Genes in Intestinal Colonization of Food-Producing Animals. *PLoS Genetics* **9**, e1003456 (2013).
428. Fath, M. J., Mahanty, H. K. & Kolter, R. Characterization of a purF operon mutation which affects colicin V production. *Journal of Bacteriology* **171**, 3158–3161 (1989).
429. Hadjifrangiskou, M. *et al.* Transposon Mutagenesis Identifies Uropathogenic Escherichia coli Biofilm Factors. *Journal of Bacteriology* **194**, 6195–6205 (2012).
430. Shaffer, C. L. *et al.* Purine Biosynthesis Metabolically Constrains Intracellular Survival of Uropathogenic Escherichia coli. *Infection and Immunity* **85**, (2017).
431. Kwon, Y. M., Ricke, S. C. & Mandal, R. K. Transposon sequencing: methods and expanding applications. *Applied Microbiology and Biotechnology* **100**, 31–43 (2016).
432. Barquist, L., Boinett, C. J. & Cain, A. K. Approaches to querying bacterial genomes with transposon-insertion sequencing. *RNA Biology* **10**, 1161–1169 (2013).
433. Donnenberg, M. S. & Kaper, J. B. Construction of an eae Deletion Mutant of Enteropathogenic Escherichia coli by Using a Positive-Selection Suicide Vector. *Infection and Immunity*. **59**, 8 (1991).
434. Poteete, A. R., Rosadini, C. & St. Pierre, C. Gentamicin and other cassettes for chromosomal gene replacement in *Escherichia coli*. *BioTechniques* **41**, 261–264 (2006).
435. Ferrieres, L. *et al.* Silent Mischief: Bacteriophage Mu Insertions Contaminate Products of Escherichia coli Random Mutagenesis Performed Using Suicidal Transposon Delivery Plasmids Mobilized by Broad-Host-Range RP4 Conjugative Machinery. *Journal of Bacteriology* **192**, 6418–6427 (2010).
436. Cameron, D. E., Urbach, J. M. & Mekalanos, J. J. A defined transposon mutant library and its use in identifying motility genes in Vibrio cholerae. *Proceedings of the National Academy of Sciences* **105**, 8736–8741 (2008).
437. Chiang, S. L. & Rubin, E. J. Construction of a mariner -based transposon for epitope-tagging and genomic targeting. *Gene* **296**, 179–185 (2002).
438. Hubbard, T. P. *et al.* A live vaccine rapidly protects against cholera in an infant rabbit model. *Science Translational Medicine* **10**, eaap8423 (2018).
439. Abel, S. *et al.* Sequence tag–based analysis of microbial population dynamics. *Nature Methods* **12**, 223–226 (2015).
440. Collington, G. K., Booth, I. W. & Knutton, S. Rapid modulation of electrolyte transport in Caco-2 cell monolayers by enteropathogenic Escherichia coli (EPEC) infection. *Gut* **42**, 200–207 (1998).
441. Stincone, A. *et al.* A systems biology approach sheds new light on Escherichia coli acid resistance. *Nucleic Acids Research* **39**, 7512–7528 (2011).
442. Dörr, T. *et al.* Differential Requirement for PBP1a and PBP1b in *In Vivo* and *In Vitro* Fitness of Vibrio cholerae. *Infection and Immunity* **82**, 2115–2124 (2014).
443. Cremers, C. M., Knoefler, D., Vitvitsky, V., Banerjee, R. & Jakob, U. Bile salts act as effective protein-unfolding agents and instigators of disulfide stress in vivo. *Proceedings of the National Academy of Sciences* **111**, E1610–E1619 (2014).
444. Morales, V. M., Bäckman, A. & Bagdasarian, M. A series of wide-host-range low-copy-number vectors that allow direct screening for recombinants. *Gene* **97**, 39–47 (1991).
445. Boyer, J. L. Bile Formation and Secretion. *Compr Physiol* **3**, 1035–1078 (2013).

446. Maldonado-Valderrama, J., Wilde, P., Macierzanka, A. & Mackie, A. The role of bile salts in digestion. *Adv Colloid Interface Sci* **165**, 36–46 (2011).
447. Prieto, A. I., Ramos-Morales, F. & Casadesús, J. Repair of DNA Damage Induced by Bile Salts in Salmonella enterica. *Genetics* **174**, 575–584 (2006).
448. Rodríguez-Beltrán, J., Rodríguez-Rojas, A., Guelfo, J. R., Couce, A. & Blázquez, J. The Escherichia coli SOS Gene dinF Protects against Oxidative Stress and Bile Salts. *PLOS ONE* **7**, e34791 (2012).
449. Bernstein, C., Bernstein, H., Payne, C. M., Beard, S. E. & Schneider, J. Bile salt activation of stress response promoters in Escherichia coli. *Curr. Microbiol.* **39**, 68–72 (1999).
450. Vijayakumar, S. R. V., Kirchhof, M. G., Patten, C. L. & Schellhorn, H. E. RpoS-Regulated Genes of Escherichia coli Identified by Random lacZ Fusion Mutagenesis. *J Bacteriol* **186**, 8499–8507 (2004).
451. Paget, M. S. Bacterial Sigma Factors and Anti-Sigma Factors: Structure, Function and Distribution. *Biomolecules* **5**, 1245–1265 (2015).
452. Weber, H., Polen, T., Heuveling, J., Wendisch, V. F. & Hengge, R. Genome-Wide Analysis of the General Stress Response Network in Escherichia coli: σS-Dependent Genes, Promoters, and Sigma Factor Selectivity. *J Bacteriol* **187**, 1591–1603 (2005).
453. Hews, C. L., Cho, T., Rowley, G. & Raivio, T. L. Maintaining Integrity Under Stress: Envelope Stress Response Regulation of Pathogenesis in Gram-Negative Bacteria. *Front. Cell. Infect. Microbiol.* **9**, (2019).
454. Ray, S., Da Costa, R., Das, M. & Nandi, D. Interplay of cold shock protein E with an uncharacterized protein, YciF, lowers porin expression and enhances bile resistance in *Salmonella* Typhimurium. *J. Biol. Chem.* **294**, 9084–9099 (2019).
455. Guest, R. L. & Raivio, T. L. Role of the Gram-Negative Envelope Stress Response in the Presence of Antimicrobial Agents. *Trends in Microbiology* **24**, 377–390 (2016).
456. Hamner, S., McInnerney, K., Williamson, K., Franklin, M. J. & Ford, T. E. Bile salts affect expression of Escherichia coli O157:H7 genes for virulence and iron acquisition, and promote growth under iron limiting conditions. *PLoS ONE* **8**, e74647 (2013).
457. Kus, J. V. *et al.* Bile Salts Induce Resistance to Polymyxin in Enterohemorrhagic Escherichia coliO157:H7 ▿. *J Bacteriol* **193**, 4509–4515 (2011).
458. Arenas-Hernández, M. M. P. *et al.* Environmental regulation of the long polar fimbriae 2 of enterohemorrhagic Escherichia coli O157:H7. *FEMS Microbiol Lett* **357**, 105–114 (2014).
459. Rosenberg, E. Y., Bertenthal, D., Nilles, M. L., Bertrand, K. P. & Nikaido, H. Bile salts and fatty acids induce the expression of Escherichia coli AcrAB multidrug efflux pump through their interaction with Rob regulatory protein. *Mol. Microbiol.* **48**, 1609–1619 (2003).
460. Thanassi, D. G., Cheng, L. W. & Nikaido, H. Active efflux of bile salts by Escherichia coli. *J Bacteriol* **179**, 2512–2518 (1997).
461. Froelich, J. M., Tran, K. & Wall, D. A pmrA Constitutive Mutant Sensitizes Escherichia coli to Deoxycholic Acid. *Journal of Bacteriology* **188**, 1180–1183 (2006).
462. Kwan, B. W. *et al.* The MqsR/MqsA toxin/antitoxin system protects Escherichia coli during bile acid stress. *Environ. Microbiol.* **17**, 3168–3181 (2015).
463. Zhang, L. H., Fath, M. J., Mahanty, H. K., Tai, P. C. & Kolter, R. Genetic analysis of the colicin V secretion pathway. *Genetics* **141**, 25–32 (1995).
464. Sueki, A., Stein, F., Savitski, M. M., Selkrig, J. & Typas, A. Systematic Localization of Escherichia coli Membrane Proteins. *mSystems* **5**, (2020).
465. Jensen, H. M., Eng, T., Chubukov, V., Herbert, R. A. & Mukhopadhyay, A. Improving membrane protein expression and function using genomic edits. *Scientific Reports* **7**, 13030 (2017).
466. Dörr, T. *et al.* A Transposon Screen Identifies Genetic Determinants of Vibrio cholerae Resistance to High-Molecular-Weight Antibiotics. *Antimicrobial Agents and Chemotherapy* **60**, 4757–4763 (2016).
467. Greene, N. G., Fumeaux, C. & Bernhardt, T. G. Conserved mechanism of cell-wall synthase regulation revealed by the identification of a new PBP activator in Pseudomonas aeruginosa. *PNAS* **115**, 3150–3155 (2018).
468. Paradis-Bleau, C. *et al.* Lipoprotein cofactors located in the outer membrane activate bacterial cell wall polymerases. *Cell* **143**, 1110–1120 (2010).
469. Fleurie, A. *et al.* A Vibrio cholerae BolA-Like Protein Is Required for Proper Cell Shape and Cell Envelope Integrity. *mBio* **10**, (2019).
470. Meeske, A. J. *et al.* MurJ and a novel lipid II flippase are required for cell wall biogenesis in Bacillus subtilis. *PNAS* **112**, 6437–6442 (2015).
471. Wivagg, C. N., Wellington, S., Gomez, J. E. & Hung, D. T. Loss of a Class A Penicillin-Binding Protein Alters β-Lactam Susceptibilities in Mycobacterium tuberculosis. *ACS Infect Dis* **2**, 104–110 (2016).

472. Klobucar, K. & Brown, E. D. Use of genetic and chemical synthetic lethality as probes of complexity in bacterial cell systems. *FEMS Microbiol Rev* **42**, (2018).
473. Harms, C. *et al.* Identification of the ABC protein SapD as the subunit that confers ATP dependence to the K+-uptake systems Trk(H) and Trk(G) from Escherichia coli K-12. *Microbiology (Reading)* **147**, 2991–3003 (2001).
474. Lee, C.-R. *et al.* Potassium mediates Escherichia coli enzyme IIANtr-dependent regulation of sigma factor selectivity. *Molecular Microbiology* **78**, 1468–1483 (2010).
475. Tam, C. & Missiakas, D. Changes in lipopolysaccharide structure induce the sigma(E)-dependent response of Escherichia coli. *Mol. Microbiol.* **55**, 1403–1412 (2005).
476. Lima, S., Guo, M. S., Chaba, R., Gross, C. A. & Sauer, R. T. Dual Molecular Signals Mediate the Bacterial Response to Outer-Membrane Stress. *Science* **340**, 837–841 (2013).
477. Testerman, T. L. *et al.* The alternative sigma factor sigmaE controls antioxidant defences required for Salmonella virulence and stationary-phase survival. *Mol. Microbiol.* **43**, 771–782 (2002).
478. Shimada, T., Tanaka, K. & Ishihama, A. The whole set of the constitutive promoters recognized by four minor sigma subunits of Escherichia coli RNA polymerase. *PLOS ONE* **12**, e0179181 (2017).
479. Ding, Y., Davis, B. M. & Waldor, M. K. Hfq is essential for Vibrio cholerae virulence and downregulates σE expression. *Molecular Microbiology* **53**, 345–354 (2004).
480. Mecsas, J., Rouviere, P. E., Erickson, J. W., Donohue, T. J. & Gross, C. A. The activity of sigma E, an Escherichia coli heat-inducible sigma-factor, is modulated by expression of outer membrane proteins. *Genes Dev.* **7**, 2618–2628 (1993).
481. Mendler, K. *et al.* AnnoTree: visualization and exploration of a functionally annotated microbial tree of life. *Nucleic Acids Res* **47**, 4442–4448 (2019).
482. Russ, W. P. & Engelman, D. M. The GxxxG motif: A framework for transmembrane helix-helix association11Edited by G. von Heijne. *Journal of Molecular Biology* **296**, 911–919 (2000).
483. Alba, B. M., Leeds, J. A., Onufryk, C., Lu, C. Z. & Gross, C. A. DegS and YaeL participate sequentially in the cleavage of RseA to activate the σE-dependent extracytoplasmic stress response. *Genes Dev* **16**, 2156–2168 (2002).
484. Walsh, N. P., Alba, B. M., Bose, B., Gross, C. A. & Sauer, R. T. OMP peptide signals initiate the envelope-stress response by activating DegS protease via relief of inhibition mediated by its PDZ domain. *Cell* **113**, 61–71 (2003).
485. Kanehara, K., Ito, K. & Akiyama, Y. YaeL (EcfE) activates the sigma(E) pathway of stress response through a site-2 cleavage of anti-sigma(E), RseA. *Genes Dev.* **16**, 2147–2155 (2002).
486. Sousa, T. *et al.* Deoxycholic acid modulates cell death signaling through changes in mitochondrial membrane properties. *J Lipid Res* **56**, 2158–2171 (2015).
487. Cabral, D. J., Small, D. M., Lilly, H. S. & Hamilton, J. A. Transbilayer movement of bile acids in model membranes. *Biochemistry* **26**, 1801–1804 (1987).
488. Yang, C. C. & Konisky, J. Colicin V-treated Escherichia coli does not generate membrane potential. *J. Bacteriol.* **158**, 757–759 (1984).
489. Gérard, F., Pradel, N. & Wu, L.-F. Bactericidal Activity of Colicin V Is Mediated by an Inner Membrane Protein, SdaC, of Escherichia coli. *J Bacteriol* **187**, 1945–1950 (2005).
490. Lazarus, J. E. *et al.* A New Suite of Allelic-Exchange Vectors for the Scarless Modification of Proteobacterial Genomes. *Appl. Environ. Microbiol.* **85**, (2019).
491. Kuru, E. *et al.* In Situ probing of newly synthesized peptidoglycan in live bacteria with fluorescent D-amino acids. *Angew. Chem. Int. Ed. Engl.* **51**, 12519–12523 (2012).
492. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nature Methods* **9**, 676–682 (2012).
493. Bhasin, M., Garg, A. & Raghava, G. P. S. PSLpred: prediction of subcellular localization of bacterial proteins. *Bioinformatics* **21**, 2522–2524 (2005).
494. Zimmermann, L. *et al.* A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. *Journal of Molecular Biology* **430**, 2237–2243 (2018).
495. Käll, L., Krogh, A. & Sonnhammer, E. L. L. A Combined Transmembrane Topology and Signal Peptide Prediction Method. *Journal of Molecular Biology* **338**, 1027–1036 (2004).
496. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols* **10**, 845–858 (2015).
497. Viklund, H., Bernsel, A., Skwark, M. & Elofsson, A. SPOCTOPUS: a combined predictor of signal peptides and membrane protein topology. *Bioinformatics* **24**, 2928–2929 (2008).
498. Omasits, U., Ahrens, C. H., Müller, S. & Wollscheid, B. Protter: interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics* **30**, 884–886 (2014).
499. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792–1797 (2004).

500. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **35**, 1547–1549 (2018).
501. Yen, H., Karino, M. & Tobe, T. Modulation of the Inflammasome Signaling Pathway by Enteropathogenic and Enterohemorrhagic Escherichia coli. *Front Cell Infect Microbiol* **6**, (2016).
502. Santos, A. S. & Finlay, B. B. Bringing down the host: enteropathogenic and enterohaemorrhagic Escherichia coli effector-mediated subversion of host innate immune pathways. *Cell. Microbiol.* **17**, 318–332 (2015).
503. Martins, F. H. *et al.* EspFu-Mediated Actin Assembly Enhances Enteropathogenic Escherichia coli Adherence and Activates Host Cell Inflammatory Signaling Pathways. *mBio* **11**, (2020).
504. Litvak, Y. *et al.* Epithelial cells detect functional type III secretion system of enteropathogenic Escherichia coli through a novel NF-κB signaling pathway. *PLoS Pathog.* **13**, e1006472 (2017).
505. He, K., Zhou, H.-R. & Pestka, J. J. Mechanisms for Ribotoxin-induced Ribosomal RNA Cleavage. *Toxicol Appl Pharmacol* **265**, 10–18 (2012).
506. Zumbrun, S. D. *et al.* Human Intestinal Tissue and Cultured Colonic Cells Contain Globotriaosylceramide Synthase mRNA and the Alternate Shiga Toxin Receptor Globotetraosylceramide. *Infection and Immunity* **78**, 4488–4499 (2010).
507. Philpott, D. J. *et al.* Translocation of verotoxin-1 across T84 monolayers: mechanism of bacterial toxin penetration of epithelium. *Am. J. Physiol.* **273**, G1349-1358 (1997).
508. Tesh, V. L., Ramegowda, B. & Samuel, J. E. Purified Shiga-like toxins induce expression of proinflammatory cytokines from murine peritoneal macrophages. *Infection and Immunity* **62**, 5085–5094 (1994).
509. Bellmeyer, A. *et al.* Enterohemorrhagic Escherichia coli suppresses inflammatory response to cytokines and its own toxin. *American Journal of Physiology-Gastrointestinal and Liver Physiology* **297**, G576–G581 (2009).
510. Annunziato, F., Romagnani, C. & Romagnani, S. The 3 major types of innate and adaptive cell-mediated effector immunity. *J. Allergy Clin. Immunol.* **135**, 626–635 (2015).
511. Reid-Yu, S. A., Small, C.-L. N. & Coombes, B. K. $CD3^{-}NK1.1^{+}$ cells aid in the early induction of a Th1 response to an attaching and effacing enteric pathogen. *Eur. J. Immunol.* **43**, 2638–2649 (2013).
512. Higgins, L. M., Frankel, G., Douce, G., Dougan, G. & MacDonald, T. T. Citrobacter rodentium infection in mice elicits a mucosal Th1 cytokine response and lesions similar to those in murine inflammatory bowel disease. *Infect. Immun.* **67**, 3031–3039 (1999).
513. Shiomi, H. *et al.* Gamma interferon produced by antigen-specific CD4+ T cells regulates the mucosal immune responses to Citrobacter rodentium infection. *Infect. Immun.* **78**, 2653–2666 (2010).
514. Chen, B., Khodadoust, M. S., Liu, C. L., Newman, A. M. & Alizadeh, A. A. Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol Biol* **1711**, 243–259 (2018).
515. Newman, A. M. *et al.* Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nature Biotechnology* **37**, 773–782 (2019).
516. Mayer, C. L., Leibowitz, C. S., Kurosawa, S. & Stearns-Kurosawa, D. J. Shiga Toxins and the Pathophysiology of Hemolytic Uremic Syndrome in Humans and Animals. *Toxins* **4**, 1261–1287 (2012).
517. Grabowski, E. F., Kushak, R. I., Liu, B. & Ingelfinger, J. R. Shiga toxin downregulates tissue factor pathway inhibitor, modulating an increase in the expression of functional tissue factor on endothelium. *Thrombosis Research* **131**, 521–528 (2013).
518. Ishii, H., Takada, K., Higuchi, T. & Sugiyama, J. Verotoxin-1 Induces Tissue Factor Expression in Human Umbilical Vein Endothelial Cells through Activation of NF-κB/Rel and AP-1. *Thromb Haemost* **84**, 712–721 (2000).
519. Nestoridi, E., Tsukurov, O., Kushak, R. I., Ingelfinger, J. R. & Grabowski, E. F. Shiga toxin enhances functional tissue factor on human glomerular endothelial cells: implications for the pathophysiology of hemolytic uremic syndrome. *J. Thromb. Haemost.* **3**, 752–762 (2005).
520. Nestoridi, E., Kushak, R. I., Duguerre, D., Grabowski, E. F. & Ingelfinger, J. R. Up-regulation of tissue factor activity on human proximal tubular epithelial cells in response to Shiga toxin. *Kidney International* **67**, 2254–2266 (2005).
521. Grabowski, E. F., Liu, B., Gerace, M. R., Kushak, R. I. & Ingelfinger, J. R. Shiga toxin-1 Decreases Endothelial Cell Tissue Factor Pathway Inhibitor Not Co-localized with Tissue Factor on the Cell Membrane. *Thrombosis Research* **135**, 1214–1217 (2015).
522. Ho, N. K. *et al.* Enterohemorrhagic Escherichia coli O157:H7 Shiga Toxins Inhibit Gamma Interferon-Mediated Cellular Activation. *Infect Immun* **80**, 2307–2315 (2012).
523. Schneider, W. M., Chevillotte, M. D. & Rice, C. M. Interferon-Stimulated Genes: A Complex Web of Host Defenses. *Annu Rev Immunol* **32**, 513–545 (2014).

524. Ruan, W. *et al.* Enhancing responsiveness of human jejunal enteroids to host and microbial stimuli. *The Journal of Physiology* **598**, 3085–3105 (2020).
525. Oppmann, B. *et al.* Novel p19 protein engages IL-12p40 to form a cytokine, IL-23, with biological activities similar as well as distinct from IL-12. *Immunity* **13**, 715–725 (2000).
526. Schuetze, N. *et al.* IL-12 family members: differential kinetics of their TLR4-mediated induction by Salmonella Enteritidis and the impact of IL-10 in bone marrow-derived macrophages. *Int Immunol* **17**, 649–659 (2005).
527. Lim, K. S. *et al.* Inflammatory and mitogenic signals drive interleukin 23 subunit alpha (IL23A) secretion independent of IL12B in intestinal epithelial cells. *J. Biol. Chem.* **295**, 6387–6400 (2020).
528. Ceponis, P. J. M., McKay, D. M., Ching, J. C. Y., Pereira, P. & Sherman, P. M. Enterohemorrhagic Escherichia coli O157:H7 Disrupts Stat1-Mediated Gamma Interferon Signal Transduction in Epithelial Cells. *Infection and Immunity* **71**, 1396–1404 (2003).
529. Brereton, C. F. *et al.* Escherichia coli heat-labile enterotoxin promotes protective Th17 responses against infection by driving innate IL-1 and IL-23 production. *J. Immunol.* **186**, 5896–5906 (2011).
530. Cowardin, C. A. *et al.* Inflammasome activation contributes to interleukin-23 production in response to Clostridium difficile. *mBio* **6**, (2015).
531. Ruf, W. & Riewald, M. *Regulation of Tissue Factor Expression*. (Landes Bioscience, 2013).
532. Guo, L. & Rondina, M. T. The Era of Thromboinflammation: Platelets Are Dynamic Sensors and Effector Cells During Infectious Diseases. *Front. Immunol.* **10**, (2019).
533. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *PNAS* **97**, 6640–6645 (2000).
534. Pacheco, A. R. *et al.* CRISPR Screen Reveals that EHEC's T3SS and Shiga Toxin Rely on Shared Host Factors for Infection. *mBio* **9**, (2018).
535. Smillie, C. S. *et al.* Intra- and Inter-cellular Rewiring of the Human Colon during Ulcerative Colitis. *Cell* **178**, 714-730.e22 (2019).
536. Afgan, E. *et al.* The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res.* **44**, W3–W10 (2016).
537. R Core Team. *R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/*. (2019).
538. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, (2014).
539. Zhu, A., Ibrahim, J. G. & Love, M. I. Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. *Bioinformatics* **35**, 2084–2092 (2019).
540. Brewer, C. A., Hatchard, G. W. & Harrower, M. A. ColorBrewer in Print: A Catalog of Color Schemes for Maps. *Cartography and Geographic Information Science* **30**, 5–32 (2003).
541. Korotkevich, G., Sukhov, V. & Sergushichev, A. Fast gene set enrichment analysis. *bioRxiv* 060012 (2019).
542. Liberzon, A. *et al.* The Molecular Signatures Database Hallmark Gene Set Collection. *cels* **1**, 417–425 (2015).
543. Godec, J. *et al.* Compendium of Immune Signatures Identifies Conserved and Species-Specific Biology in Response to Inflammation. *Immunity* **44**, 194–206 (2016).
544. Liberzon, A. *et al.* Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740 (2011).
545. Hulsen, T., de Vlieg, J. & Alkema, W. BioVenn – a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. *BMC Genomics* **9**, 488 (2008).
546. Raudvere, U. *et al.* g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res* **47**, W191–W198 (2019).
547. Sturm, G. *et al.* Comprehensive evaluation of transcriptome-based cell-type quantification methods for immuno-oncology. *Bioinformatics* **35**, i436–i445 (2019).
548. Hayek, N. Lateral transfer and GC content of bacterial resistance genes. *Front. Microbiol.* **4**, (2013).
549. Navarre, W. W. Selective Silencing of Foreign DNA with Low GC Content by the H-NS Protein in Salmonella. *Science* **313**, 236–238 (2006).
550. Navarre, W. W., McClelland, M., Libby, S. J. & Fang, F. C. Silencing of xenogeneic DNA by H-NS—facilitation of lateral gene transfer in bacteria by a defense system that recognizes foreign DNA. *Genes Dev.* **21**, 1456–1471 (2007).
551. Chao, M. C. *et al.* High-resolution definition of the Vibrio cholerae essential gene set with hidden Markov model–based analyses of transposon-insertion sequencing data. *Nucleic Acids Research* **41**, 9033–9048 (2013).
552. Francino, M. P. Antibiotics and the Human Gut Microbiome: Dysbioses and Accumulation of Resistances. *Front Microbiol* **6**, (2016).

553. Fleitas Martínez, O., Cardoso, M. H., Ribeiro, S. M. & Franco, O. L. Recent Advances in Anti-virulence Therapeutic Strategies With a Focus on Dismantling Bacterial Membrane Microdomains, Toxin Neutralization, Quorum-Sensing Interference and Biofilm Inhibition. *Front Cell Infect Microbiol* **9**, (2019).
554. Kuehl, C. J., D'Gama, J. D., Warr, A. R. & Waldor, M. K. An Oral Inoculation Infant Rabbit Model for Shigella Infection. *mBio* **11**, (2020).
555. Subuddhi, U. & Mishra, A. K. Effect of sodium deoxycholate and sodium cholate on DPPC vesicles: A fluorescence anisotropy study with diphenylhexatriene. *J Chem Sci* **119**, 169–174 (2007).
556. Devor, D. C., Sekar, M. C., Frizzell, R. A. & Duffey, M. E. Taurodeoxycholate activates potassium and chloride conductances via an IP3-mediated release of calcium from intracellular stores in a colonic cell line (T84). *J Clin Invest* **92**, 2173–2181 (1993).
557. Lau, B. W. *et al.* Deoxycholic acid activates protein kinase C and phospholipase C via increased Ca2+ entry at plasma membrane. *Gastroenterology* **128**, 695–707 (2005).
558. Waise, S. *et al.* An optimised tissue disaggregation and data processing pipeline for characterising fibroblast phenotypes using single-cell RNA sequencing. *Scientific Reports* **9**, 9580 (2019).
559. Nuss, A. M. *et al.* Tissue dual RNA-seq allows fast discovery of infection-specific functions and riboregulators shaping host–pathogen transcriptomes. *PNAS* **114**, E791–E800 (2017).
560. Ho, N. K. *et al.* Enterohemorrhagic Escherichia coli O157:H7 Shiga toxins inhibit gamma interferon-mediated cellular activation. *Infect. Immun.* **80**, 2307–2315 (2012).
561. Vossenkämper, A., MacDonald, T. T. & Marchès, O. Always one step ahead: How pathogenic bacteria use the type III secretion system to manipulate the intestinal mucosal immune system. *Journal of Inflammation* **8**, 11 (2011).
562. Li, Y., Yu, X., Ma, Y. & Hua, S. IL-23 and dendritic cells: What are the roles of their mutual attachment in immune response and immunotherapy? *Cytokine* **120**, 78–84 (2019).
563. Silberger, D. J., Zindl, C. L. & Weaver, C. T. Citrobacter rodentium: A Model Enteropathogen for Understanding the Interplay of Innate and Adaptive Components of Type 3 Immunity. *Mucosal Immunol* **10**, 1108–1117 (2017).
564. Heyderman, R. S., Soriani, M. & Hirst, T. R. Is immune cell activation the missing link in the pathogenesis of post-diarrhoeal HUS? *Trends in Microbiology* **9**, 262–266 (2001).