Understanding host-microbiome interactions and influence on STEC colonization in cattle using integrated omics

by

Zhe Pan

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in
Animal Science

Department of Agricultural, Food and Nutritional Science
University of Alberta

Abstract

Shiga toxin *producing Escherichia coli* (STEC) is the major foodborne pathogen in humans with Shiga toxin 1 (stx1) and 2 (stx2) being the main virulence factors. Cattle are the major reservoir of STEC with those shedding >10^4 CFU/g STEC being defined as super shedders (SS). The rectal anal junction (RAJ) is the primary colonization site of STEC and previous studies have revealed that both fecal and rectal mucosal microbiota impact STEC colonization at the RAJ. To date, the extent to which stx in STEC affects host-microbial interactions remains unidenfined. This thesis aimed to identify how host and fecal/mucosal microbiota respond to stxs and STEC colonization. Study 1 (Chapter 2) consisted of an epidemiological survey to reveal the abundance (DNA) and expression (RNA) of *stx1* and *stx2* in STEC as it was affected by sampling type (fecal vs. rectal mucosa) and breed (Angus, Charolais, Kinsella Composite Expression of *stx2* was influenced by the expressions of host immune genes previously reported to be downregulated in SS including *MS4A1*, *CCL21*, *CD19*, and *LTB*. The random forest model and Boruta method further revealed that *MS4A1* was the most predictive of *stx2* expression, a response that appeared to be linked to host immunity. Study 2 (Chapter 3) performed amplicon sequencing to characterize differences in rectal digesta microbial profiles and interactions using samples collected from steers in which *stx2* was not expressed (Stx2- group) and those with *stx2* expressed (Stx2+ group). Although microbial diversities and similarities did not differ between the two groups, microbial networks were remarkably different, with group-specific microbes being the most connected taxa within the network. These results suggested that the expression of *stx2* in bacteria altered microbial community structures even when their diversity

and composition were comparable to the Stx2- group. Study 3 (Chapter 4) identified the variation in host transcriptomes in veal calves challenged with STEC O157 that lacked stx2a (WT group) or possessed stx2a (RE group) using rectal mucosa samples collected at pre- (T1), 7 days (T2), and 26 days post-challenge (T5). The stx2a is a subtype of stx2 and is critical for STEC pathogenicity. A total of 214 downregulated differentially expressed genes (DEGs) were identified in WT-T2 compared to RE-T2. No upregulated DEGs were identified in WT and RE at T2. At T5, a total of 152 upregulated DEGs and 45 GO terms were shared between WT and RE, while no downregulated DEGs were identified for WT and RE. Functional analysis revealed that WT inhibited responses at extracellular regions and impaired tissue barrier integrity at T2, while those responses were enhanced at T5. For RE, no functional variations were found at T2, with the aforementioned functions enhanced at T5. In study 4 (Chapter 5), cDNA amplicon sequencing was performed to characterize the activity of the rectal mucosa microbial profiles, interactions, and assembly as well as host-microbial interactions related to the expression of *stx2a* gene in RAJ mucosa colonized STEC O157. The rectal mucosa microbial diversities were not affected by the presence of stx2a in STEC. Instead, the dynamics of microbial interactions and assembly patterns differed in response to strain-specific STEC O157 colonization. The relative abundance of *Paeniclostridium* and *Gallibacterium* were identified as connectors in microbial networks and specialists in microbial assembly. Host immune responses varied after challenge with B-cell and T-cell signaling receptor pathways, antigen processing and presentations being upregulated regardless of stx2a production. The beneficial microbes (e.g. *Prevotella*) dominated interactions with host immune genes, while the

opportunistic pathogen *Paeniclostridium* dominated interactions with expressions of host immune genes post challenge and such relationship depended on the production of stx2a. In summary, this thesis provides knowledge of host-microbial interactions in response to *stx* gene expression and STEC colonization. Our findings suggest that STEC colonization and *stx* gene expression could be systematically attributed to differences in genetic variations, host responses, and fecal/mucosal microbiota.

Preface

This thesis is an original work by Zhe Pan, and it is a research project conducted by Zhe Pan and supervised by Dr. Le Luo Guan. This research project was funded by Beef Cattle Research Council (BCRC, FOS.07.17). The animal trial for Chapters 2 and 3 followed Canadian Council of Animal Care Guidelines and was approved by the Animal Care and Use Committee, University of Alberta (Animal Care Committee protocol number AUP00000882). The animal trial for Chapters 4 and 5 was carried out by Dr. Tom McNeilly's group at the Moredun Research Institute (MRI) under Home Office License 70/7914 granted by the UK Home Office under the Animal (Scientific Procedures) Act 1986 and was approved by the Moredun Research Institute (MRI) animal care and ethics review committee.

Chapter 2 has been published as Pan Z, Chen Y, McAllister TA, Gänzle M, Plastow G, Guan LL. Abundance and Expression of Shiga Toxin Genes in *Escherichia coli* at the Recto-Anal Junction Relates to Host Immune Genes. Front Cell Infect Mi. 2021;11:633573. Under Dr. Le Luo Guan's supervision, I was responsible for data generation and analysis, as well as drafting and editing the manuscript. All co-authors contributed to editing and composing the manuscript.

Chapter 3 has been published as Pan Z, Chen Y, Zhou M, McAllister TA, Guan LL. Microbial interaction-driven community differences as revealed by network analysis. Comput Struct Biotechnology J. 2021;19:6000–8. Under Dr. Le Luo Guan's supervision, I was responsible for data generation and analysis, as well as drafting and editing the manuscript. All co-authors contributed to editing and composing the manuscript.

Dedication

In the hopes that this research may in some way contribute to the sustainability of the beef industry. This thesis is dedicated to my family, friends and supervisor who have supported me throughout the program. Thanks for making me see this adventure through to the end.

Acknowledgements

Words cannot express my gratitude to my supervisor Dr. Le Luo Guan and committee members Drs. Tim McAllister and Paul Stothard for the invaluable patience and feedback on my research. I appreciate my supervisor Dr. Guan as the guide for functional genomics and bioinformatics and Dr. Guan greatly helped me in shaping the way to conduct rigorous and meaningful research. Dr. McAllister, as a knowledgeable expert in STEC, taught me the way to construct the bridge between bioinformatic data and their biological impacts and Dr. Stothard furthered my knowledge in bioinformatics and how to address issues from a computational perspective. I am also grateful to Dr. Tom N.Mcneilly at Moredum Research Institute in UK for the collaboration of the animal study. Additionally, this endeavor would not have been possible without the support from the funding agency Beef Cattle Research Council (BCRC), NSERC, Lallemand.

I am also grateful to lab members and visiting scholars whom I worked with including Yanhong Chen, Mi Zhou, Yajing Ban, Sang Weon Na, Jian Wang, Wei Guo, Yangyi Hao, Yixin Wang, Tao Mao, Eóin O'Hara, Anna Widenmann, Anusha Bulumulla, Zilin Fu for their help and feedback in my research.

Lastly, I would be remiss in not mentioning my friends and family, especially my parents and grandparents. Their belief in me has kept my spirits and motivation high during the PhD program. I would also like to thank my cat Momo and Tantan for all the entertainment and emotional support.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

A/E lesions	attaching and effacing lesions
AI	artificial intelligence
ALDOA	fructose-bisphosphate aldolase A
ANOSIM	Analysis of similarities
ASD	autism spectrum disorder
ASVs	amplicon sequence variants
AUC	area under the ROC curve
AUPR	area under the precision-recall curves
BoLA-DQB	BoLa Class II histocompatibility antigen, DQB*101 beta chain
bTB	bovine tuberculosis
CA	correspondence analysis
CCL21	chemokine (C-C motif) ligand 21
CD19	CD19 molecule
CD77/Gb3	the cellular receptor globotriaosylceramide
CDC	the Center for Disease Control
DBIndex	Davis-Bouldin-Index
E.coli	*Escherichia coli*
EPEC	Enteropathogenic *Escherichia coli*
ER	endoplasmic reticulum
ERAD pathway	ER-associated protein degradation pathway
EspB	*Escherichia coli* secretion protein B
ETEC	Enterotoxigenic *Escherichia coli*
ExPEC	Extraintestinal *Escherichia coli*
FSIS	the Food Safety and Inspection Service
GI	gastrointestinal
GOE	Gaussian orthogonal ensemble
GSEA	gene set enrichment analysis
GWAS	genome wide association studies
HACCP	the Hazard Analysis and Critical Control Points
HC	hemorrhagic colitis
HCP	hemorrhagic coli pili
HUS	hemolytic uremic syndrome

IKK	IκB kinase
IL-17	interleukin 17 family
ILF	lymphoid follicle
KEGG	Kyoto Encyclopedia of Genes and Genomes
KS test	Kolmogorov-Smirnov test
LEE	locus of enterocyte effacement
LPS	lipopolysaccharide
LSA	local similarity analysis
LTB	lymphotoxin beta
MAPK	mitogen-activated protein kinase
MAR	multiple antimicrobial resistance
MENA	molecular ecological network analysis
MIC	maximal information coefficient
mRMR	minimum redundancy-maximum relevance
MS4A1	4-domains, subfamily A, member 1
MyD88	myeloid differentiation factor 88
NF-κB pathway	nuclear factor-κB pathway
Nle	non-Lee effectors
NS	non-shedders
PCA	principal component analysis
PCoA	principal coordinate analysis
PCR	polymerase chain reaction
PCs	principal components
Pi	among-module connectivity
PT	phage type
qPCR	quantitative PCR
qRT-PCR	real-time quantitative reverse transcription PCR
RAJ	rectal-anal junction
RBBC	repeated bead beating and a column
RMT	random matrix theory
SIP	stable isotope probing
SNPs	single nucleotide polymorphisms
SOM	self-organizing map
SS	super-shedders
STEC	Shiga toxin-producing *Escherichia coli*

Stxs	Shiga toxins
T3SS	Type III secretion system
Tir	translocated intimin receptor
TPM	transcripts per million
USDA	the U.S. Department of Agriculture
VFAs	violate fatty acids
Zi	within-module connectivity

Chapter 1. Literature Review

1.1 Introduction

Shiga toxin-producing *Escherichia coli* (STEC) is an important foodborne bacterium that can cause many severe human illnesses such as bloody diarrhea, hemolytic uremic syndrome (HUS), and death (Croxen et al., 2013). STEC infections are frequently reported in the United States, Canada, Japan, and the European Union (Chase-Topping et al., 2008), among which Scotland (Dundas et al., 2001), Canada (Waters et al., 1994), USA (Griffin and Tauxe, 1991) exhibit the higest prevalence of STEC. In Canada, the incidence of STEC infection from 1990 to 1999 has remained stable at 4.1 to 7.1 per 100,000 persons with an average of 1,407 cases reported annually (Lisboa et al., 2019). From 2006 to 2015, approximately 730 annual STEC cases were reported in Canada, however, for each reported STEC case, it is estimated 10-47 cases unreported each year, suggesting the annual incidence of STEC cases maybe ten times higher (Kate et al., 2006; Reynolds et al., 2020). Of the Canadian provinces, Alberta has one of the highest rates of STEC infection (Galanis et al., 2003). From 2006 to 2016, 1,526 cases of STEC were reported in Alberta, with an average of 139 reported cases per 100,000 people (Galanis et al., 2003). Outbreaks of STEC not only cause health issues but also result in remarkable economic losses. A 2003-2006 epidemiological study in Australia estimated the annual cost of reported STEC cases AUD$ 2,633,181 per year (McPherson et al., 2011), even with the low incidence of STEC infection (0.4 per 100,000 in 2006) (Group, 2018).

Ruminants, especially cattle, are regarded as the primary asymptomatic carriers of STEC and the major source of STEC infections in humans (Chapman et al., 1993). STEC contamination of the environment can occur on farms as a result of defecation by carrier cattle (Rangel et al., 2005; Heiman et al., 2015). Therefore, preharvest controls are desirable for lowering STEC contamination and cattle-human STEC transmission. Consequently, the focus on exploring cattle-STEC interactions and their potential in the development of on-farm STEC mitigation technologies is important as STEC remains one of the leading threats to human health.

1.2 Stx positive Shiga toxin-producing *Escherichia coli*

Escherichia coli (*E.coli*) are gram-negative facultative anaerobic bacilli. These bacteria are commensals that are normally found in the intestine of mammals but are also found in the gut microbiome of birds, and fish as well as associated with soil, water, plants, and food (Hartl and Dykhuizen, 1984; Leimbach et al., 2013). *E.coli* can be classified into three groups based on its pathogenicity: commensal, potentially pathogenic as well as pathogenic *E.coli* (Figure 1.1A) (Proença et al., 2017). Commensal *E.coli* is one of the most common bacteria inhabiting the lower intestine of mammals (Blount, 2015), while only accounting for approximately 0.1-5% of the total bacterial flora within the gut (Eckburg et al., 2005; Blount, 2015). They play a beneficial role to their hosts such as they can exclude pathogenic bacteria (*e.g. Salmonella typhimurium*) (Schierack et al., 2011) as well as produce vitamins (*e.g.* K2, B12)(Blount, 2015). However, pathogenic *E.coli* can possess virulence factors that contribute to a number of human

diseases including urinary tract infections, diarrhea, and meningitis (Nataro and Kaper, 1998; Gomes et al., 2016). Some *E.coli* can be considered to be potentially pathogenic. Based on mechanisms by which pathogenic *E.coli* causes disease, pathogenic *E.coli* can be typically classified into the following four pathotypes including Shiga toxin producing *Escherichia coli* (STEC), Enterotoxigenic *Escherichia coli* (ETEC), Extraintestinal *Escherichia coli* (ExPEC) and Enteropathogenic *Escherichia coli* (EPEC, Table 1.1, Figure 1.1B). *E.coli* pathotypes can cause various animal diseases (Table 1.2), among which STEC is the major pathotype that causes severe disease in humans, including hemolytic uremic syndrome, bloody diarrhea, and even death (Croxen et al., 2013). The clinical outcomes caused by STEC were typically due to the production of Shiga toxins in STEC infected humans.

1.3 Virulence factors in STEC

1.3.1 Shiga toxins (Stxs)

1.3.1.1 Stxs classification and structure

Shiga toxin is the main virulence factor in STEC and can be classified into two prototypes, stx1, and stx2 (Scheutz et al., 2012). Stx1 and stx2 share only 56% amino acid sequence similarity and are immunologically distinct (Jackson et al., 1987). Based on the phylogenetic analysis of stx sequences, several subtypes of stx1 and stx2 were identified which play a differential role in the severity of human illnesses (Table 1.3, Melton-Celsa, 2014a). The stx2 is the main attribute causing human severe infections (*i.e.* HUS) (Cimolai et al., 1994). The stx2-producing STEC has been reported to be found in 71 % (34 out of 48) of children with HUS while only 40% (4 out of 10) of patients were associated with stx1-producing STEC strains (Ludwig et al.,

2001). Furthermore, stx2 is 400 times more toxic (as quantified by LD50 in mice) than stx1 (Tesh et al., 1993).

Stx1 and stx2 share the same AB5 protein structure composed of a single A subunit (32 kDa) and five binding B subunits (7.7kDa) (Figure 1.2A, B) (Croxen et al., 2013). The trypsin-sensitive region enables stx1 to be asymmetrically cleaved into the A1 subunit and A2 peptide connected by a disulphide bridge (Figure 1.2C, Melton-Celsa, 2014a). The enzymatic activity of stx is related to the subunit A1, while the A2 peptide is linked to A1 to the binding moiety and connected to the B pentamer (Figure 1.2C, Melton-Celsa, 2014a). The identical B subunits are responsible for binding to the cellular receptor globotriaosylceramide (Gb3 or CD77) in humans that is present in several organs such as the intestinal tract and kidney (Etcheverría and Padola, 2013). In addition to binding affinity to the Gb3 receptor of stx, one exceptional variant toxin stx2e preferentially recognizes Gb4 (Gb4 is derived from Gb3 and is composed of tetra-saccharide) over Gb3 (Gb3 is composed of tri- saccharide.) (DeGrandis et al., 1989; Samuel et al., 1990).

1.3.1.2 Pathogenesis of stx in human

The hemolytic uremic syndrome is the typical clinical outcome of tissue damage induced by stxs in humans (Nataro and Kaper, 1998). The HUS can be classified as typical and atypical HUS (Sheerin and Glover, 2019). The typical HUS was induced by the binding affinity of stx to the Gb3 receptor on the vessels of the kidney followed by the endothelial injury (Andreoli et al., 2002), while atypical HUS can be attributed to the acquired defect in the control of complement activation and such atypical HUS also predominantly affect the kidney (Sheerin

and Glover, 2019). In this review, the molecular pathogenesis of stx inducing the typical HUS is illustrated.

In STEC, the *stx1* and *stx2* genes are encoded in the genome of lambdoid prophages (Łoś et al., 2020). Without prophage induction, the expression of *stx* genes is frequently suppressed (Łoś et al., 2020). Hence, in most cases, the prerequisite for effective production of Shiga toxin requires prophage induction and its further lytic development including replication of the phage genome as an extrachromosomal element (Schmidt, 2001; Krause et al., 2018) (Figure 1.3). The specific location between the genes encoding the late anti-terminator Q and the lysis enzymes in the phage genome allows *stx* to be co-transcribed with the late phage genes (Wagner et al., 2002) (Figure 1.4), with the translation of the *stx* gene occurring when the prophages are induced. Prophage induction is associated with the bacterial SOS response, which mediates one of the main regulation pathways for stx production (Serra-Moreno et al., 2008). Particularly, the production of stx1 can be induced by lower iron levels in the gut (Weinstein et al., 1988). Nonetheless, stx1 cannot be transported outside the cell without prophage induction and host cell lysis as *E. coli* lacks an appropriate secretion system for stx1 (Weinstein et al., 1988). The production of stx2 is dependent on transcriptional activity of the late region of phage which is expressed at later stages of induction. In addition to phage induction and bacterial lysis affecting stx production, environmental factors including temperature, oxidative stress, quorum sensing, and antibiotics also impact stx production (Mühldorfer et al., 1996; Krause et al., 2018).

After bacterial lysis and phage induction, stx binds to the Gb3 receptor present on gut epithelial cells in humans, which further initiates the retrograde pathway of stx (Figure 1.5) and results in host cell death (Melton-Celsa, 2014a). The outline of the retrograde pathway is summarized as follows. After stx binds to the Gb3 receptor, the stx-Gb3 complex then enters into the early endosome through the lipid raft and moves to the Golgi apparatus where it can reach the endoplasmic reticulum (ER) (Sandvig et al., 2010). It is noticeable that the Gb3 receptor is necessary for the stx-receptor complex to move through the retrograde pathway since a non-receptor mediated mechanism prohibits stx from reaching the Golgi apparatus (Philpott et al., 1997; Raa et al., 2009). The disulphide bond connecting A1 and A2 subunits (Figure 1.2C) is reduced once the toxin enters the endoplasmic reticulum with only the A1 toxin subunit entering the cytosol (Melton-Celsa, 2014a). The unfolded A1 subunit exits ER through the ER-associated protein degradation (ERAD) pathway to reach the target ribosome (Spooner and Lord, 2011). The A1 subunit further removes an adenine from the 28S ribosomal subunit, inhibiting protein synthesis and causing a disconnection with elongation factor 1 (Melton-Celsa, 2014a). Furthermore, the ribosome induces a pro-inflammatory and pro-apoptotic 'ribotoxic stress response' in cells as a result of stx-mediated damage (Spooner and Lord, 2011). At the same time, macrophages can induce innate immunity by upregulating cytokine and chemokine expression through the stx-regulated signalling pathway (Lee and Tesh, 2019). The pathophysiological consequences of stx-induced innate immunity include changes in cell morphology and intracellular tight junctions to facilitate stxs crossing into the lamina propria where they damage colonic blood vessels and initiate hematogenous spread which can result in

watery or bloody diarrhea. The toxins can facilitate the infiltration of inflammatory cells into the gut lamina propria and kidneys and can up-regulate the expression of Gb3 on microvascular endothelial cells, deconstructing red blood cells and damaging the blood clotting functions of the kidney in the infected people (Wang et al., 2014; Brandelli et al., 2015; Legros et al., 2018). Overall, stx can destroy the host cell structure and alter their morphology in a manner that enables the toxins to circulate throughout the body.

1.3.2 Other virulence factors

It is noticeable that STEC colonized on the epitheliums and the ability of extracellular appendages fimbriae to adhere to host surfaces is important for successful colonization by such pathogen (Antão et al., 2009). The protein associated with the fimbriae called adhesins, which directs high-affinity binding to specific cell surface components (Antão et al., 2009). There are different types of fimbriaes including Type 1 fimbriae (*fim*), P fimbriae (*pap*), S fimbriae (*sfa*), etc and these fimbriaes are host cell- specific (Antão et al., 2009). For instance, Type 1 fimbriae can adhere to human balder epithelium, chicken tracheal and gut explants, while S fimbriae can bind to human brain endothelium (Antão et al., 2009). The type 1 fimbriae plays a significant role as mediators of attachments by *E.coli* infections in the pathogenesis of Gram negative bacteria. And the presence of such fimbriae system could likely ensure successful bacteria colonization and the host cell-specific receptor production will then determine the site of action for a given fimbriae during STEC colonization and infection (Ahmed et al., 2008, Antão et al., 2009).

Most disease-related STEC serogroups contain a chromosomal pathogenicity island termed the locus of enterocyte effacement (LEE) (Perna et al., 1998). The LEE enables STEC to attach to the host intestinal mucosa and initiate the destruction of the microvillus brush border (Frankel et al., 1998). The attachment is enhanced by the interactions between intimin, a receptor expressed on the cell membrane of STEC, and the nucleolin of host cells (Puente and Finlay, 2001; Croxen et al., 2013). Once attached, STEC injects various effector proteins into the host cell as mediated by Type III secretion system (T3SS) (Figure 1.6), which enhances the attachment of STEC and induces cytoskeletal rearrangements (Deng et al., 2017). The translocated intimin receptor (Tir), a receptor encoded by STEC can be translocated into host cells via T3SS (Gaytán et al., 2016). The T3SS also dysregulates actin polymerization within host cells, resulting in the formation of an actin-rich pedestal structure that facilitates the adherence of STEC to the host intestinal epithelium (Gaytán et al., 2016). Formation of actin-rich pedestal structure and STEC colonization further alter host cell morphology including the removal of microvilli on the surface of epithelium resulting in the formation of attaching and effacing lesions (A/E lesions) (Ji and Dong, 2015).

In addition, effector proteins have anti-phagocytosis functions that halt host immunities. For instance, *Escherichia coli* secretes protein B (EspB) which interacts with actin-binding domains of myosin proteins to prevent phagocytosis (Ji and Dong, 2015). The nuclear factor-κB (NF-κB) and mitogen-activated protein kinase (MAPK) signalling pathway in host cells can also be interrupted through non-Lee effectors injections (NLeE, B, C, D, H) produced by T3SS, leading to the potential inhibition of host innate and adaptive immunity (Newton et al., 2009;

Bliska et al., 2013; Deng et al., 2017). The T3SS components together with secreted effector proteins, intimin, and its receptor *Tir*, as well as attaching and effacing (*eae*) genes are encoded within a genomic pathogenicity island termed as the aforementioned Lotus of Enterocyte Effacement (LEE) (Gaytán et al., 2016). A previous epidemiological study characterized 60 STEC isolates from bovine feces, among which 56 (93%) isolates were LEE- and stx-positive (Mainil et al., 1993). The LEE-positive STEC is commonly detected in cattle which is regarded as the natural reservoir for STEC and is responsible for most human STEC infections (Mainil et al., 1993; Newton et al., 2009). In addition to Shiga toxins, other virulence factors (e.g. T3SS) that participate in human STEC pathogenesis and cattle colonization warrant further explorations.

1.4 STEC Epidemiology

1.4.1 Cattle as the major reservoir of STEC and super-shedders

Most human STEC infections are attributed to the consumption of contaminated beef products, as indicated by survey data across 27 countries from 1998 to 2017, where beef products were linked to about 22% (210 out of 957) of STEC outbreaks. What's more, fecal shedding of STEC varies among cattle, ranging from 10 to 10^7 CFU per gram of feces. The frequency of STEC in cattle also varies across different studies. Venegas-Vargas et al. (2016) surveyed fecal samples from 378 beef cattle and found 80 (21%) were positive for STEC, while Onyeka et al. (2021) observed that 147 out of 419 (35.1%) fecal samples from beef cattle were positive for STEC in South Africa.

Cattle that shed fecal STEC > 10^4 CFU per gram of feces are designated as super-shedders (SS, Wang et al., 2016). SS usually account for approximately 10% of cattle in a herd, from which these individuals are responsible for more than 90% of STEC that enters the ambient environment. Shedding of STEC is often intermittent and the level of shedding often varies substantially. Baines et al. (2008) suggested that SS can be classified into three types based on the fecal shedding duration, including non-persistent (shed less than 14 days), moderately persistent (shed about 30 days), and persistent SS (shed for several months) (Baines et al., 2008). STEC can colonize the intestinal tract of cattle without causing hemorrhagic diseases since Gb3 receptors are not expressed on the surface of bovine epithelium (Paton and Paton, 1998; Menge, 2020). Among the digestive tract, the lower gastrointestinal tract is an ideal niche for STEC colonization and proliferation (Stein and Katz, 2017). Researchers have suggested that the rectal-anal junction (RAJ) is the major colonization site for STEC, and that colonization of the RAJ is often associated with a high level of STEC shedding (Wang et al., 2016).

E.coli O157:H7/HM (STEC O157), is a key pathogenic STEC that colonizes beef cattle and can cause contaminations of food products and human illness (Lim et al., 2010). Particularly, the O represents the O-polysaccharide, a part of the lipopolysaccharide (LPS) on the cell membrane of Gram-negative bacteria. The H refers to the flagellar antigens and combinations of O and H factors define the serotypes of isolated STEC. In addition to STEC O157, recent studies suggest that the 'big six' including STEC O26, O45, O103, O111, O121, O145 represent the emerging foodborne STEC that colonize beef cattle and threaten human

health (Alharbi et al., 2022). Identifying STEC O157 and non-O157 strains and super-shedders in cattle is important as SS can continuously spread STEC into the surrounding environment and be inadvertently transmitted to farm workers and increase cattle-cattle and cattle-human transmission (Paton and Paton, 1998; Menge, 2020).

1.4.2 Cattle-human STEC transmission and STEC serotypes

The fecal-oral route is the major transmission route for human STEC infection (Fairbrother and Nadeau, 2006; Maluta et al., 2014; Browne et al., 2021). Zoonotic STEC can be ingested by cattle and other ruminants and then colonize the intestinal tract without causing disease in these hosts. As STEC is shed in the feces, it can contaminate the environment including water used for drinking and recreation. The contamination of fruits, vegetables, sprouts, and lettuce via run-off water, manure, or slurry can all act as a source of transmission. Milk and meat can also be a source of transmission of STEC if they become adulterated. People working on farms or in slaughterhouses or visiting farms or petting zoos may also acquire STEC through direct contact with the host. The direct person to person transmission can also occur (Figure 1.7).

The STEC genome can also contain different stx prophages and non-stx prophages (Rodríguez-Rubio et al., 2021), among them STEC O157 phage types 21/28 (PT 21/28) were most prevalent in SS and frequently detected in patients with STEC infections in Scotland (Corbishley et al., 2014; Fitzgerald et al., 2019). In Scotland between 1997 and 2001, 61 % of HUS cases in children were caused by PT21/28 (Lynn et al., 2005). In one study of 88 Scottish farms in 2006, approximately half of STEC O157 isolated were PT 21/28 (Halliday et al., 2006). Furthermore, the PT21/28 was also more likely to be associated with high levels of shedding in

cattle, increased prevalence of super-shedders and cattle-human transmission (Halliday et al., 2006).

In addition to STEC O157 which causes human illnesses, more than 200 non-O157 serotypes of STEC have been identified from HC and HUS patients. The Center for Disease Control (CDC) in the US estimated that approximately 64 % (169,600 out of 265,000 cases) of STEC infections were attributed to non-O157 infections each year (Scallan et al., 2011). Among non-O157 serotypes, O26 (22%), O111 (16%), O103 (12%), O121 (8%), O45 (7%), and O145 (5%) were the most common serotypes isolated from infected people and were designated as the 'Big Six'(Yoon et al., 2013). The Food Safety and Inspection Service (FSIS) of the U.S. Department of Agriculture (USDA) further declared the 'Big Six' as food adulterants. The Hazard Analysis and Critical Control Points (HACCP) program was implemented to prevent the sale of raw beef and its products contaminated by these STEC serotypes (Yoon et al., 2013). Therefore, identifying and developing mitigation strategies for STEC O157 and non-O157 in beef cattle could be critical to preventing cattle-human STEC transmission.

1.5 Host-STEC interaction affects STEC colonization in cattle

1.5.1 Host factors affecting STEC colonization

1.5.1.1 Host genetics

A study found that STEC O157 isolates were more likely to cause human clinical disease when human had polymorphisms in the 255th nucleotide of the Tir gene (T>A T allele), an observation often observed with isolates acquired from SS vs. NS (Peroutka-Bigus et al., 2022).

This study also confirmed that the host genetic mutation was associated with *E.coli* O157's colonization of the bovine epithelium, which increased the occurrence of SS. Wang et al. (2017) evaluated the differential expression of genes in rectal tissues between SS and NS and found 33 single nucleotide polymorphisms (SNPs) in seven DE genes associated with host immunities and cholesterol transport which potentially contributed to STEC colonization. The same study also revealed that these SNPs could lead to higher level of B-cell signaling, T-cell responses, and cholesterol absorption in the gastrointestinal tract in SS compared to NS, resulting in different host responses.

Mir et al. (2016) conducted a cohort study for assessing the prevalence of STEC from a multi-breed beef calf population derived from Brahman and Angus, and collected fecal samples from March (n = 259), June (n = 263), August (n = 261) and December (n = 193). There was no significant difference in STEC prevalence across each hybrid breed group, suggesting between-breed variation plays a limited role in STEC colonization and possibly super-shedding. However, presumably STEC colonization is affected through various mechanisms such as immune responses and gut microbiota. Therefore, further epidemiological studies together with genotypes and/or SNPs analysis are needed to define the role of breed/cattle genetics in STEC epithelium colonization and super-shedding.

1.5.1.2 Host physiological stages

The host physiological stages can be manifested by host ages, which together with diets and environmental factors collectively influence fecal shedding. A previous study collected fecal samples in March, June, August, and December from calves that were 1–3, 4–6, 7–9, and 10–

12 months of age, respectively (Mir et al., 2016). There was a significantly higher prevalence of STEC in young calves (1–3 months old, 60.3%) compared to older calves (39.5% at 4–6 months age, 20.3% at 7–9 months age, 20.7% at 12 months old, $P < 0.00$, Mir et al., 2016). While Jeong et al. (2015) observed that the prevalence of STEC in cows was not age dependent as the prevalence was comparable between older (15 months old, 26 out of 42, 61.9%) and younger cows (3 months old, 30 out of 42, 71.4%, P=0.32). These results contrasted with those of Mir et al. (2016), suggesting host physiological stages together with diet, environmental factors influence STEC colonization and super-shedding.

1.5.1.3 Dietary factors

Diet is known to impact the ruminal and intestinal microbiota in ruminants and is suggested to be associated with STEC shedding in cattle. A previous study explored the effect of diet cofounded with physiological changes on *E.coli* O157 shedding and found that changes in diet from pre-weaning to weaning could increase STEC colonization and *E.coli* O157 shedding (Venegas-Vargas et al., 2016). Thirty calves were weaned and fed a corn silage–based diet (High Moisture, HM) during the weaning and preconditioning period. The author reported that *E.coli* O157 increased from 16.6% before weaning to 38.3% at 14 days after weaning ($p<0.05$) and stayed at a higher level during the preconditioning period (56 days after weaning). One possible reason for this is that weaning stress (physiological changes) might decrease host immune functions (Kim et al.,2012), increasing O157 colonization (Naylor et al.,2003). Diez-Gonzalez et al. (1998) and Herriott et al. (1998) also reported that a low-grain diet resulted in lower O157 shedding, and that feeding corn silage significantly increased the risk of

enterohemorrhagic *E.coli* shedding among heifers. Inconsistencies in the impact of diet on the shedding of *E. coli* O157 in cattle have also been reported that dry grain-fed cattle had reduced *E. coli* shedding levels. Callaway et al. (2009) suggested that this inconstancy may arise from intrinsic factors of forages such as the forage quality, and the presence of antimicrobials may also account for the variability observed across different studies. What's more, the concentrations of metabolizable substrates available for fermentation in the lower bovine intestinal tract might be a contributing factor to the higher level of O157 shedding (Callaway et al.,2009). Buchko et al. (2000) reported that feeding forage 48h after fasting increased *E. coli* O157 abundance in the intestinal tract and shedding duration, resulting in greater colonization and higher prevalence of *E. coli* O157 in cattle. It is reasonable to assume that fasting which decreases violate fatty acids (VFAs) concentrations in the rumen and intestinal contents, may reduce the extent to which they inhibit the growth of *E. coli* O157. These studies highlight the complicated role of dietary factors in STEC colonization and shedding in cattle generated by internal diet factors, diet composition, as well as host physiology.

1.5.2 STEC colonization affects host immunity

1.5.2.1 Host innate response

The activation of the innate immune system is the frontline of the host defence against bacterial invasion. The effective stimulation of the ruminant innate response occurs since STEC closely interacts with intestinal epithelial cells. Previous studies observed activated host innate immunity in cattle challenged with *E.coli* O157. For example, *E. coli* O157 challenge via stomach tube can cause neutrophil infiltration in the colon, ileum, and rectum of neonatal calves

(Dean-Nystrom et al., 1997). Calves challenged with *E. coli* O157 have also exhibited infiltration of neutrophils and eosinophils in the mucosa of the gallbladder (Reinstein et al., 2007).

As aforementioned STEC could colonize the bovine epithelium via T3SS, which translocates effector proteins (*i.e.* Tir, EspA encoded within the LEE and non-LEE genomic regions) into infected cells (Deng et al., 2017). Simultaneously, H7 flagella can initiate interactions with the bovine epitheliums that can be recognized by Toll-like receptors 5 (TLR5, (Deng et al., 2017). A signaling cascade mediated by the myeloid differentiation factor 88 (MyD88) adaptor molecule is initiated, leading to the activation of MAP kinases and IκB kinase (IKK) (Hayashi et al., 2001). The proteasomal degradation of IκB liberates the IκB-bound NF-κB transcription factors, which translocates to the nucleus to drive the expression of pro-inflammatory cytokines such as IL-8, IL-1β and TNF-α (Häcker and Karin, 2006, Miyamoto et al., 2006). The host develops an immune response when STEC injects the non-LEE encoded NleB, NleC, NleE, and NleH effectors into host cells to disrupt the host NF-κB pathway (Newton et al., 2009; Nadler et al., 2010). Overall, the balance between the pro-inflammatory bacterial extracellular components such as H7 flagellin, and the anti-inflammatory effector proteins such as the Nle family shapes the outcome of host immune response.

In addition to the H7 flagellin recognized by the bovine innate immune system, STEC LPS is also recognized by bovine monocyte-derived macrophages (Bono et al., 2004) as well as bovine colonic cells expressing TLR1, TLR3, TLR4 and TLR5 (Walle et al., 2013). The

signaling pathway induced by LPS has not been elucidated to date, but TLR4-mediated signaling could be the target as the binding of LPS to TLR4 has been demonstrated for *E.coli* O111:B4 in bovine monocyte-derived macrophages (Magee et al., 2012). In summary, STEC colonization in beef cattle could trigger varied host innate responses, leading to a number of host-STEC interactions.

1.5.2.2 Host cellular response

Studies on host cellular immune response in cattle to STEC are limited and therefore the role of host cellular immunity and pathways involved in STEC colonization is unclear. Corbishley et al. (2014) suggested that T-cell responses were induced in calves challenged with two different STEC O157 strains (a PT32 and a PT21/28 STEC O157 strain), resulting in the enhanced expressions of IFN-γ within the rectal mucosa of calves. In vitro stimulation of rectal lymph node cells from the same calves with T3SS proteins led to the proliferation of CD4, CD8, and γδ T cells in PT21/28 challenged calves, while the proliferation of NK, CD8, γδ T cells occurred in calves challenged with PT32 (Corbishley et al., 2014; Fitzgerald et al., 2019). Hence, these results suggest that cattle can develop strain-dependent cellular responses during STEC colonization but further clarification of the cellular immunity responses in cattle colonized with STEC is needed.

1.5.2.3 Host mucosa and humoral antibody response

Since STEC are able to colonize the epithelium, mucosal antibody responses are considered to be the front line of the host defense against STEC colonization. Significant rectal mucosal IgA antibody responses have been characterized in response to EspA, EspD, EspD, Tir, H7, OmpC,

and O157:H7 LPS in calves experimentally challenged with an stx-negatived *E.coli* O157 strain (Nart et al., 2008). Whereas rectal mucosal IgA antibody responses were also detected in calves following challenge with a stx-positive STEC O157 strain (Nart et al., 2008). However, the titers of IgA antibodies were inconsistent across different individuals, suggesting variation in host protective responses. To date, the extent and mechanisms whereby mucosal antibodies protect against STEC colonization are largely unknown.

STEC-specific humoral antibody responses were demonstrated to develop in cattle following oral challenge. A significant increase in serum IgG specific to STEC O157 inimin, Tir, EspA, EspB, and O157 LPS was observed in cattle orally challenged with a stx-positive STEC O157 strain (Bretschneider et al., 2007), suggesting the host can serologically respond to STEC O157 T3SS proteins and LPS. Bretschneider et al (2007) also highlighted Tir-, intimin-, and EspB- specific serum IgA were decreased following oral challenge with a STEC O157 stx-positive strain due to the downregulation of mucosal IgA. Antibody responses specific to stx1 and stx2 have also been found in sera in calves orally challenged with STEC O157 but the development of these antibody responses is delayed (Bretschneider et al., 2007).

Researchers have also demonstrated that serum IgG levels (*i.e.* Tir-, intimin- and O157 LPS- specific) are highly correlated with fecal shedding of STEC O157 (Bretschneider et al., 2007). However, systemic antibody responses against STEC antigens are not always associated with bacterial shedding or clearance. Wray et al. (2007) demonstrated no serological response to STEC O157 LPS antigen following oral challenge with STEC O157 in adult cattle. Nalyor et al (2007) showed that increased serum IgG and IgA for O157 LPS and H7 antigen production

were observed in calves orally challenged with a stx-positive STEC O157 strain, but these antibody responses were not correlated with bacterial shedding. Since the shedding of STEC is a complex process that is affected by multiple factors (*i.e.* microbial communities, environments) (Williams et al., 2015), further studies about STEC shedding and host immune response interactions are needed.

1.5.3 Rectum microbiota affects STEC colonization

1.5.3.1 Cattle gut microbiota and microbial community assembly

Bovine gut microbiota play a critical role in affecting STEC shedding and has been proposed to regulate the host immune system, and nutrient metabolism (Nicholson et al., 2012). Gut commensals are capable of inhibiting the growth of pathogenic microbes through direct (*i.e.* releasing antimicrobials) and indirect effects (*i.e.* host immunity activation, Buffie and Pamer, 2013). Gut microbial communities were gradually established with two theories for explaining microbiome assembly: the niche theory and neutral theory (Zhou and Ning, 2017). The deterministic process is a key feature of the niche theory that proposes microbial community assembly is structured by environmental heterogeneity and biotic (inter taxa interactions) factors (Zhou and Ning, 2017; Liu et al., 2019). Unlike the niche theory which assumes that taxa are functionally different, the neutral theory assumes that all individuals are ecologically equivalent and microbial interactions and abiotic factors-microbial interactions are ignored (Hubbell and Borda-de-Água, 2004). Under the niche theory, microbial diversity is controlled by stochastic processes (also termed historical contingency) such as ecological drift (change in the relative abundance of taxa in a location due to chance demographic fluctuations) and

dispersal (movement of taxa across spaces, Chave, 2004; Hanson et al., 2012; Herbert et al., 2014). Previous studies explored how stochastic and deterministic factors contribute to microbial community assembly (Table 1.4). For instance, a recent study revealed that stochasticity is the major force driving the rumen microbiome assembly from birth to maturity, a process that is affected by both external (*i.e.* diet) and host factors (*i.e.* age), suggesting that neutral theory based microbial community assembly occurs in beef cattle (Furman et al., 2020). Since STEC colonizes the hindgut of beef cattle along with other gut microbiota, they could influence the assembly of gut microbial communities, but the nature of this interaction remains unclear. Therefore, integration of the microbial community assembly approach and community analysis could provide a useful framework for understanding the dynamics of gut microbiota in response to STEC colonization in beef cattle.

1.5.3.2 Rectum microbiota and its functions

The rectal-anal junction (RAJ) is the major colonization site of STEC, and the RAJ is known to be colonized by commensal microbiota. Previous studies identified that the rectal mucosa and fecal microbiota differ in terms of their composition and function and were both related with cattle STEC colonization (Xu et al., 2014; Wang et al., 2016; Zaheer et al., 2017). Therefore, the following section mainly focuses on the differed composition and functions of hindgut mucosa microbiota and fecal microbiota in beef cattle.

A study examined the microbial communities collected from rectal mucosa and feces from Holstein dairy calves aged five years old and found a remarkably higher Chao1 and Shannon index (refer to microbial richness and evenness, respectively) in rectal mucosa-

attached microbial communities compared to fecal microbiota (Mucosa: Chao1=2346, Shannon=5.63; Fecal: Chao1=1526, Shannon=3.49) (Mao et al., 2015). The unweighted UniFrac distance-based principal coordinate analysis (PCoA) further revealed distinct microbial cluster profiles between mucosa and fecal microbiota (Mao et al., 2015). Further analysis from the same study revealed varied microbial community composition at the phylum and genera levels between rectal mucosa- and fecal microbiota. Particularly, Firmicutes (91.26%) was the dominant phyla with *Turicibacter* (14.23%) and *Clostridium* (14.72%) being the predominant genera in fecal microbial communities. While Firmicutes (46.57%) and Bacteroidetes (30.94%) were the predominant phyla with *Ruminococcaceae* (18.16%) and *Treponema* from Spirochaetae (9.13%) being the predominant genera in rectal mucosa microbial communities. However, functional analysis revealed that both rectal mucosa and fecal microbiota shared similar functions related to membrane transport, carbohydrate metabolism, replication and repair, amino acid metabolism, and energy metabolism (Mao et al., 2015).

1.5.3.3 Rectum microbiota and its role in STEC shedding and colonization

Some bacteria (*i.e. Lactobacillus*) have been reported to reduce *E.coli* O157 shedding in cattle due to their ability to stimulate the development of host immunity and lower intestinal pH (Sherman et al., 2005). Further studies proved that rectum mucosa-attached microbial communities interacted with host immunity in a manner that could affect *E.coli* O157 fecal shedding (Wang et al., 2017). This suggested that there could be a potential role of rectum mucosa-attached microbiota in regulating STEC colonization and *E.coli* O157 fecal shedding.

Wang et al 2018 reported significant variations in the microbial composition and function of the rectal mucosal microbiota between SS and NS. For example, two taxa were unique to SS which were members of *Bacteroides* and *Clostridium*, while seven taxa were unique to NS which were members of *Coprococcus*, *Provetella*, *Clostridium*, and *Paludibacter* (Wang et al., 2018). These unique microbes in the NS group, for example, *Coprococcus* spp., play a role in butyrate production, possibly contributing to a lower gut pH and an environment unfavorable for *E.coli* O157 colonization (Holdeman and Morre, 1974).

Comparisons of bacterial diversity and microbial composition in fecal microbiota between SS and NS have been controversial. A previous study revealed that microbial community structures were affected by shedding status and were different in terms of bacterial compositions in the first 21 days in postpartum dairy cattle, *i.e.* persistent shedders were seen to have lower average fecal microbial richness compared to NS (Stenkamp-Strahm et al., 2018). In contrast, another study using beef steers suggested that fecal bacterial structures were not associated with shedding status but differed in terms of microbial composition between SS and NS (Zaheer et al., 2017). Given different outcomes of fecal microbiota variations in SS compared to NS, host factors (*i.e.* breed, age, sex) could also affect STEC colonization, suggesting that supper shedding could be driven by host-microbe interactions. Further studies considering host factors (*i.e.* age, genetics, host immunity) for exploring fecal microbiota- super shedding are warranted.

1.5.3.4 Host-microbial interactions at rectum mucosa and relations with STEC shedding

Host- gut microbiome interactions play a critical role in maintaining host gut homeostasis and during host gut dysbiosis (Zhou et al., 2022). Specifically, the gut microbes in the gut exert control at the epithelial barrier. During homeostasis, commensal microbiota can produce metabolites through the fermentation of dietary polysaccharides and the microbial molecular ligands (such as LPS) and therefore impact intestinal barrier cells (Zhou et al., 2022). The host produces IgA in plasma cells in the lamina propria which are then transferred across the epithelium as secretory IgA (Zhou et al., 2022). Induced IgA can bind to commensals, forming a reciprocal feedback circle between host-symbiont during homeostasis. During host gut dysbiosis, the commensal microbial community can become unbalanced with unwanted outgrowth of gut microbes and invasion of pathogens (*i.e. Salmonella* spp. is an invasive flagellated pathogen that can evade commensal resistance.), promoting host immune cell recruitments to the epithelium to combat these gut microbes (Kaiko and Stappenbeck, 2014). The knowledge of gut microbes exerting effects on host gut health provides us the biological basis for furthering our understanding on host-microbial interactions in beef cattle.

However, the interaction of microbes with the hindgut epitlhelim in cattle has been rarely studied. A previous study found that host-hindgut microbial interactions affect bovine growth and immunity, and host genetics in turn exert lifelong effects on hindgut microbiota (Fan et al., 2021b). A total of 278 Angus-Brahman beef cattle were designated to six subgroups based on their breed composition (percentage of Angus and Brahman). Fecal samples were collected from calves at preweaning, postweaning, and fattening stages. The study revealed that

the hindgut microbiota structure was affected by host growth stage and cofounded with diet transactions and breeds throughout life. Butyrate-producing bacteria, *Roseburia* and *Oscillospira*, were related to nine SNPs located in the genes involved in host immunity and metabolism regulation in the hindgut, suggesting host-microbial interactions in the hindgut are critical for host functional variation and the establishment of microbial communities. A previous study in our research group revealed that hindgut mucosa-attached microbiota were associated with host immunity in response to STEC O157 colonization in SS, and that host-microbial interactions regulated *E.coli* O157:H7 fecal shedding in beef cattle (Wang 2018). However, the nature of host-microbe interactions in the hindgut of beef cattle in response to strain-specific STEC O157 is largely unknown.

1.6 Approaches to study host-microbial interactions in response to STEC colonization

1.6.1 Molecular techniques for Shiga toxin gene quantification

1.6.1.1 Real-time quantitative PCR (qPCR) and real-time quantitative reverse transcription PCR (qRT-PCR)

The polymerase chain reaction (PCR) is a fast and reliable technique used to detect DNA, and real-time quantitative PCR (qPCR) can be used to measure the abundance of amplified DNA (Bustin et al., 2009). While real-time quantitative reverse transcription PCR (qRT-PCR) was developed to measure targeted expressed genes (mRNA) based on the reverse transcription of DNA and q-PCR. The q-PCR and qRT-PCR are techniques that enable reliable detection and measurements of gene expression generated during each cycle of the PCR process (Bustin et

al., 2009). These procedures are also capable of detecting genes at low abundance more effectively than other molecular techniques (*i.e.* microarrays). As a result, they are considered the gold standard for measuring gene expression (Bustin, 2005). However, the selection of suitable primers is the key to generating reliable results since primers can target specific sequences and the optimization of reaction conditions can also be time-consuming. SYBR Green is the major economical DNA-binding dye used for q-PCR and qRT-PCR (Bulcke et al., 2010), however, it can interact with all double-stranded DNA, including non-specific amplicons which can cause errors. Using fluorescent reporter probes, such as Taqman (Ponchel et al., 2003), can improve specificity and accuracy because fluorescence can only be detected in amplicons that contain complementary sequences to the reporter probe. However, qPCR and qRT-PCR are not suitable to detect the abundance or expressions of a large number of genes, or to detect unknown genes/transcripts.

1.6.2 Studying host gene expression and microbial communities using high-throughput sequencing

1.6.2.1 RNA-seq for host gene expressions

RNA-seq refers to the analysis of gene expression profiles using next-generation sequencing techniques (Stark et al., 2019). It is an in-depth genome sequencing technique that investigates all genes expressed within a biological system and uncovers novel and rare transcripts. However, there are some challenges for RNA-seq including a lack of quantification of absolute gene expression as well-annotated reference genomes are acquired to obtain accurate information from RNA-seq data (Stark et al., 2019). For instance, most of RNA-seq analysis

require that sequenced gene fragments are mapped to genomic data and therefore a complete and up to date reference genome is needed to acquire precise results.

RNA-seq has been wildly adopted in many research areas including the understanding of disease etiology and characterization of host-microbial interaction, and host-pathogen interaction. Wang et al. (2016) identified 58 differentially expressed genes including 11 up-regulated and 47 down-regulated genes in super-shedders vs. non-shedders using RNA-seq, and further revealed that host innate and adaptive immune functions (e.g. Immigration of immune cells) were inhibited based on the functional annotation of down-regulated differential expressed genes (Wang et al., 2016). Casey et al. (2015) revealed altered signaling pathways in bovine macrophages including CD40 and IL-15 signaling in cattle infected by *Mycobacterium avium* subspecies *paratuberculosis* (Casey et al., 2015). Overall, these studies indicate that RNA-seq derived transcriptomic analysis informs our understanding of host-microbial interactions, and therefore can be used as a practical technique to study host-microbial interactions contributing to STEC colonization and super-shedding in cattle.

1.6.2.2 Amplicon seq for microbial community analysis

Conventional culture methods for identifying bacteria in human or animal samples enable the classification of bacteria at the species or strain level (Dowd et al., 2008). However, it has the disadvantages in that it can only identify bacteria that readily grow under laboratory conditions (Dowd et al., 2008) which represents only a small fraction of the total microbial populations (Rhoads et al., 2012).

Compared to culture methods, molecular techniques such as 16S rRNA gene amplicon seq, are sensitive and cost-effective (Deurenberg et al., 2017). For 16S rRNA gene amplicon seq, DNA is first extracted and a particular region of 16S rRNA gene is amplified, sequenced and the generated sequences are identified based on the similarity to a reference microbial genome. The 16S amplicon seq does not require that bacteria within the sample be culturable, and it can detect many different types of bacteria simultaneously (Poretsky et al., 2014). A cohort study collected fecal samples from children and compared the effectiveness of bacteria identification using 16S amplicon seq and culture methods (Gupta et al., 2019). The study revealed that culture methods identified a maximum of 8 bacterial species per sample while 16S amplicon seq identified up to 140 unique species per sample, making amplicon seq more suitable for characterizing microbial communities as compared to culture methods (Gupta et al., 2019).

Amplicon seq has been widely used in STEC research, *i.e.* a study collected fecal samples from beef cattle in a commercial feedlot and used 16S rRNA gene amplicon seq to identify that fecal samples from cattle that were positive for STEC and had lower bacterial diversity (Chopyk et al., 2016). Another study performed amplicon seq of fecal samples collected from 11 SS and 11 NS and revealed higher bacterial richness in SS compared to NS and 72 differentially abundant microbes in SS (Xu et al., 2014). These studies provide new insight into the variation in host microbial populations in response to STEC colonization of beef cattle. By adopting amplicon seq, differential abundant microbes can be determined which could be potential markers for STEC colonization.

1.6.3 Machine learning-based approaches to explore host-microbial interactions and keystone markers

Machine learning refers to the application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from previous experience without intensive programming (Rebala et al., 2019). It was first developed for pattern recognition in the 1960s and to date, the term machine learning summarizes two major objectives: classification of data based on models which have been developed and making predictions for future outcomes based on established models (Kotsiantis et al., 2006). Supervised machine learning approaches use labeled datasets and these datasets are designed to train algorithms for data classification or outcome prediction (Kotsiantis et al., 2006). While unsupervised machine learning is used for clustering unlabeled data sets and these algorithms can discover the hidden patterns in the datasets without human intervention (Lee et al., 2022). Unsupervised machine learning is used for clustering (that is grouping unlabeled data into different clusters based on their similarities or differences) and dimensionality reduction (that is reducing the number of input data to a manageable size while keeping the data integrity) (Glielmo et al., 2021).

Microbiome studies that have adopted machine learning approaches focus mainly on the classification and prediction of microbial taxa, predictions of host phenotypes by linking microbiome, and uncovering disease mechanisms through markers identified within microbial communities (Marcos-Zambrano et al., 2021). For instance, a previous study using the minimum redundancy-maximum relevance (mRMR) feature selection method identified a set of 20 microbial genes from the fecal microbiome that can be predictive of colorectal cancer (Yu

et al., 2017). This suggests that there maybe value in adopting machine learning approaches to investigate networks within microbiomes. Integrating machine learning methods into STEC studies is promising since it can effectively combine multi-omics data and host phenotypes in support of the possible identification of signature markers (*i.e.* host genetic markers, gut microbial markers) that could predict STEC colonization in beef cattle.

1.6.3.1 Supervised machine learning approaches for STEC research

1.6.3.1.1 Regression models

The regression model is a classic method for studying relationships between interested dependent and independent variables (Harrell et al., 1985). It can predict an outcome from a binary variable (that is Y), from one or more response categorical or continuous variable (that is X) (Harrell et al., 1985). Therefore, this approach can predict potential host and microbial factors from STEC abundance and prevalence. For example, a generalized linear regression model predicted that increased age was a potential risk factor that was positively related to *E.coli* O157 prevalence in feces on farm, while farm size and introduction of new beef cattle did not affect *E.coli* O157 prevalence (Cobbaut et al., 2009). Another linear regression model-based analysis revealed that multiple antimicrobial resistance (MAR) in bovine *E.coli* isolates was age-specific, being the highest in *E.coli* isolates from calves and progressively lower in isolates from adult cattle (Berge et al., 2010). The regression models are easy to understand and build on from known packages and can include all interested variables for prediction. However, it requires a high-quality dataset and is susceptible to colinear problems (that is there exists a strong linear correlation between independent variables).

1.6.3.1.2 Decision tree and random forest

The decision tree classifies a population into branch-like segments that construct an inverted tree with a root node, internal nodes (that denote a test on an attribute), and leaf nodes (that hold a class label) (Kingsford and Salzberg, 2008). The dataset can be split into subsets based on the chosen attribute to construct the decision tree and the Gini index is used to evaluate how accurate a split is among the classified groups (Kingsford and Salzberg, 2008). The random forest is developed from the decision tree, consisting of a large number of individual decision trees that are regarded as an ensemble (Couronné et al., 2018). Each individual tree in the random forest generates a class prediction and the class with the most votes becomes the model's prediction. Since the decision tree and random forest can handle large non-linear datasets efficiently and produce relatively robust predictions, it has been widely adopted for microbiome research. For instance, Han et al 2021 suggested that the survival of *E.coli* O157:H7 can be determined by soil salinity, pH, and bacterial community interactions based on a random forest analysis (Han et al., 2021). Bakshy et al. (2021) adopted decision tree and random forest models for selection of genetic markers for prediction of bovine tuberculosis (bTB) and revealed that 124 SNPs were predictable for bTB phenotyping identification. These studies suggested the feasibility of adopting random forest models for resolving biological questions, with the potential method to study factors associated with STEC colonization in beef cattle.

1.6.3.2 Unsupervised machine learning approaches

1.6.3.2.1 Principal component analysis

Principal component analysis (PCA) is a dimensionality reduction method that is used to reduce the dimensionality of a large dataset by transforming a large set of variables into smaller ones but still maintaining most of the information associated with the large set (David and Jacobs, 2013; Jolliffe and Cadima, 2016). Principal components (PCs) can be identified from eigenvectors and eigenvalues defined by the large set of an initial variable matrix. These PCs are uncorrelated and most of the information is squeezed into the first PC that represents the initial variables. This approach is not affected by variable collinearity and can reduce overfitting while being highly applicable for large datasets. Wang et al (2016) revealed a significant separation of host transcriptome between SS and NS in mucosa samples collected from the duodenum, jejunum, cecum, and colon using PCA analysis (Wang et al., 2016). This suggestd that applying PCA to uncover the substructure of genome data could be a useful tool for studying host transcriptome and amplicon sequencing data substructures.

1.6.3.2.2 Self-organizing map

The self-organizing map (SOM) is a dimensionality reduction method that reduces high dimension data to a lower dimensional while preserving the topological structure of the data (Kohonen, 1990). The SOM first trains an input dataset (that is the input space) to generate a lower-dimensional representation of the input data (that is the map space) and it then classifies the rest of the input data for mapping the generated map space (Kohonen, 1990). This neutral network-based algorithm, similar to PCA, plays a critical role in dimensionality reduction for

large datasets. While SOM is superior in terms of its ability to identify distinct patterns (PCA leads to overlapping of patterns), and it is more accurate for clustering large scale data compared to other dimensionality reduction methods and is prior assumption- free (PCA requires the normality of data) (Reusch et al., 2005; Astel et al., 2007). This approach can perform an unsupervised clustering process that the patterns are organized considering only their homology, without knowing the class to which they belong (Delgado et al., 2015). Therefore, it is suitable for clustering genomic data and for exploration of potential genetic markers. For instance, previous studied signature genes associated with immune responses using SOM in 3-month-old Tau22 (a pathological hallmark of Alzheimer's disease) mutated mice, confirmed the involvement of immunological processes in the onset of at Alzheimers (Ising et al., 2019). This approach is superior at sorting out genes with similar expression patterns that have a critical impact on host phenotypes. Considering the fact that STEC colonization in beef cattle triggers host gene and gut microbiome responses (Xu et al., 2014; Wang et al., 2021), certain genetic/microbial markers could in turn affect STEC colonization. Self-organizing maps could reveal genetic/microbial markers that affect host homeostasis. Therefore, the SOM could be the potential approach for identifying markers for host-microbiome interactions and to identify genetic/microbial markers that affect STEC O157 colonization in beef cattle.

1.6.4 Network-based approaches for understanding microbial community interactions and variations in responses to STEC O157 colonization

1.6.4.1 Microbial networks definition

Microbial taxa within the ecological niches could interact with each other and form 'micro-

communities' that affect microbial abundance and community structures (Faust and Raes, 2012). Interactions between microbial taxa could have either positive (win), negative (loss), or no impact (neutral) on microbial abundance and community structure (Lidicker, 1979; Faust and Raes, 2012). The potential outcomes of two or more interaction partners could result in various types of interaction (Figure 1.8). For instance, bacteria (of different taxonomic groups) may cooperate to build a biofilm, which confers antibiotic resistance to its members — a win–win relationship that is known as mutualism (Høiby et al., 2010). Predator-prey relationships and host-parasite relationships are classical loss-win interactions (Faust and Raes, 2012). Amensalism refers to if one partner is detrimental without any advantage to others, e.g., *lactobacilli* can lower the pH of the surrounding environment which can be harmful for other microorganisms (Faust and Raes, 2012). With commensalism one partner can be beneficial without helping or damaging the other, while the competition (loss-loss relationship) represents two species that occupy the same niche with the potential to exclude each other (Faust and Raes, 2012). Microbial interactions are vital for a successful establishment and maintenance of microbial populations in different environments. However, predicting microbial relationships can be complex since many ecological factors (*i.e.* physiochemical changes, metabolite exchanges, signaling, genetic exchanges resulting in genotype selections) can affect microbial interactions.

1.6.4.2 Microbial interaction predictions and methods for microbial network construction

Microbial interaction patterns can be predicted based on the microbial abundance under the premise that strongly non-random distribution patterns exist among microbial communities

(Faust and Raes, 2012). Predicting relationships from this principle is straightforward: when two species (or any taxonomically relevant units) co-occur or show a similar abundance pattern as, a positive relationship is assumed when they show mutual exclusion, anticorrelation, or a negative correlation (Faust and Raes, 2012). The prediction of microbial interactions from presence-absence or abundance data can be a problem known as network inference (Smet and Marchal, 2010). A few algorithms have been developed to predict microbial interaction networks such as similarity-based models, mainly using Pearson or Spearman correlation for abundance data. For example, Roehe et al. (2020) identified microbial taxa and interactions that could affect bovine methane emissions based on 63 rumen samples using Pearson correlation (Martínez-Álvaro et al., 2020). Another popular similarity-based network inference methodology is local similarity analysis (LSA), which can detect similarity between shifted abundance profiles and is therefore frequently used to build association networks from time series data (Ruan et al., 2006).

However, the selection of network construction methods can present complications. Relative abundance is usually used for correlations, while computing correlations between relative abundance can significantly distort results, a response known as the compositional bias which can yield artefactual correlations (Aitchison and Egozcue, 2005; Faust and Raes, 2012). In particular, these artefactual correlations can occur among non-correlated low abundance taxa. The distance-based method such as Bray-Curtis distance as well as normalization methods for compositional data could be applicable to assessing relative abundance-based microbial networks (Aitchison and Egozcue, 2005; Faust and Raes, 2012). ‘Rare biosphere’ termed as the

presence of a large percentage of zero abundant members within microbial taxa is also problematic because of its ambiguous interpretation as it is not possible to differentiate between absence vs below detectable levels (Sogin et al., 2006). This problem can be alleviated by data filtering or dedicated selections of distance measures.

What's more, the popular similarity-based methods (Pearson or Spearman correlation), and maximal information coefficient (MIC) have been proven to be less applicable for inferring microbial ecological networks as assessed by area under the precision-recall curves (AUPR) (Hirano and Takemoto, 2019). These metrics (*i.e.*, correlation-based approaches, MIC) are less applicable to microbial compositional data as the assumption of independent variables cannot be satisfied, leading to the generation of spurious correlations (Hirano and Takemoto, 2019), making it less suitable for inferring real microbial interactions. The random matrix theory-based approach could be adopted for inferring accurate microbial networks (Deng et al., 2012). This method is robust to noise and zero biosphere due to its nature of identifying transition points from Gaussian orthogonal ensemble (that is the non-random true relations) to Poisson distribution (false correlation) (Deng et al., 2012). The dedicated selection of network construction approaches and interpretation is warranted for uncovering microbial mechanisms for beef cattle microbiome research.

Furthermore, the structure of microbial networks also provides insight into the organization of microbial communities. For example, most microbial networks are scale-free, suggesting the presence of many taxa with only a few links and a few highly connected hub taxa (Chaffron et al., 2010; Faust and Raes, 2012). Networks can be partitioned into

clusters/modules, which are the densely interconnected nodes among networks. Clusters in the network have specific and different functions that enable microbes to respond to different environmental conditions (Chaffron et al., 2010; Jiao et al., 2016). For instance, microbial communities extracted from soil contaminated with oil identified three major clusters along with their diverse functions (electron-transfer, biogeochemical C- and N- cycles, organic contaminant degradation, respectively, Jiao et al., 2016). These clusters were shown to be crucial components of microbial communities in the network.

1.6.4.3 Microbial networks relate to host-microbial interactions

Understanding microbial interactions in host-associated communities is critical for the in-depth understanding of mechanisms of host health maintenance. For example, *Bacteroides* was the key genera identified from microbial interactions in the feces from autism spectrum disorder (ASD) children which can produce propionic acid and other short-chain fatty acids (Wan et al., 2021). The production of propionic acid showed increased restrictive behaviors and impaired social behavior in rat models (MacFabe et al., 2011). The other key taxa *Porphyromonas* identified from microbial interactions in ASD could induce cognitive dysfunction mediated by neuronal inflammation in mice models (MacFabe et al., 2011). Our previous studies indicated that rectal mucosa and fecal microbiota were remarkably different in terms of microbial diversity and composition (Xu et al., 2014; Wang et al., 2018). However, how microbial interactions can be driven by STEC colonization and therefore alter the microbial community structures and functions is unclear. Therefore, approaches towards the construction of microbial networks and selection of key taxa driven network differences can open the way towards the

understanding of STEC O157 colonization mechanisms from a microbial perspective. In the long run, this may enable microbial communities to be engineered in manner that mitigates STEC O157 in the digestive tract of cattle.

1.7 Knowledge Gaps, hypothesis, and objectives

On-farm STEC detection is one of the most critical steps during pre-harvest interventions, as it would further ensure that raw beef products safely enter the processing plant and are suitable to enter the food production chain for consumption by humans. However, current identification methods largely rely on bacterial culture, which is time-consuming and can only detect the O157 strain. It is known that expressions of *stx* combined with other virulence factors promote STEC colonization in cattle. Also, the abundance of *stx1* and *stx2* genes in cattle is important as they can be used as indicators of potential STEC. However, information on the abundance and expression of *stx1* and *stx2* genes *in vivo* (e.g. in RAJ) of feedlot cattle is lacking. Although immune responses in cattle have been proven to be altered in cattle colonized with STEC O157, the relationship between host immune genes and Shiga toxin gene remains unexplored and the potential of host immune genes to be used as markers of STEC colonization is unclear. The gut microbial community could be affected by STEC colonization, while how fecal microbial community structures and interactions vary in response to stxs gene expression is unknown.

I first hypothesize that the expression and abundance of *stx* genes at the RAJ are influenced by cattle breed and the expression of host immune genes. Secondly, host immune gene expression can be utilized as potential genetic markers for the identification of *stx*

expression. Furthermore, the *stx* gene expressions in STEC could drive the variations of fecal and mucosa microbial community structures and interactions with keystone microbes being identified as markers for *stx* gene expression.

Since microbial communities were previously reported to be related to STEC O157 colonization, whether the mucosa microbial community assembly and interactions could affect O157 strain-specific challenge in calves is also unknown. Previous studies identified differed host immune responses in calves challenged with STEC O157 PT 21/28 (producing functional stx2a) in comparison to calves following RE21/28 (producing functional stx2a and stx2c) challenge. I further hypothesized that different magnitudes of host immune responses will be identified upon strain-specific STEC O157 colonization. Secondly, mucosa microbial community structures, assembly, and interactions will be significantly different in terms of specific STEC O157 challenges in calves. At last, varied host-microbial interactions driven by strain-specific STEC O157 colonization will be identified, and signature host genetic/microbial markers will be identified to assess expression of differed stx2 subtypes.

The specific objectives are: (1) to investigate factors affecting the abundance and expression of *stx* at the RAJ from feedlot cattle and their associations with the expression of host immune genes previously reported to be altered in SS using machine learning-derived models (Chapter 2); and (2) to investigate variations of fecal microbial community structures and interactions in response to *stx* gene expression in STEC at the RAJ from feedlot cattle and identification of keystone microbes that could be utilized as potential microbial markers for *stx* gene expression (Chapter 3); (3) to investigate the transcriptome of RAJ and to identify

differentially expressed genes and their functions in calves challenged with different strains of STEC O157 (Chapter 4); (4) to investigate mucosa microbial community variations in response to the O157 strain-specific challenge (Chapter 5), and (5) to investigate variation in host-microbial interactions in calves challenged with STEC O157 differed in terms of the production of stx2a and to preliminary identify genes/microbes among host-microbial interactions that can be employed as genetic/microbes markers to differentiate strain-dependent STEC O157 colonization in cattle (Chapter 5).

The long-term goal of this project is to further understand host-STEC interactions in cattle from a perspective of host genomics and gut microbiome as well as to gain insights into the development of strategies that can rapidly identify potential SS and differentiate STEC O157 strains based on differences in gene expressions and gut microbes.

1.8 References

Ahlqvist, E., Storm, P., Käräjämäki, A., Martinell, M., Dorkhan, M., Carlsson, A., et al. (2018). Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. *Lancet Diabetes Endocrinol* 6, 361–369. doi: 10.1016/s2213-8587(18)30051-2.

Ahmed, N., Dobrindt, U., Hacker, J., and Hasnain, S. E. (2008). Genomic fluidity andpathogenic bacteria: applications in diagnostics, epidemiology and intervention. Nat. Rev. Microbiol. 6, 387–394. doi: 10.1038/nrmicro1889.

Aitchison, J., and Egozcue, J. J. (2005). Compositional Data Analysis: Where Are We and Where Should We Be Heading? *Math Geol* 37, 829–850. doi: 10.1007/s11004-005-7383-7.

Alharbi, M. G., Al-Hindi, R. R., Esmael, A., Alotibi, I. A., Azhari, S. A., Alseghayer, M. S., et al. (2022). The "Big Six": Hidden Emerging Foodborne Bacterial Pathogens. *Tropical Medicine Infect Dis* 7, 356. doi: 10.3390/tropicalmed7110356.

Andreoli, S. P., Trachtman, H., Acheson, D. W. K., Siegler, R. L., and Obrig, T. G. (2002). Hemolytic uremic syndrome: epidemiology, pathophysiology, and therapy. *Pediatr Nephrol* 17, 293–298. doi: 10.1007/s00467-001-0783-0.

Antão, E.-M., Wieler, L. H., and Ewers, C. (2009). Adhesive threads of extraintestinal pathogenic Escherichia coli. *Gut Pathog.* 1, 22. doi: 10.1186/1757-4749-1-22.

Astel, A., Tsakovski, S., Barbieri, P., and Simeonov, V. (2007). Comparison of self-organizing maps classification approach with cluster and principal components analysis for large environmental data sets. *Water Res* 41, 4566–4578. doi: 10.1016/j.watres.2007.06.030.

Baines, D., Lee, B., and McAllister, T. (2008). Heterogeneity in enterohemorrhagic Escherichia coli O157:H7 fecal shedding in cattle is related to Escherichia coli O157:H7 colonization of the small and large intestine. *Can J Microbiol* 54, 984–995. doi: 10.1139/w08-090.

Bakshy, K., Heimeier, D., Schwartz, J. C., Glass, E. J., Wilkinson, S., Skuce, R. A., et al. (2021). Development of polymorphic markers in the immune gene complex loci of cattle. *J Dairy Sci*. doi: 10.3168/jds.2020-19809.

Berge, A. C., Hancock, D. D., Sischo, W. M., and Besser, T. E. (2010). Geographic, farm, and animal factors associated with multiple antimicrobial resistance in fecal Escherichia coli isolates from cattle in the western United States. *J Am Vet Med Assoc* 236, 1338–1344. doi: 10.2460/javma.236.12.1338.

Bliska, J. B., Wang, X., Viboud, G. I., and Brodsky, I. E. (2013). Modulation of innate immune responses by Yersinia type III secretion system translocators and effectors. *Cell Microbiol* 15, 1622–1631. doi: 10.1111/cmi.12164.

Blount, Z. D. (2015). The unexhausted potential of E. coli. *Elife* 4, e05826. doi: 10.7554/elife.05826.

Bono, J. L., Keen, J. E., Miller, L. C., Fox, J. M., Chitko-McKown, C. G., Heaton, M. P., et al. (2004). Evaluation of a Real-Time PCR Kit for Detecting Escherichia coli O157 in Bovine Fecal Samples. *Appl Environ Microb* 70, 1855–1857. doi: 10.1128/aem.70.3.1855-1857.2004.

Brandelli, J. R., Griener, T. P., Laing, A., Mulvey, G., and Armstrong, G. D. (2015). The Effects of Shiga Toxin 1, 2 and Their Subunits on Cytokine and Chemokine Expression by Human Macrophage-Like THP-1 Cells. *Toxins* 7, 4054–4066. doi: 10.3390/toxins7104054.

Bretschneider, G., Berberov, E. M., and Moxley, R. A. (2007). Reduced intestinal colonization of adult beef cattle by Escherichia coli O157:H7 tir deletion and nalidixic-acid-resistant mutants lacking flagellar expression. *Vet Microbiol* 125, 381–386. doi: 10.1016/j.vetmic.2007.06.009.

Browne, A. S., Midwinter, A. C., Withers, H., Cookson, A. L., Biggs, P. J., Marshall, J. C., et al. (2021). Transmission Dynamics of Shiga Toxin-Producing Escherichia coli in New Zealand Cattle from Farm to Slaughter. *Appl Environ Microb* 87, e02907-20. doi: 10.1128/aem.02907-20.

Buffie, C. G., and Pamer, E. G. (2013). Microbiota-mediated colonization resistance against intestinal pathogens. *Nat Rev Immunol* 13, 790–801. doi: 10.1038/nri3535.

Bulcke, M. V. den, Lievens, A., Barbau-Piednoir, E., MbongoloMbella, G., Roosens, N., Sneyers, M., et al. (2010). A theoretical introduction to “Combinatory SYBR®Green qPCR Screening”, a matrix-based approach for the detection of materials derived from genetically modified plants. *Anal Bioanal Chem* 396, 2113–2123. doi: 10.1007/s00216-009-3286-7.

Bürk, C., Dietrich, R., Açar, G., Moravek, M., Bülte, M., and Märtlbauer, E. (2003). Identification and Characterization of a New Variant of Shiga Toxin 1 in Escherichia coli ONT:H19 of Bovine Origin. *J Clin Microbiol* 41, 2106–2112. doi: 10.1128/jcm.41.5.2106-2112.2003.

Bustin, S. A. (2005). Real-time, fluorescence-based quantitative PCR: a snapshot of current procedures and preferences. *Expert Rev Mol Diagn* 5, 493–498. doi: 10.1586/14737159.5.4.493.

Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., et al. (2009). The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. *Clin Chem* 55, 611–622. doi: 10.1373/clinchem.2008.112797.

Casey, M. E., Meade, K. G., Nalpas, N. C., Taraktsoglou, M., Browne, J. A., Killick, K. E., et al. (2015). Analysis of the Bovine Monocyte-Derived Macrophage Response to Mycobacterium avium Subspecies Paratuberculosis Infection Using RNA-seq. *Front Immunol* 6, 23. doi: 10.3389/fimmu.2015.00023.

Castro, V. S., Carvalho, R. C. T., Conte-Junior, C. A., and Figuiredo, E. E. S. (2017). Shiga-toxin Producing Escherichia coli: Pathogenicity, Supershedding, Diagnostic Methods, Occurrence, and Foodborne Outbreaks. *Compr Rev Food Sci F* 16, 1269–1280. doi: 10.1111/1541-4337.12302.

Chaffron, S., Rehrauer, H., Pernthaler, J., and Mering, C. von (2010). A global network of coexisting microbes from environmental and whole-genome sequence data. *Genome Res* 20, 947–959. doi: 10.1101/gr.104521.109.

Chapman, P., Siddons, C., Wright, D., et al. (1993). Cattle as a possible source of verocytotoxin-producing Escherichia coli O157 infections in man | Epidemiology & Infection | Cambridge Core. *Cattle as a possible source of verocytotoxin-producing Escherichia coli O157 infections in man.* Available at: https://www.cambridge.org/core/journals/epidemiology-and-infection/article/cattle-as-a-possible-source-of-verocytotoxinproducing-escherichia-coli-o157-infections-in-man/E350CA666395D1B61BEA2A68E9419BE1 [Accessed March 20, 2021].

Chase-Topping, M., Gally, D., Low, C., Matthews, L., and Woolhouse, M. (2008). Super-shedding and the link between human infection and livestock carriage of Escherichia coli O157. *Nat Rev Microbiol* 6, 904–912. doi: 10.1038/nrmicro2029.

Chave, J. (2004). Neutral theory and community ecology. *Ecol Lett* 7, 241–253. doi: 10.1111/j.1461-0248.2003.00566.x.

Chopyk, J., Moore, R. M., DiSpirito, Z., Stromberg, Z. R., Lewis, G. L., Renter, D. G., et al. (2016). Presence of pathogenic Escherichia coli is correlated with bacterial community diversity and composition on pre-harvest cattle hides. *Microbiome* 4, 9. doi: 10.1186/s40168-016-0155-4.

Cimolai, N., Basalyga, S., Mah, D. G., Morrison, B. J., and Carter, J. E. (1994). A continuing assessment of risk factors for the development of Escherichia coli O157:H7-associated hemolytic uremic syndrome. *Clin Nephrol* 42, 85–9.

Cobbaut, K., Berkvens, D., Houf, K., Deken, R. D., and Zutter, L. D. (2009). Escherichia coli O157 Prevalence in Different Cattle Farm Types and Identification of Potential Risk Factors. *J Food Protect* 72, 1848–1853. doi: 10.4315/0362-028x-72.9.1848.

Corbishley, A., Ahmad, N. I., Hughes, K., Hutchings, M. R., McAteer, S. P., Connelley, T. K., et al. (2014). Strain-Dependent Cellular Immune Responses in Cattle following Escherichia coli O157:H7 Colonization. *Infect Immun* 82, 5117–5131. doi: 10.1128/iai.02462-14.

Couronné, R., Probst, P., and Boulesteix, A.-L. (2018). Random forest versus logistic regression: a large-scale benchmark experiment. *Bmc Bioinformatics* 19, 270. doi: 10.1186/s12859-018-2264-5.

Croxen, M. A., Law, R. J., Scholz, R., Keeney, K. M., Wlodarska, M., and Finlay, B. B. (2013). Recent Advances in Understanding Enteric Pathogenic Escherichia coli. *Clin Microbiol Rev* 26, 822–880. doi: 10.1128/cmr.00022-13.

David, C. C., and Jacobs, D. J. (2013). Protein Dynamics, Methods and Protocols. *Methods Mol Biology* 1084, 193–226. doi: 10.1007/978-1-62703-658-0_11.

Dean-Nystrom, E. A., Bosworth, B. T., Cray, W. C., and Moon, H. W. (1997). Pathogenicity of Escherichia coli O157:H7 in the intestines of neonatal calves. *Infect Immun* 65, 1842–1848. doi: 10.1128/iai.65.5.1842-1848.1997.

DeGrandis, S., Law, H., Brunton, J., Gyles, C., and Lingwood, C. A. (1989). Globotetraosylceramide Is Recognized by the Pig Edema Disease Toxin. *J Biol Chem* 264, 12520–12525. doi: 10.1016/s0021-9258(18)63888-8.

Delgado, S., Morán, F., Mora, A., Merelo, J. J., and Briones, C. (2015). A novel representation of genomic sequences for taxonomic clustering and visualization by means of self-organizing maps. *Bioinformatics* 31, 736–744. doi: 10.1093/bioinformatics/btu708.

Deng, W., Marshall, N. C., Rowland, J. L., McCoy, J. M., Worrall, L. J., Santos, A. S., et al. (2017). Assembly, structure, function and regulation of type III secretion systems. *Nat Rev Microbiol* 15, 323–337. doi: 10.1038/nrmicro.2017.20.

Deng, Y., Jiang, Y.-H., Yang, Y., He, Z., Luo, F., and Zhou, J. (2012). Molecular ecological network analyses. *Bmc Bioinformatics* 13, 113. doi: 10.1186/1471-2105-13-113.

Deurenberg, R. H., Bathoorn, E., Chlebowicz, M. A., Couto, N., Ferdous, M., García-Cobos, S., et al. (2017). Application of next generation sequencing in clinical microbiology and infection prevention. *J Biotechnol* 243, 16–24. doi: 10.1016/j.jbiotec.2016.12.022.

Dowd, S. E., Sun, Y., Secor, P. R., Rhoads, D. D., Wolcott, B. M., James, G. A., et al. (2008). Survey of bacterial diversity in chronic wounds using Pyrosequencing, DGGE, and full ribosome shotgun sequencing. *Bmc Microbiol* 8, 43. doi: 10.1186/1471-2180-8-43.

Dundas, S., Todd, W. T. A., Stewart, A. I., Murdoch, P. S., Chaudhuri, A. K. R., and Hutchinson, S. J. (2001). The Central Scotland Escherichia coli O157:H7 Outbreak: Risk Factors for the Hemolytic Uremic Syndrome and Death among Hospitalized Patients. *Clin Infect Dis* 33, 923–931. doi: 10.1086/322598.

Eckburg, P. B., Bik, E. M., Bernstein, C. N., Purdom, E., Dethlefsen, L., Sargent, M., et al. (2005). Diversity of the Human Intestinal Microbial Flora. *Science* 308, 1635–1638. doi: 10.1126/science.1110591.

Eeckhaut, V., Machiels, K., Perrier, C., Romero, C., Maes, S., Flahou, B., et al. (2013). Butyricicoccus pullicaecorum in inflammatory bowel disease. *Gut* 62, 1745. doi: 10.1136/gutjnl-2012-303611.

EKIRI, A. B., LANDBLOM, D., DOETKOTT, D., OLET, S., SHELVER, W. L., and KHAITSA, M. L. (2016). Isolation and Characterization of Shiga Toxin–Producing Escherichia coli Serogroups O26, O45, O103, O111, O113, O121, O145, and O157 Shed from Range and Feedlot Cattle from Postweaning to Slaughter. *J Food Protect* 77, 1052–1061. doi: 10.4315/0362-028x.jfp-13-373.

Etcheverría, A. I., and Padola, N. L. (2013). Shiga toxin-producing Escherichia coli. *Virulence* 4, 366–372. doi: 10.4161/viru.24642.

Fairbrother, J. M., and Nadeau, E. (2006). Escherichia coli: on-farm contamination of animals. *Revue Sci Et Technique Int Office Epizootics* 25, 555–69.

Fan, P., Kim, M., Liu, G., Zhai, Y., Liu, T., Driver, J. D., et al. (2021a). The Gut Microbiota of Newborn Calves and Influence of Potential Probiotics on Reducing Diarrheic Disease by Inhibition of Pathogen Colonization. *Front Microbiol* 12, 772863. doi: 10.3389/fmicb.2021.772863.

Fan, P., Nelson, C. D., Driver, J. D., Elzo, M. A., Peñagaricano, F., and Jeong, K. C. (2021b). Host genetics exerts lifelong effects upon hindgut microbiota and its association with bovine growth and immunity. *Isme J* 15, 2306–2321. doi: 10.1038/s41396-021-00925-x.

Faust, K., and Raes, J. (2012). Microbial interactions: from networks to models. *Nat Rev Microbiol* 10, 538–550. doi: 10.1038/nrmicro2832.

Fitzgerald, S. F., Beckett, A. E., Palarea-Albaladejo, J., McAteer, S., Shaaban, S., Morgan, J., et al. (2019). Shiga toxin sub-type 2a increases the efficiency of Escherichia coli O157 transmission between animals and restricts epithelial regeneration in bovine enteroids. *Plos Pathog* 15, e1008003. doi: 10.1371/journal.ppat.1008003.

Frankel, G., Phillips, A. D., Rosenshine, I., Dougan, G., Kaper, J. B., and Knutton, S. (1998). Enteropathogenic and enterohaemorrhagic Escherichia coli : more subversive elements. *Mol Microbiol* 30, 911–921. doi: 10.1046/j.1365-2958.1998.01144.x.

Fraser, M. E., Fujinaga, M., Cherney, M. M., Melton-Celsa, A. R., Twiddy, E. M., O'Brien, A. D., et al. (2004). Structure of Shiga Toxin Type 2 (Stx2) from Escherichia coli O157:H7*. *J Biol Chem* 279, 27511–27517. doi: 10.1074/jbc.m401939200.

Furman, O., Shenhav, L., Sasson, G., Kokou, F., Honig, H., Jacoby, S., et al. (2020). Stochasticity constrained by deterministic effects of diet and age drive rumen microbiome assembly dynamics. *Nat Commun* 11, 1904. doi: 10.1038/s41467-020-15652-8.

Galanis, E., Longmore, K., Hasselback, P., Swann, D., Ellis, A., and Panaro, L. (2003). Investigation of an E. coli O157:H7 outbreak in Brooks, Alberta, June-July 2002: the role of occult cases in the spread of infection within a daycare setting. *Can Commun Dis Rep Relevé Des Maladies Transm Au Can* 29, 21–8.

Gaytán, M. O., Martínez-Santos, V. I., Soto, E., and González-Pedrajo, B. (2016). Type Three Secretion System in Attaching and Effacing Pathogens. *Front Cell Infect Mi* 6, 129. doi: 10.3389/fcimb.2016.00129.

Glielmo, A., Husic, B. E., Rodriguez, A., Clementi, C., Noé, F., and Laio, A. (2021). Unsupervised Learning Methods for Molecular Simulation Data. *Chem Rev* 121, 9722–9758. doi: 10.1021/acs.chemrev.0c01195.

Gomes, T. A. T., Elias, W. P., Scaletsky, I. C. A., Guth, B. E. C., Rodrigues, J. F., Piazza, R. M. F., et al. (2016). Diarrheagenic Escherichia coli. *Braz J Microbiol* 47, 3–30. doi: 10.1016/j.bjm.2016.10.015.

Gomez, D. E., Arroyo, L. G., Costa, M. C., Viel, L., and Weese, J. S. (2017). Characterization of the Fecal Bacterial Microbiota of Healthy and Diarrheic Dairy Calves. *J Vet Intern Med* 31, 928–939. doi: 10.1111/jvim.14695.

Griffin, P. M., and Tauxe, R. V. (1991). 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. doi: 10.1093/oxfordjournals.epirev.a036079.

Group, O. W. (2018). Monitoring the incidence and causes of diseases potentially transmitted by food in Australia: Annual report of the OzFoodNet network, 2012. *Commun Dis Intell 2018* 42.

Gupta, S., Mortensen, M. S., Schjørring, S., Trivedi, U., Vestergaard, G., Stokholm, J., et al. (2019). Amplicon sequencing provides more accurate microbiome information in healthy children compared to culturing. *Commun Biology* 2, 291. doi: 10.1038/s42003-019-0540-1.

Hale, T. L., and Formal, S. B. (1980). Cytotoxicity of Shigella dysenteriae 1 for cultured mammalian cells. *Am J Clin Nutrition* 33, 2485–2490. doi: 10.1093/ajcn/33.11.2485.

Hall, G., Kurosawa, S., and Stearns-Kurosawa, D. J. (2017). Shiga Toxin Therapeutics: Beyond Neutralization. *Toxins* 9, 291. doi: 10.3390/toxins9090291.

Halliday, J. E., Chase-Topping, M. E., Pearce, M. C., McKendrick, I. J., Allison, L., Fenlon, D., et al. (2006). Herd-level risk factors associated with the presence of Phage type 21/28 E. coli O157 on Scottish cattle farms. *Bmc Microbiol* 6, 99. doi: 10.1186/1471-2180-6-99.

Han, Z., Ma, J., Yang, C.-H., and Ibekwe, A. M. (2021). Soil salinity, pH, and indigenous bacterial community interactively influence the survival of E. coli O157:H7 revealed by multivariate statistics. *Environ Sci Pollut R* 28, 5575–5586. doi: 10.1007/s11356-020-10942-6.

Hanson, C. A., Fuhrman, J. A., Horner-Devine, M. C., and Martiny, J. B. H. (2012). Beyond biogeographic patterns: processes shaping the microbial landscape. *Nat Rev Microbiol* 10, 497–506. doi: 10.1038/nrmicro2795.

Harrell, F. E., Lee, K. L., Matchar, D. B., and Reichert, T. A. (1985). Regression models for prognostic prediction: advantages, problems, and suggested solutions. *Cancer Treat Rep* 69, 1071–77.

Hartl, D. L., and Dykhuizen, D. E. (1984). The Population Genetics of Escherichia Coli. *Annu Rev Genet* 18, 31–68. doi: 10.1146/annurev.ge.18.120184.000335.

Heiman, K. E., Mody, R. K., Johnson, S. D., Griffin, P. M., and Gould, L. H. (2015). Escherichia coli O157 Outbreaks in the United States, 2003–2012. *Emerg Infect Dis* 21, 1293–1301. doi: 10.3201/eid2108.141364.

Herbert, L. J., Vali, L., Hoyle, D. V., Innocent, G., McKendrick, I. J., Pearce, M. C., et al. (2014). E. coli O157 on Scottish cattle farms: Evidence of local spread and persistence using repeat cross-sectional data. *Bmc Vet Res* 10, 95–95. doi: 10.1186/1746-6148-10-95.

Hirano, H., and Takemoto, K. (2019). Difficulty in inferring microbial community structure based on co-occurrence network approaches. *Bmc Bioinformatics* 20, 329. doi: 10.1186/s12859-019-2915-1.

Høiby, N., Bjarnsholt, T., Givskov, M., Molin, S., and Ciofu, O. (2010). Antibiotic resistance of bacterial biofilms. *Int J Antimicrob Ag* 35, 322–332. doi: 10.1016/j.ijantimicag.2009.12.011.

Hsu, B. B., Way, J. C., and Silver, P. A. (2019). Stable neutralization of virulent bacteria using temperate phage in the mammalian gut. *Biorxiv*, 794222. doi: 10.1101/794222.

Hubbell, S. P., and Borda-de-Água, L. (2004). THE UNIFIED NEUTRAL THEORY OF BIODIVERSITY AND BIOGEOGRAPHY: REPLY. *Ecology* 85, 3175–3178. doi: 10.1890/04-0808.

Ising, C., Venegas, C., Zhang, S., Scheiblich, H., Schmidt, S. V., Vieira-Saecker, A., et al. (2019). NLRP3 inflammasome activation drives tau pathology. *Nature* 575, 669–673. doi: 10.1038/s41586-019-1769-z.

Jackson, M. P., Neill, R. J., O'Brien, A. D., Holmes, R. K., and Newland, J. W. (1987). Nucleotide sequence analysis and comparison of the structural genes for Shiga-like toxin I and

Shiga-like toxin II encoded by bacteriophages from Escherichia coli 933. *Fems Microbiol Lett* 44, 109–114. doi: 10.1111/j.1574-6968.1987.tb02252.x.

Ji, H., and Dong, H. (2015). Key steps in type III secretion system (T3SS) towards translocon assembly with potential sensor at plant plasma membrane. *Mol Plant Pathol* 16, 762–773. doi: 10.1111/mpp.12223.

Jiao, S., Liu, Z., Lin, Y., Yang, J., Chen, W., and Wei, G. (2016). Bacterial communities in oil contaminated soils: Biogeography and co-occurrence patterns. *Soil Biology Biochem* 98, 64–73. doi: 10.1016/j.soilbio.2016.04.005.

Jolliffe, I. T., and Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical Transactions Royal Soc Math Phys Eng Sci* 374, 20150202. doi: 10.1098/rsta.2015.0202.

Kaiko, G. E., and Stappenbeck, T. S. (2014). Host–microbe interactions shaping the gastrointestinal environment. *Trends Immunol* 35, 538–548. doi: 10.1016/j.it.2014.08.002.

Kaper, J. B., Nataro, J. P., and Mobley, H. L. T. (2004). Pathogenic Escherichia coli. *Nat Rev Microbiol* 2, 123–140. doi: 10.1038/nrmicro818.

Karmali, MohamedA., Petric, M., Steele, BrianT., and Lim, C. (1983). SPORADIC CASES OF HAEMOLYTIC-URAEMIC SYNDROME ASSOCIATED WITH FAECAL CYTOTOXIN AND CYTOTOXIN-PRODUCING ESCHERICHIA COLI IN STOOLS. *Lancet* 321, 619–620. doi: 10.1016/s0140-6736(83)91795-6.

Kate, T., M., E, M., Shannon, N, S., Paul, Aamir, F., Frank, P., Kathryn, D., et al. (2006). Estimated Numbers of Community Cases of Illness Due to Salmonella, Campylobacter and Verotoxigenic Escherichia Coli: Pathogen-Specific Community Rates. *Can J Infect Dis Medical Microbiol* 17, 229–234. doi: 10.1155/2006/806874.

Kingsford, C., and Salzberg, S. L. (2008). What are decision trees? *Nat Biotechnol* 26, 1011–1013. doi: 10.1038/nbt0908-1011.

Koch, C., Hertwig, S., Lurz, R., Appel, B., and Beutin, L. (2001). Isolation of a Lysogenic Bacteriophage Carrying thestx 1OX3 Gene, Which Is Closely Associated with Shiga Toxin-Producing Escherichia coli Strains from Sheep and Humans. *J Clin Microbiol* 39, 3992–3998. doi: 10.1128/jcm.39.11.3992-3998.2001.

Kohonen, T. (1990). The self-organizing map. *P Ieee* 78, 1464–1480. doi: 10.1109/5.58325.

Kotsiantis, S. B., Zaharakis, I. D., and Pintelas, P. E. (2006). Machine learning: a review of classification and combining techniques. *Artif Intell Rev* 26, 159–190. doi: 10.1007/s10462-007-9052-3.

Krause, M., Barth, H., and Schmidt, H. (2018). Toxins of Locus of Enterocyte Effacement-Negative Shiga Toxin-Producing Escherichia coli. *Toxins* 10, 241. doi: 10.3390/toxins10060241.

Lee, J., Warner, E., Shaikhouni, S., Bitzer, M., Kretzler, M., Gipson, D., et al. (2022). Unsupervised machine learning for identifying important visual features through bag-of-words

using histopathology data from chronic kidney disease. *Sci Rep-uk* 12, 4832. doi: 10.1038/s41598-022-08974-8.

Lee, M.-S., and Tesh, V. L. (2019). Roles of Shiga Toxins in Immunopathology. *Toxins* 11, 212. doi: 10.3390/toxins11040212.

Legros, N., Pohlentz, G., Steil, D., and Müthing, J. (2018). Shiga toxin-glycosphingolipid interaction: status quo of research with focus on primary human brain and kidney endothelial cells. *Int J Med Microbiol* 308, 1073–1084. doi: 10.1016/j.ijmm.2018.09.003.

Leibold, M. A. (1995). The Niche Concept Revisited: Mechanistic Models and Community Context. *Ecology* 76, 1371–1382. doi: 10.2307/1938141.

Leibold, M. A., and McPeek, M. A. (2006). COEXISTENCE OF THE NICHE AND NEUTRAL PERSPECTIVES IN COMMUNITY ECOLOGY. *Ecology* 87, 1399–1410. doi: 10.1890/0012-9658(2006)87[1399:cotnan]2.0.co;2.

Leimbach, A., Hacker, J., and Dobrindt, U. (2013). Between Pathogenicity and Commensalism. *Curr Top Microbiol* 358, 3–32. doi: 10.1007/82_2012_303.

Leung, P. H. M., Peiris, J. S. M., Ng, W. W. S., Robins-Browne, R. M., Bettelheim, K. A., and Yam, W. C. (2003). A Newly Discovered Verotoxin Variant, VT2g, Produced by Bovine Verocytotoxigenic Escherichia coli. *Appl Environ Microb* 69, 7549–7553. doi: 10.1128/aem.69.12.7549-7553.2003.

Li, P., Luo, H., Ji, B., and Nielsen, J. (2022). Machine learning for data integration in human gut microbiome. *Microb Cell Fact* 21, 241. doi: 10.1186/s12934-022-01973-4.

Lidicker, W. Z. (1979). A Clarification of Interactions in Ecological Systems. *Bioscience* 29, 475–477. doi: 10.2307/1307540.

Lim, J. Y., Yoon, J., and Hovde, C. J. (2010). A brief overview of Escherichia coli O157:H7 and its plasmid O157. *J Microbiol Biotechn* 20, 5–14.

Lisboa, Szelewicki, Lin, Latonas, Li, Zhi, et al. (2019). Epidemiology of Shiga Toxin-Producing Escherichia coli O157 in the Province of Alberta, Canada, 2009–2016. *Toxins* 11, 613. doi: 10.3390/toxins11100613.

Liu, J., Meng, Z., Liu, X., and Zhang, X.-H. (2019). Microbial assembly, interaction, functioning, activity and diversification: a review derived from community compositional data. *Mar Life Sci Technology* 1, 112–128. doi: 10.1007/s42995-019-00004-3.

Loef, B., Wong, A., Janssen, N. A. H., Strak, M., Hoekstra, J., Picavet, H. S. J., et al. (2022). Using random forest to identify longitudinal predictors of health in a 30-year cohort study. *Sci Rep-uk* 12, 10372. doi: 10.1038/s41598-022-14632-w.

Łoś, M., Kuzio, J., McConnell, M. R., Kropinski, A. M., Węgrzyn, G., and Christie, G. E. (2020). Bacteriophages in the Control of Food- and Waterborne Pathogens. 157–198. doi: 10.1128/9781555816629.ch9.

Ludwig, K., Karmali, M. A., Sarkim, V., Bobrowski, C., Petric, M., Karch, H., et al. (2001). Antibody Response to Shiga Toxins Stx2 and Stx1 in Children with Enteropathic Hemolytic-Uremic Syndrome. *J Clin Microbiol* 39, 2272–2279. doi: 10.1128/jcm.39.6.2272-2279.2001.

Lynn, R. M., O'Brien, S. J., Taylor, C. M., Adak, G. K., Chart, H., Cheasty, T., et al. (2005). Childhood Hemolytic Uremic Syndrome, United Kingdom and Ireland - Volume 11, Number 4—April 2005 - Emerging Infectious Diseases journal - CDC. *Emerg Infect Dis* 11, 590–596. doi: 10.3201/eid1104.040833.

MacFabe, D. F., Cain, N. E., Boon, F., Ossenkopp, K.-P., and Cain, D. P. (2011). Effects of the enteric bacterial metabolic product propionic acid on object-directed behavior, social behavior, cognition, and neuroinflammation in adolescent rats: Relevance to autism spectrum disorder. *Behav Brain Res* 217, 47–54. doi: 10.1016/j.bbr.2010.10.005.

Magee, D. A., Taraktsoglou, M., Killick, K. E., Nalpas, N. C., Browne, J. A., Park, S. D. E., et al. (2012). Global Gene Expression and Systems Biology Analysis of Bovine Monocyte-Derived Macrophages in Response to In Vitro Challenge with Mycobacterium bovis. *Plos One* 7, e32034. doi: 10.1371/journal.pone.0032034.

Mainil, J. G., Jacquemin, E. R., Kaeckenbeeck, A. E., and Pohl, P. H. (1993). Association between the effacing (eae) gene and the Shiga-like toxin-encoding genes in Escherichia coli isolates from cattle. *Am J Vet Res* 54, 1064–8.

Maluta, R. P., Fairbrother, J. M., Stella, A. E., Rigobelo, E. C., Martinez, R., and Ávila, F. A. de (2014). Potentially pathogenic Escherichia coli in healthy, pasture-raised sheep on farms and at the abattoir in Brazil. *Vet Microbiol* 169, 89–95. doi: 10.1016/j.vetmic.2013.12.013.

Mao, S., Zhang, M., Liu, J., and Zhu, W. (2015). Characterising the bacterial microbiota across the gastrointestinal tracts of dairy cattle: membership and potential function. *Sci Rep-uk* 5, 16116. doi: 10.1038/srep16116.

Marcos-Zambrano, L. J., Karaduzovic-Hadziabdic, K., Turukalo, T. L., Przymus, P., Trajkovik, V., Aasmets, O., et al. (2021). Applications of Machine Learning in Human Microbiome Studies: A Review on Feature Selection, Biomarker Identification, Disease Prediction and Treatment. *Front Microbiol* 12, 634511. doi: 10.3389/fmicb.2021.634511.

Martínez-Álvaro, M., Auffret, M. D., Stewart, R. D., Dewhurst, R. J., Duthie, C.-A., Rooke, J. A., et al. (2020). Identification of Complex Rumen Microbiome Interaction Within Diverse Functional Niches as Mechanisms Affecting the Variation of Methane Emissions in Bovine. *Front Microbiol* 11, 659. doi: 10.3389/fmicb.2020.00659.

McPherson, M., Kirk, M. D., Raupach, J., Combs, B., and Butler, J. R. G. (2011). Economic Costs of Shiga Toxin–Producing Escherichia coli Infection in Australia. *Foodborne Pathog Dis* 8, 55–62. doi: 10.1089/fpd.2010.0608.

Melton-Celsa, A. R. (2014a). Shiga Toxin (Stx) Classification, Structure, and Function. *Microbiol Spectr* 2. doi: 10.1128/microbiolspec.ehec-0024-2013.

Melton-Celsa, A. R. (2014b). Shiga Toxin (Stx) Classification, Structure, and Function. *Microbiol Spectr* 2, EHEC-0024-2013.

Menge, C. (2020). The Role of Escherichia coli Shiga Toxins in STEC Colonization of Cattle. *Toxins* 12, 607. doi: 10.3390/toxins12090607.

Mühldorfer, I., Hacker, J., Keusch, G. T., Acheson, D. W., Tschäpe, H., Kane, A. V., et al. (1996). Regulation of the Shiga-like toxin II operon in Escherichia coli. *Infect Immun* 64, 495–502.

Munns, K. D., Zaheer, R., Xu, Y., Stanford, K., Laing, C. R., Gannon, V. P. J., et al. (2016). Comparative Genomic Analysis of Escherichia coli O157:H7 Isolated from Super-Shedder and Low-Shedder Cattle. *Plos One* 11, e0151673. doi: 10.1371/journal.pone.0151673.

Nadler, C., Baruch, K., Kobi, S., Mills, E., Haviv, G., Farago, M., et al. (2010). The Type III Secretion Effector NleE Inhibits NF-κB Activation. *Plos Pathog* 6, e1000743. doi: 10.1371/journal.ppat.1000743.

Nart, P., Naylor, S. W., Huntley, J. F., McKendrick, I. J., Gally, D. L., and Low, J. C. (2008). Responses of Cattle to Gastrointestinal Colonization by Escherichia coli O157:H7 ▿ . *Infect Immun* 76, 5366–5372. doi: 10.1128/iai.01223-07.

Nataro, J. P., and Kaper, J. B. (1998). Diarrheagenic Escherichia coli. *Clin Microbiol Rev* 11, 142–201.

Newton, H. J., Sloan, J., Bulach, D. M., Seemann, T., Allison, C. C., Tauschek, M., et al. (2009). Shiga Toxin–producing Escherichia coli Strains Negative for Locus of Enterocyte Effacement. *Emerg Infect Dis* 15, 372–380. doi: 10.3201/eid1502.080631.

Nicholson, J. K., Holmes, E., Kinross, J., Burcelin, R., Gibson, G., Jia, W., et al. (2012). Host-Gut Microbiota Metabolic Interactions. *Science* 336, 1262–1267. doi: 10.1126/science.1223813.

Onyeka, L. O., Adesiyun, A. A., Keddy, K. H., Manqele, A., Madoroba, E., and Thompson, P. N. (2021). Prevalence, risk factors and molecular characteristics of Shiga toxin-producing Escherichia coli in beef abattoirs in Gauteng, South Africa. *Food Control* 123, 107746. doi: 10.1016/j.foodcont.2020.107746.

Paton, A. W., Beutin, L., and Paton, J. C. (1995). Heterogeneity of the amino-acid sequences of Escherichia coli shiga-like toxin type-I operons. *Gene* 153, 71–74. doi: 10.1016/0378-1119(94)00777-p.

Paton, A. W., Paton, J. C., and Manning, P. A. (1993). Polymerase chain reaction amplification, cloning and sequencing of variant Escherichia coli Shiga-like toxin type II operons. *Microb Pathogenesis* 15, 77–82. doi: 10.1006/mpat.1993.1058.

Paton, J. C., and Paton, A. W. (1998). Pathogenesis and Diagnosis of Shiga Toxin-Producing Escherichia coli Infections. *Clin Microbiol Rev* 11, 450–479. doi: 10.1128/cmr.11.3.450.

Perna, N. T., Mayhew, G. F., Pósfai, G., Elliott, S., Donnenberg, M. S., Kaper, J. B., et al. (1998). Molecular Evolution of a Pathogenicity Island from Enterohemorrhagic Escherichia coli O157:H7. *Infect Immun* 66, 3810–3817. doi: 10.1128/iai.66.8.3810-3817.1998.

Peroutka-Bigus, N., Nielsen, D. W., Trachsel, J., Mou, K. T., Sharma, V. K., Kudva, I. T., et al. (2022). Phenotypic and genomic comparison of human outbreak and cattle-associated Shiga toxin-producing Escherichia coli O157:H7. *Biorxiv*, 2022.09.30.510420. doi: 10.1101/2022.09.30.510420.

Persson, S., Olsen, K. E. P., Ethelberg, S., and Scheutz, F. (2007). Subtyping Method for Escherichia coli Shiga Toxin (Verocytotoxin) 2 Variants and Correlations to Clinical Manifestations ▿ . *J Clin Microbiol* 45, 2020–2024. doi: 10.1128/jcm.02591-06.

Philpott, D. J., Ackerley, C. A., Kiliaan, A. J., Karmali, M. A., Perdue, M. H., and Sherman, P. M. (1997). Translocation of verotoxin-1 across T84 monolayers: mechanism of bacterial toxin penetration of epithelium. *Am J Physiol-gastr L* 273, G1349–G1358. doi: 10.1152/ajpgi.1997.273.6.g1349.

Piérard, D., Muyldermans, G., Moriau, L., Stevens, D., and Lauwers, S. (1998). Identification of new verocytotoxin type 2 variant B-subunit genes in human and animal Escherichia coli isolates. *J Clin Microbiol* 36, 3317–22.

Ponchel, F., Toomes, C., Bransfield, K., Leong, F. T., Douglas, S. H., Field, S. L., et al. (2003). Real-time PCR based on SYBR-Green I fluorescence: An alternative to the TaqMan assay for

a relative quantification of gene rearrangements, gene amplifications and micro gene deletions. *Bmc Biotechnol* 3, 18. doi: 10.1186/1472-6750-3-18.

Poretsky, R., Rodriguez-R, L. M., Luo, C., Tsementzi, D., and Konstantinidis, K. T. (2014). Strengths and Limitations of 16S rRNA Gene Amplicon Sequencing in Revealing Temporal Microbial Community Dynamics. *Plos One* 9, e93827. doi: 10.1371/journal.pone.0093827.

Proença, J. T., Barral, D. C., and Gordo, I. (2017). Commensal-to-pathogen transition: One-single transposon insertion results in two pathoadaptive traits in Escherichia coli -macrophage interaction. *Sci Rep-uk* 7, 4504. doi: 10.1038/s41598-017-04081-1.

Puente, J. L., and Finlay, B. B. (2001). Principles of Bacterial Pathogenesis. 387–456. doi: 10.1016/b978-012304220-0/50010-8.

Raa, H., Grimmer, S., Schwudke, D., Bergan, J., Wälchli, S., Skotland, T., et al. (2009). Glycosphingolipid Requirements for Endosome-to-Golgi Transport of Shiga Toxin. *Traffic* 10, 868–882. doi: 10.1111/j.1600-0854.2009.00919.x.

Rangel, J. M., Sparling, P. H., Crowe, C., Griffin, P. M., and Swerdlow, D. L. (2005). Epidemiology of Escherichia coli O157:H7 Outbreaks, United States, 1982–2002 - Volume 11, Number 4—April 2005 - Emerging Infectious Diseases journal - CDC. *Emerg Infect Dis* 11, 603–609. doi: 10.3201/eid1104.040739.

Rebala, G., Ravi, A., and Churiwala, S. (2019). An Introduction to Machine Learning. 1–17. doi: 10.1007/978-3-030-15729-6_1.

Reinstein, S., Fox, J. T., Shi, X., and Nagaraja, T. G. (2007). Prevalence of Escherichia coli O157:H7 in Gallbladders of Beef Cattle. *Appl Environ Microb* 73, 1002–1004. doi: 10.1128/aem.02037-06.

Reusch, D. B., Alley, R. B., and Hewitson, B. C. (2005). Relative Performance of Self-Organizing Maps and Principal Component Analysis in Pattern Extraction from Synthetic Climatological Data. *Polar Geogr* 29, 188–212. doi: 10.1080/789610199.

Reynolds, C., Checkley, S., Chui, L., Otto, S., and Neumann, N. F. (2020). Evaluating the risks associated with Shiga-toxin-producing Escherichia coli (STEC) in private well waters in Canada. *Can J Microbiol* 66, 337–350. doi: 10.1139/cjm-2019-0329.

Rhoads, D. D., Cox, S. B., Rees, E. J., Sun, Y., and Wolcott, R. D. (2012). Clinical identification of bacteria in human chronic wound infections: culturing vs. 16S ribosomal DNA sequencing. *Bmc Infect Dis* 12, 321. doi: 10.1186/1471-2334-12-321.

Riley, L. W., Remis, R. S., Helgerson, S. D., McGee, H. B., Wells, J. G., Davis, B. R., et al. (1983). Hemorrhagic Colitis Associated with a Rare Escherichia coli Serotype. *New Engl J Medicine* 308, 681–685. doi: 10.1056/nejm198303243081203.

Rodríguez-Rubio, L., Haarmann, N., Schwidder, M., Muniesa, M., and Schmidt, H. (2021). Bacteriophages of Shiga Toxin-Producing Escherichia coli and Their Contribution to Pathogenicity. *Pathogens* 10, 404. doi: 10.3390/pathogens10040404.

Roussel, C., Cordonnier, C., Livrelli, V., de Wiele, T. V., and Blanquet-Diot, S. (2017). *Escherichia coli* - Recent Advances on Physiology, Pathogenesis and Biotechnological Applications. doi: 10.5772/intechopen.68309.

Ruan, Q., Dutta, D., Schwalbach, M. S., Steele, J. A., Fuhrman, J. A., and Sun, F. (2006). Local similarity analysis reveals unique associations among marine bacterioplankton species and environmental factors. *Bioinformatics* 22, 2532–2538. doi: 10.1093/bioinformatics/btl417.

Samuel, J. E., Perera, L. P., Ward, S., O'Brien, A. D., Ginsburg, V., and Krivan, H. C. (1990). Comparison of the glycolipid receptor specificities of Shiga-like toxin type II and Shiga-like toxin type II variants. *Infect Immun* 58, 611–618. doi: 10.1128/iai.58.3.611-618.1990.

Sandvig, K., Bergan, J., Dyve, A.-B., Skotland, T., and Torgersen, M. L. (2010). Endocytosis and retrograde transport of Shiga toxin. *Toxicon* 56, 1181–1185. doi: 10.1016/j.toxicon.2009.11.021.

Scallan, E., Hoekstra, R. M., Angulo, F. J., Tauxe, R. V., Widdowson, M.-A., Roy, S. L., et al. (2011). Foodborne Illness Acquired in the United States—Major Pathogens. *Emerg Infect Dis* 17, 7–15. doi: 10.3201/eid1701.p11101.

Scheutz, F., Teel, L. D., Beutin, L., Piérard, D., Buvens, G., Karch, H., et al. (2012). Multicenter Evaluation of a Sequence-Based Protocol for Subtyping Shiga Toxins and Standardizing Stx Nomenclature. *J Clin Microbiol* 50, 2951–2963. doi: 10.1128/jcm.00860-12.

Schierack, P., Kleta, S., Tedin, K., Babila, J. T., Oswald, S., Oelschlaeger, T. A., et al. (2011). E. coli Nissle 1917 Affects Salmonella Adhesion to Porcine Intestinal Epithelial Cells. *Plos One* 6, e14712. doi: 10.1371/journal.pone.0014712.

Schmidt, H. (2001). Shiga-toxin-converting bacteriophages. *Res Microbiol* 152, 687–695. doi: 10.1016/s0923-2508(01)01249-9.

Schmidt, H., Scheef, J., Morabito, S., Caprioli, A., Wieler, L. H., and Karch, H. (2000). A New Shiga Toxin 2 Variant (Stx2f) fromEscherichia coli Isolated from Pigeons. *Appl Environ Microb* 66, 1205–1208. doi: 10.1128/aem.66.3.1205-1208.2000.

Serra-Moreno, R., Jofre, J., and Muniesa, M. (2008). The CI Repressors of Shiga Toxin-Converting Prophages Are Involved in Coinfection of Escherichia coli Strains, Which Causes a Down Regulation in the Production of Shiga Toxin 2 ▿ . *J Bacteriol* 190, 4722–4735. doi: 10.1128/jb.00069-08.

Sheerin, N. S., and Glover, E. (2019). Haemolytic uremic syndrome: diagnosis and management. *F1000Research* 8, F1000 Faculty Rev-1690. doi: 10.12688/f1000research.19957.1.

Sherman, P. M., Johnson-Henry, K. C., Yeung, H. P., Ngo, P. S. C., Goulet, J., and Tompkins, T. A. (2005). Probiotics Reduce Enterohemorrhagic Escherichia coli O157:H7- and Enteropathogenic E. coli O127:H6-Induced Changes in Polarized T84 Epithelial Cell

Monolayers by Reducing Bacterial Adhesion and Cytoskeletal Rearrangements. *Infect Immun* 73, 5183–5188. doi: 10.1128/iai.73.8.5183-5188.2005.

Smet, R. D., and Marchal, K. (2010). Advantages and limitations of current network inference methods. *Nat Rev Microbiol* 8, 717–729. doi: 10.1038/nrmicro2419.

Sogin, M. L., Morrison, H. G., Huber, J. A., Welch, D. M., Huse, S. M., Neal, P. R., et al. (2006). Microbial diversity in the deep sea and the underexplored "rare biosphere." *Proc National Acad Sci* 103, 12115–12120. doi: 10.1073/pnas.0605127103.

Spooner, R. A., and Lord, J. M. (2011). Ricin and Shiga Toxins, Pathogenesis, Immunity, Vaccines and Therapeutics. *Curr Top Microbiol* 357, 19–40. doi: 10.1007/82_2011_154.

Stark, R., Grzelak, M., and Hadfield, J. (2019). RNA sequencing: the teenage years. *Nat Rev Genet* 20, 631–656. doi: 10.1038/s41576-019-0150-2.

Stein, R. A., and Katz, D. E. (2017). Escherichia coli, cattle and the propagation of disease. *Fems Microbiol Lett* 364, fnx050. doi: 10.1093/femsle/fnx050.

Stenkamp-Strahm, C., McConnel, C., Magzamen, S., Abdo, Z., and Reynolds, S. (2018). Associations between Escherichia coli O157 shedding and the faecal microbiota of dairy cows. *J Appl Microbiol* 124, 881–898. doi: 10.1111/jam.13679.

Tesh, V. L., Burris, J. A., Owens, J. W., Gordon, V. M., Wadolkowski, E. A., O'Brien, A. D., et al. (1993). Comparison of the relative toxicities of Shiga-like toxins type I and type II for mice. *Infect Immun* 61, 3392–402.

Venegas-Vargas, C., Henderson, S., Khare, A., Mosci, R. E., Lehnert, J. D., Singh, P., et al. (2016). Factors Associated with Shiga Toxin-Producing Escherichia coli Shedding by Dairy and Beef Cattle. *Appl Environ Microb* 82, 5049–5056. doi: 10.1128/aem.00829-16.

Wagner, P. L., Livny, J., Neely, M. N., Acheson, D. W. K., Friedman, D. I., and Waldor, M. K. (2002). Bacteriophage control of Shiga toxin 1 production and release by Escherichia coli. *Mol Microbiol* 44, 957–970. doi: 10.1046/j.1365-2958.2002.02950.x.

Walle, K. V., Vanrompay, D., and Cox, E. (2013). Bovine innate and adaptive immune responses against Escherichia coli O157:H7 and vaccination strategies to reduce faecal shedding in ruminants. *Vet Immunol Immunop* 152, 109–120. doi: 10.1016/j.vetimm.2012.09.028.

Wan, Y., Zuo, T., Xu, Z., Zhang, F., Zhan, H., CHAN, D., et al. (2021). Underdevelopment of the gut microbiota and bacteria species as non-invasive markers of prediction in children with autism spectrum disorder. *Gut*, gutjnl-2020-324015. doi: 10.1136/gutjnl-2020-324015.

Wang, H., Rogers, T. J., Paton, J. C., and Paton, A. W. (2014). Differential Effects of Escherichia coli Subtilase Cytotoxin and Shiga Toxin 2 on Chemokine and Proinflammatory

Cytokine Expression in Human Macrophage, Colonic Epithelial, and Brain Microvascular Endothelial Cell Lines. *Infect Immun* 82, 3567–3579. doi: 10.1128/iai.02120-14.

Wang, O., Liang, G., McAllister, T. A., Plastow, G., Stanford, K., Selinger, B., et al. (2016). Comparative Transcriptomic Analysis of Rectal Tissue from Beef Steers Revealed Reduced Host Immunity in Escherichia coli O157:H7 Super-Shedders. *Plos One* 11, e0151284. doi: 10.1371/journal.pone.0151284.

Wang, O., McAllister, T. A., Plastow, G., Stanford, K., Selinger, B., and Guan, L. L. (2017). Host mechanisms involved in cattle Escherichia coli O157 shedding: a fundamental understanding for reducing foodborne pathogen in food animal production. *Sci Rep-uk* 7, 7630. doi: 10.1038/s41598-017-06737-4.

Wang, O., McAllister, T. A., Plastow, G., Stanford, K., Selinger, B., and Guan, L. L. (2018). Interactions of the Hindgut Mucosa-Associated Microbiome with Its Host Regulate Shedding of Escherichia coli O157:H7 by Cattle. *Appl Environ Microb* 84, e01738-17. doi: 10.1128/aem.01738-17.

Wang, O., Zhou, M., Chen, Y., McAllister, T. A., Plastow, G., Stanford, K., et al. (2021). MicroRNAomes of Cattle Intestinal Tissues Revealed Possible miRNA Regulated Mechanisms Involved in Escherichia coli O157 Fecal Shedding. *Front Cell Infect Mi* 11, 634505. doi: 10.3389/fcimb.2021.634505.

Waters, J. R., Sharp, J. C. M., and Dev, V. J. (1994). Infection Caused by Escherichia coli O157:H7 in Alberta, Canada, and in Scotland: A Five-Year Review, 1987-1991. *Clin Infect Dis* 19, 834–843. doi: 10.1093/clinids/19.5.834.

Weinstein, D. L., Jackson, M. P., Samuel, J. E., Holmes, R. K., and O'Brien, A. D. (1988). Cloning and sequencing of a Shiga-like toxin type II variant from Escherichia coli strain responsible for edema disease of swine. *J Bacteriol* 170, 4223–4230. doi: 10.1128/jb.170.9.4223-4230.1988.

Wells, J. G., Davis, B. R., Wachsmuth, I. K., Riley, L. W., Remis, R. S., Sokolow, R., et al. (1983). Laboratory investigation of hemorrhagic colitis outbreaks associated with a rare Escherichia coli serotype. *J Clin Microbiol* 18, 512–520. doi: 10.1128/jcm.18.3.512-520.1983.

WILLIAMS, K. J., WARD, M. P., DHUNGYEL, O. P., and HALL, E. J. S. (2015). Risk factors for Escherichia coli O157 shedding and super-shedding by dairy heifers at pasture. *Epidemiol Infect* 143, 1004–1015. doi: 10.1017/s0950268814001630.

Xu, Y., Dugat-Bony, E., Zaheer, R., Selinger, L., Barbieri, R., Munns, K., et al. (2014). Escherichia coli O157:H7 Super-Shedder and Non-Shedder Feedlot Steers Harbour Distinct Fecal Bacterial Communities. *Plos One* 9, e98115. doi: 10.1371/journal.pone.0098115.

Yoon, S. C., Windham, W. R., Ladely, S., Heitschmidt, G. W., Lawrence, K. C., Park, B., et al. (2013). Differentiation of big-six non-O157 Shiga-toxin producing Escherichia coli (STEC) on

spread plates of mixed cultures using hyperspectral imaging. *J Food Meas Charact* 7, 47–59. doi: 10.1007/s11694-013-9137-4.

Zaheer, R., Dugat-Bony, E., Holman, D., Cousteix, E., Xu, Y., Munns, K., et al. (2017). Changes in bacterial community composition of Escherichia coli O157:H7 super-shedder cattle occur in the lower intestine. *Plos One* 12, e0170050. doi: 10.1371/journal.pone.0170050.

Zhou, H., Beltrán, J. F., and Brito, I. L. (2022). Host-microbiome protein-protein interactions capture disease-relevant pathways. *Genome Biol* 23, 72. doi: 10.1186/s13059-022-02643-9.

Zhou, J., and Ning, D. (2017). Stochastic Community Assembly: Does It Matter in Microbial Ecology? *Microbiol Mol Biol R* 81, e00002-17. doi: 10.1128/mmbr.00002-17.

1.9 Tables and figures

Table 1.1. Major pathotypes and their subset of pathogenic *Escherichia coli*

Type	Pathotypes	Subset
ETEC	Enterotoxigenic *Escherichia coli*	N/A
STEC	Shiga toxin-producing *Escherichia coli*	EDEC: Edema disease *Escherichia coli*
EPEC	Enteropathogenic *Escherichia coli*	N/A
ExPEC	Extraintestinal *Escherichia coli*	APEC: Avian pathogenic *Escherichia coli* SEPEC: Septicemic *Escherichia coli* UPEC: Uropathogenic *Escherichia coli*

Table 1.2. Animal disease caused by pathogenic *E.coli*

Species	Disease	Pathotype	Localization	Age
Poultry	Embryonic mortality	-	Egg	-
	Colisepticemia	ExPEC (APEC)	Upperrespiratory tract, air sacs, systemic	5-10 weeks
	Swollen head, dermatitis, cellulitis	-	Localized infections	adult
	Diarrhea	-	Intestine	-
Cattle	Newborn diarrhea	ETEC	Small intestine	0-1 week
	Hemorrhagic dysentery	EPEC, STEC	Colon	1-6 weeks
	Septicemia	ExPEC	Systemic	0-1 week
	Mastitis	-	Mammary gland	adult
	Diarrhea	ETEC, EPEC	Small and large intestines	young animal
	Urinary tract infection	ExPEC, UPEC	Kidney, urinary tract	
Pig	Newborn diarrhea	ETEC	Small intestine	0-1 week
	Young pig diarrhea	ETEC		2-4 weeks
	Postweaning diarrhea	ETEC, EPEC		Post-weaning
	Edema disease	STEC, EDEC		4-8 weeks
	Hemorrhagic gastroenteritis	ETEC		1-8 weeks
	Septicemia	ExPEC (SEPEC), ETEC	Systemic	0-4 weeks
Rabbit	Newborn and Weaning diarrhea	EPEC	Small and large intestines	-

Table 1.3. Prototype toxins and strains that produce those toxins

Toxin type(s)	Linked with serious human disease; difference(s) from prototype toxin[a]	Reference(s)
Stx	Yes	(Hale and Formal, 1980)
Stx1a	Yes	(Riley et al., 1983)
Stx1c	No; immunologically distinct	(Paton et al., 1995; Koch et al., 2001)
Stx1d	No; immunologically distinct, less potent	(Bürk et al., 2003)
Stx2a	Yes	(Riley et al., 1983)
Stx2b (originally named VT-2d or Stx2d)	No; the B subunit gene was not detected by methods used to detect other *stx2* B subunit genes	(Piérard et al., 1998)
Stx2c	Yes, less toxic to Vero cells and mice	(Paton et al., 1993)
Stx2d (Stx2dact)	Yes; more toxic after incubation with elastase, less toxic to Vero cells	(Persson et al., 2007)
Stx2e	No; binds to Gb4, associated with disease in pigs	(Weinstein et al., 1988)
Stx2f	No; originally isolated in STEC from pigeons; immunologically distinct	(Schmidt et al., 2000)
Stx2g	No; the *stx2g* gene is not amplified by primers specific for *stx2a*	(Leung et al., 2003)

[a]Prototype toxin indicated in bold. This table is cited and modified from (Melton-Celsa, 2014a)

Table 1.4. Examples of studies examining the relative role of deterministic and stochastic processes in structured bacterial communities

Ecosystem	Stochasticity measurement	Relative importance		References
		Determinism	Stochasticity	
Freshwater lakes	Neutral model	-	✔	Roguet et al. (2015)
Soil spanning 105 years succession	Newly designed model	Community when succession proceeded	Community at initial stage	Dini-Andreote et al. (2015)
Coastal water	Stegen's framework	High eutrophication	Low eutrophication	Dai et al. (2017)
Coastal water	Neutral model and variation partitioning	-	Both abundant and rare communities	Mo et al. (2018)

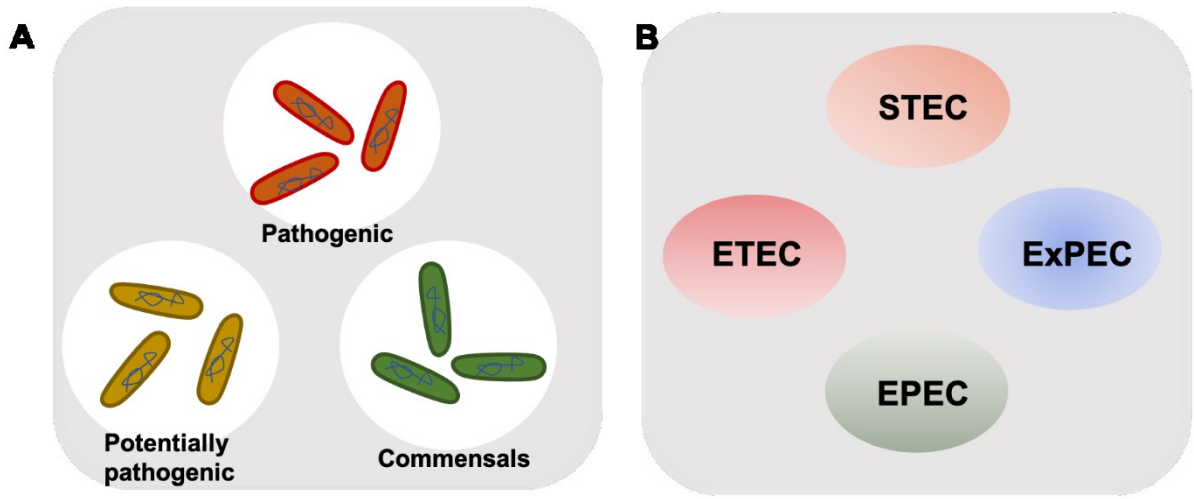


Figure 1.1. *E.coli* (A) and Pathogenic *E.coli* (B) classification

STEC: Shiga toxin-producing *Escherichia coli;* ETEC*:* Enterotoxigenic *Escherichia coli;* ExPEC*:* Extraintestinal *Escherichia coli;* EPEC*:* Enteropathogenic *Escherichia coli*

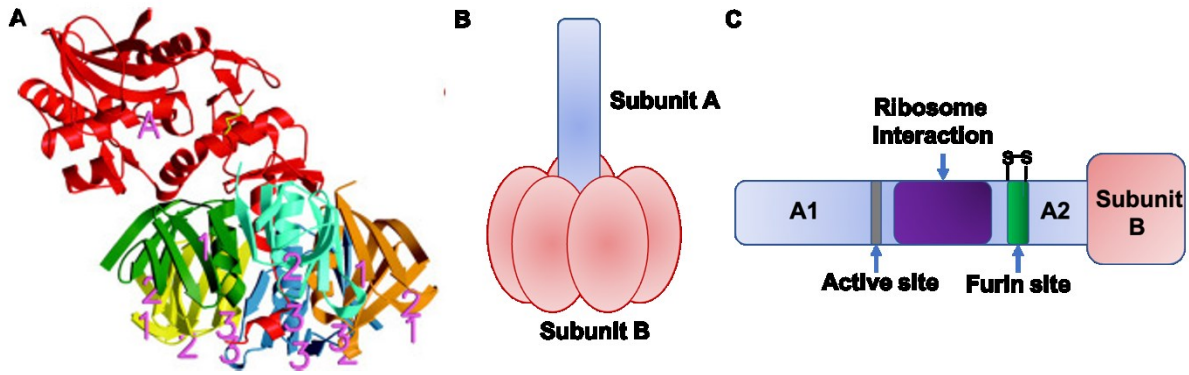


Figure 1.2. Structure of stx and subunit molecules

A. A ribbon diagram of stx2 adapted from (Fraser et al., 2004). The A-subunit is red, whereas the B-subunits are orange, cyan, green, yellow, and blue. The active site in the A-subunit is marked by the magenta letter A. The side chains of the cysteine residues that link A1 and A2 are depicted in yellow. The sites equivalent to the Gb_3-binding sites on the B-pentamer of Stx1 are shown by magenta numbers that distinguish the type of binding site.

B. A schematic diagram of a generic stx molecule adapted from (Fraser et al., 2004), the A subunit is blue, the B subunit is red.

C. A schematic diagram of subunit A revised from (Melton-Celsa, 2014a). The active site glutamic acid is indicated as a vertical grey line, the ribosome interaction region is shown in purple, the protease (furin) sensitive site is depicted in green, and the B pentamer is a red block. The disulfide bridge that connects the A1 subunit and the A2 peptide is shown above the protease sensitive site. Not to scale.

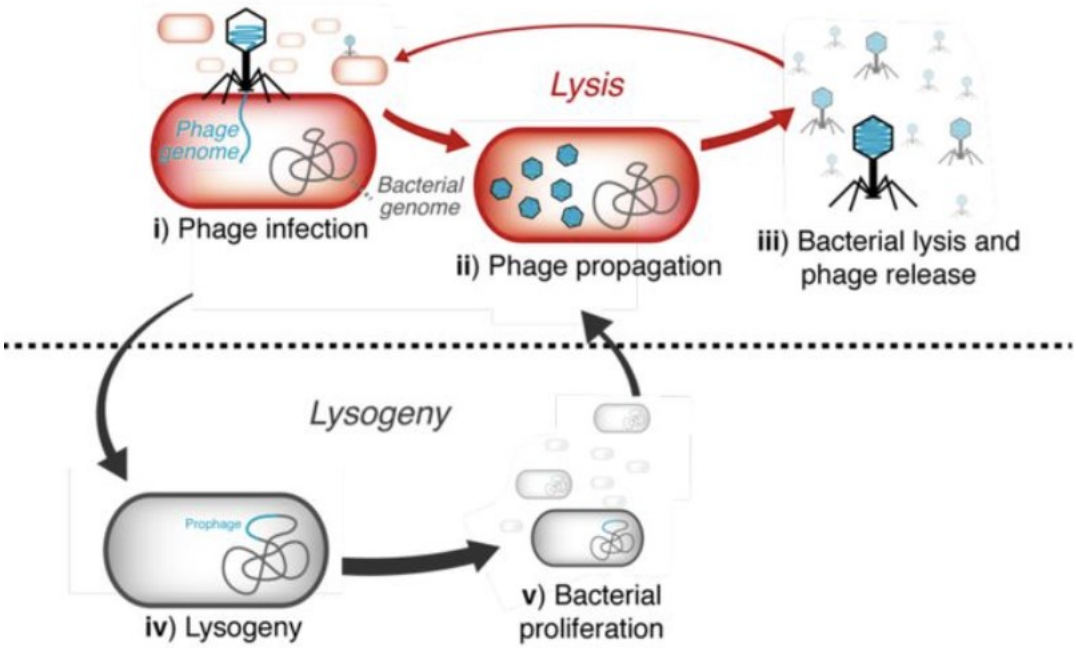


Figure 1.3. The lytic and lysogenic cycle of phages in STEC

i) In lytic phage infection, the phage particle injects its genetic material into a bacterium and ii) directs the cell to produce phage components, iii) which are released upon cell lysis to continued infection. iv) Temperate phages can also participate in the lysogenic life cycle where they integrate their DNA into the bacterial genome and v) remain as a prophage during bacterial proliferation with the possibility of entering lytic replication in the future adopted from (Hsu et al., 2019).

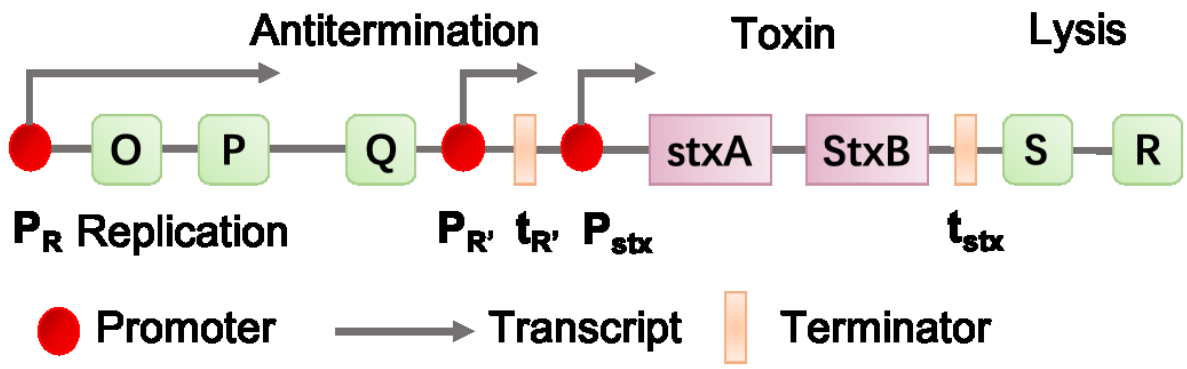


Figure 1.4. The late region of Shiga toxin converting lambdoid phage

Genes coding for replication proteins (O and P), a gene for anti-terminator Q, the Shiga toxin genes (stxA, stxB), and genes coding for proteins causing cell lysis (S and R), are marked. Promoters are shown by ovals, terminators by rectangles, and arrows indicate the directionality of transcription. Revised based on (Łoś et al., 2020).

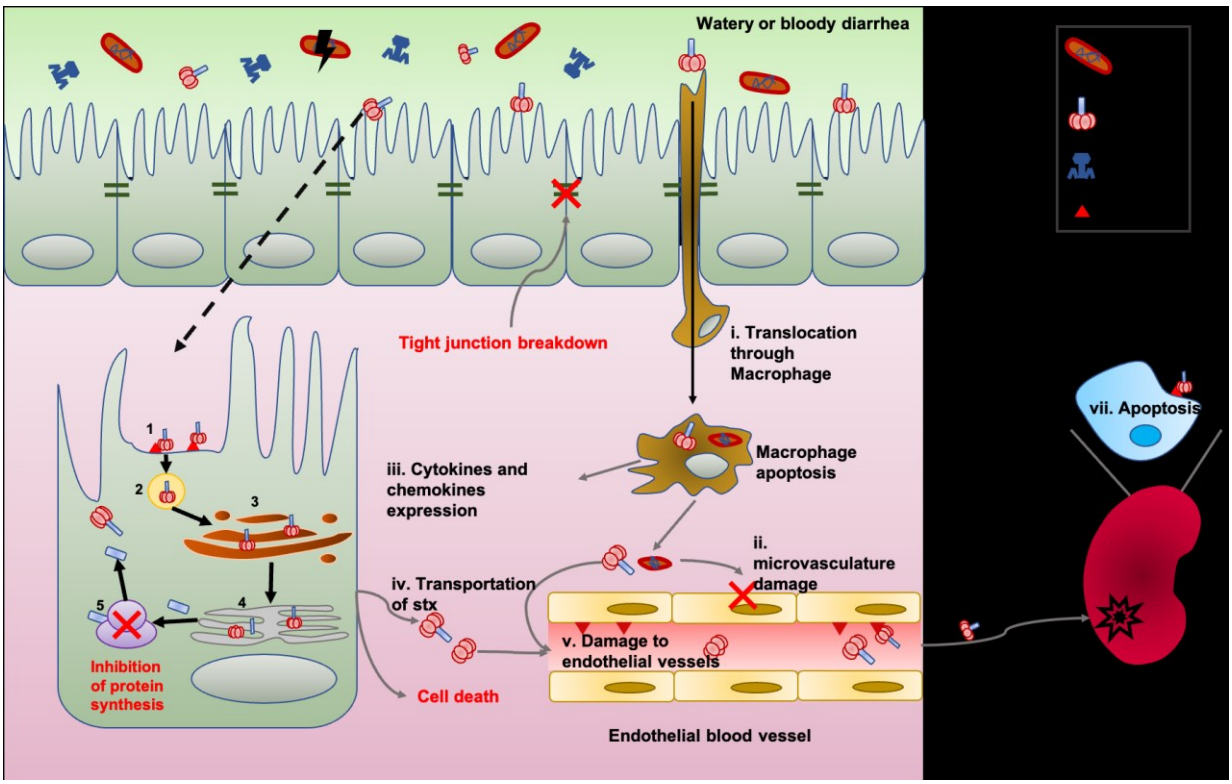


Figure 1.5. Pathogenesis of shiga toxin in human gastrointestinal tract and kidney

The figure is revised based on (Kaper et al., 2004; Melton-Celsa, 2014b; Castro et al., 2017; Hall et al., 2017; Roussel et al., 2017; Lee and Tesh, 2019).

Mechanisms of action of Shiga toxin (Refers to retrograde pathway) are marked by numbers : (1) Binding of stx to the Gb3 receptor within lipid raft on the surface of intestinal epithelial cells; (2) stx-Gb3 complex internalizes within an endosome (3) stx traffics to the Golgi complex (4) stx transports to ER, and the disulfide bridge that keeps the A1 tethered to A2B5 is reduced (5) Action of the A1 portion on the rRNA in the 28S portion, replacing an Adenine moiety (6) inhibition of protein synthesis and cell death.

Shiga toxin may also translocate through i) macrophage uptake. Once in the submucosa, the toxins may directly ii) damage the intestinal microvasculature and iii) elicit cytokine and chemokine production by resident macrophages. Macrophage activation results in the

infiltration of neutrophils and monocytes which may further exacerbate tissue damage. Neutrophils and monocytes may also act as “carrier” cells to iv) transport toxins in the bloodstream. Once in microvessels that are rich in the toxin receptor, Gb3, the toxins may be transferred from the carrier cells to v) damage endothelial cells. Toxins will be further vi) transported to kidney where tubular epithelial cells express a higher level of Gb3 through blood circulation. The localized epithelial cells in the kidney will be vii) damaged and promote the pathogenesis of HUS in humans.

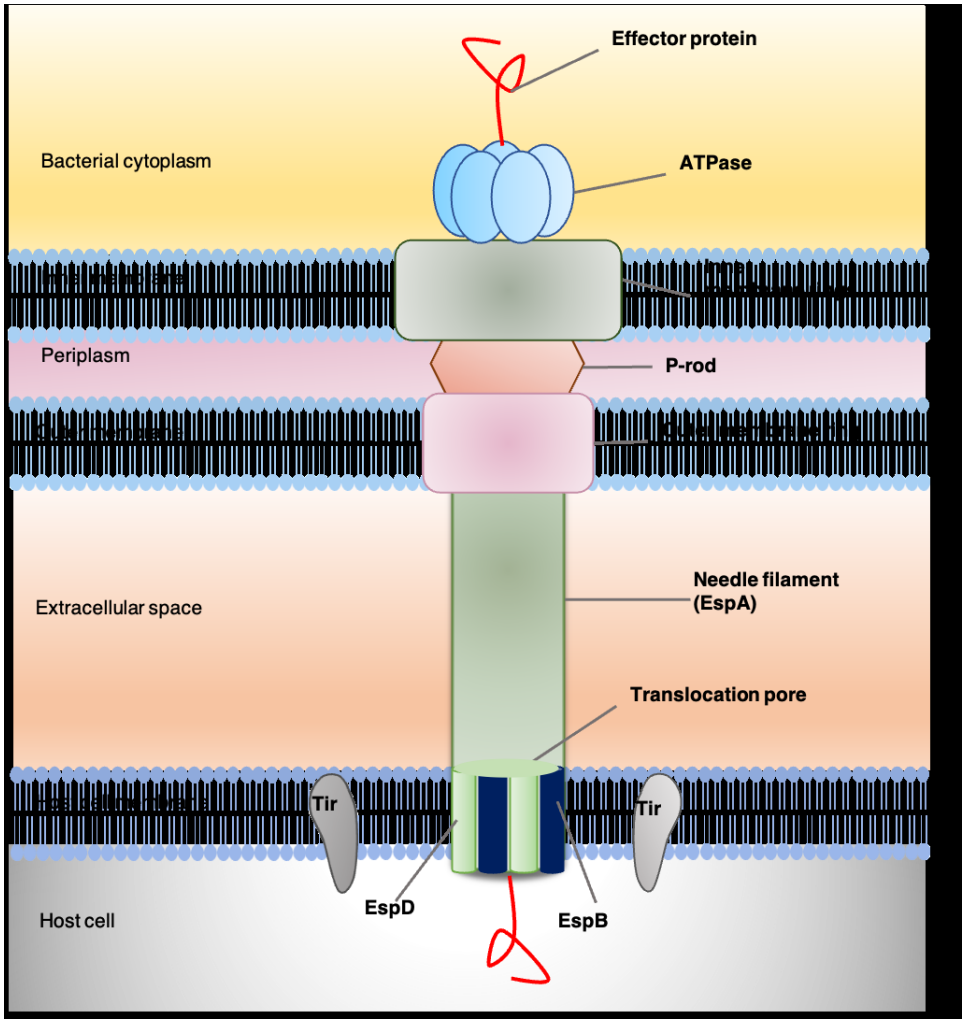


Figure 1.6. Schematic diagram of T3SS architectures of STEC- epithelium interactions

The basal body of the T3SS spans the bacterial inner membrane (IM) and outer membrane (OM), and comprises ring structures that are connected by a periplasmic rod (P-rod) across the periplasm. The basal body associates with an extracellular needle (needle filament, *E.coli* secreted proteins A, EspA) and forms a channel-like translocon (EspB and EspD) that inserts into the host cell membrane. The energy for the docking and unfolding of T3SS substrates, including effector proteins and T3SS-accessory extracellular proteins, is provided by a cytoplasmic ATPase associated with the T3SS. The figure is revised based on (Ji and Dong, 2015; Gaytán et al., 2016; Deng et al., 2017).

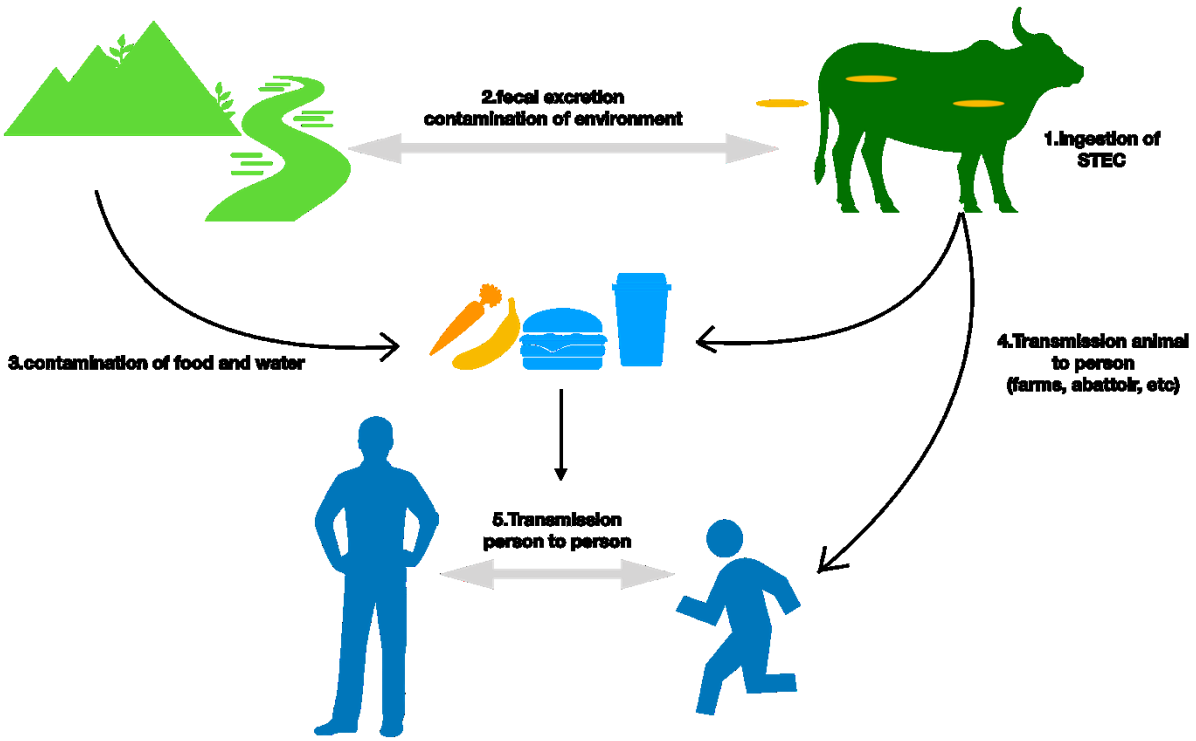


Figure 1.7. Zoonotic STEC transmission

(1) Cattle can ingest STEC through oral routes and then (2) secrete STEC through feces and contaminate environments. (3) The secreted STEC can survive in the environments and then contaminate food and water nearby. (4) The STEC transmission can occur directly from infected cattle to humans on farms or abattoirs or directly from contaminated food and water. (5) Once human ingest STEC, it can spread through human-human transmission as well. This figure is modified based on (Fairbrother and Nadeau, 2006).

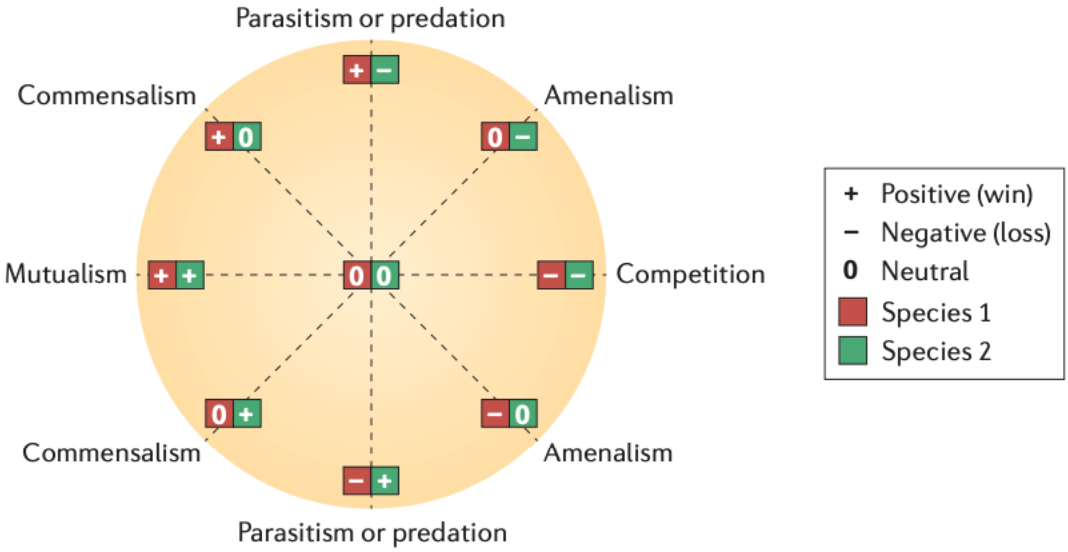


Figure 1.8. Summary of ecological interactions between members of different species

For each interaction partner, there are three possible outcomes: positive (+), negative (–), and neutral (0). For instance, in parasitism, the parasite benefits from the relationship (+), whereas the host is harmed (–); this relationship is thus represented by the symbol pair+–. Figure is adopted from Faust and Raes (2012).

Chapter 2. Abundance and expression of Shiga toxin genes in *Escherichia coli* at the recto-anal junction relates to host immune genes[1]

2.1 Introduction

Shiga toxin-producing *Escherichia coli* (STEC) cause foodborne disease that can lead to hemolytic uremic syndrome (HUS) and hemorrhagic colitis (HC)(Karmali et al., 1983). Approximately, 2.8 million acute illnesses around the world are attributed to STEC (Majowicz et al., 2014), with 60,000 of these occurring in the US annually (Scallan et al., 2011). Many infections in humans are attributed to direct or indirect contact with food or water contaminated with cattle feces (Mir et al., 2016). Ruminants, especially cattle are the main reservoir who are asymptomatic carriers of O157 and non-O157 STEC strains with the recto-anal junction (RAJ) as the main colonization site (Wang et al., 2016). Most *E.coli* strains are commensals within the gut of cattle (Mir et al., 2016; Wang et al., 2016), and are shed into the environment through feces. Cattle that shed more than 10^4 CFU STEC per gram of feces are defined as "super-shedders" (SS), which are considered the primary source of STEC transmission on farms (Matthews et al., 2005). Although the incidence of *E. coli* O157:H7 causing disease in cattle is low, the prevalence of STEC including both *E. coli* O157:H7 and non-O157:H7 serotypes is not low in cattle ranging from 38.5%-75.0% (Cho et al., 2007). Both *E. coli* O157:H7 and non-O157:H7 serotypes can cause human disease and among non-O157 infections, up to 70% of

[1] Chapter 2 was published as a part of a paper: Pan Z, Chen Y, McAllister TA, Gänzle M, Plastow G, Guan LL. Abundance and Expression of Shiga Toxin Genes in Escherichia coli at the Recto-Anal Junction Relates to Host Immune Genes. Front Cell Infect Mi. 2021;11:633573.

human infections are attributed to six non-O157 STEC serogroups (O26, O45, O103, O111, O121, and O145) (Bosilevac and Koohmaraie, 2012).

Shiga toxins are the main virulence factors in STEC and other pathogenic bacterial species with the prototype toxins being designated as Shiga toxin 1a (stx1a) and Shiga toxin 2a (stx2a, Melton-Celsa, 2014). These toxins differ in their virulence and host specificity (Fuller et al., 2011; Lee and Tesh, 2019; Petro et al., 2019) with stx2 being most commonly associated with severe illness (HUS, hospitalization, and bloody diarrhea) in humans (Karmali et al., 1983; Panel et al., 2020). For example, 40% HUS, 41% hospitalization, and 43% bloody diarrhea cases reported in humans can be attributed to detectable stx2 (Panel et al., 2020). Therefore, identifying the abundance of *stx*1 and *stx*2 genes in cattle is important as they could harbour and shed STEC. However, information on the abundance and expression of *stx*1 and *stx*2 genes i*n vivo* (e.g. in RAJ) of feedlot cattle is lacking. We hypothesize that the expression and abundance of *stx* genes at the RAJ is influenced by cattle breed and expression of host immune genes. Genetic variation in the host has previously been linked to the level of expression of immune genes in SS, which also affected the attachment and the colonization of the mucosa by STEC (Wang et al., 2018). The understanding of abundance and expression of *stx* genes in STEC at the main colonization site and its linkage with host immune gene expression will provide insight into the host-STEC interactions at the RAJ of feedlot cattle.

2.2 Materials and methods

2.2.1 Animal populations and sample collection

In total, rectal tissue and contents were collected over two consecutive years (2014 and 2015) from 143 cattle representing Angus (AN, n=47), Charolais (CH, n=48), and a crossbreed named Kinsella Composite (KC, n=48) breed that were reared at the University of Alberta Roy Berg Kinsella Research Station. Sampling was performed when animals were slaughtered at a comparable age (Year 2014: 492d ± 30d; Year 2015: 496d ± 22d; P=0.11) in each year. Ten cm^2 of rectal tissue was collected from RAJ and 10 mL rectum contents were squeezed from each steer within 30 min after slaughter at a federally approved abattoir. The samples were deep-frozen immediately in liquid nitrogen and stored at -80 °C until use.

2.2.2 DNA and RNA extraction

Tissue and content samples of RAJ were ground into fine powder in liquid nitrogen and mixed homogeneously before DNA and RNA extraction. DNA was isolated from 0.1g powdered tissue using repeated bead beating and a column (RBBC) method (Yu and Morrison, 2004) and purified using the QIAmp Stool Mini Kit (Qiagen, Germany). The quantity and quality of DNA were assessed based on absorbance at 260 and 280 nm using the ND-1000 spectrophotometer (NanoDropTechnologies, Wilmington, USA). Trizol reagent (Invitrogen Corporation, Carlsbad, CA, USA) was used to isolate total RNA from 0.1 g powdered tissue following the manufacturer's protocol. RNA was purified using the RNeasy MinElute Cleanup kit (Qiagen, Valencia, CA, USA). Quality and quantity of RNA were assessed using Agilent 2200 TapeStation (Agilent Technologies, Santa Clara, CA, USA) and Qubit 3.0 Fluorometer (Invitrogen, Carlsbad, CA, USA), respectively. DNA was extracted from 0.5g of the RAJ

contents from each steer using the same bead beating method described above. DNA was obtained from contents of 131 steers and was used for downstream analysis.

2.2.3 Assessment of Shiga toxin gene abundance using qPCR

The DNA extracted from contents and tissues was used to evaluate the abundance of *Stx* genes using quantitative PCR (qPCR) with primers for the detection of all subtypes of *stx1* and *stx2* (Table 1) and SYBR Green I reagent (Fast SYBR green master mix; Applied Biosystems, Foster City, CA, USA). The qPCR was conducted in triplicates for each sample on a StepOnePlus™ Real-Time PCR System (Applied Biosystems, Foster City, CA, USA) with the program of one cycle at 95°C for 20 s followed by 40 cycles of 3 s at 95°C, 30 s at 60°C. Melting curve analysis with a temperature gradient of 0.1°C/s from 60 to 95°C with fluorescence signal measurement at 0.1°C intervals was performed to make sure targeted products were amplified specifically. The standard curve method was used to quantify *stx1* and *stx2* copy numbers. The standard curve was constructed by genomic DNAs isolated from strain *E.coli* FUA 1403 and *E.coli* FUA 1400, which contain *stx1* and *stx2*, respectively. The formula to calculate the absolute copy number of standard curves is described as follows (Li et al., 2009):

$$Absolute\ copy\ number\ \left(\frac{\#}{g\ Sample}\right) = \frac{Amount\left(\frac{g\ DNA}{g\ Sample}\right) * 6.022 * 10^{23}(\frac{\#}{mol})}{Length\ (bp) * 660\ (\frac{g\ DNA}{mol * bp})}$$

where 6.022 * 10^{23} represents the Avogadro's constant (# / mol); Length (bp) is the length of template DNA; 660 represents the average mass of 1bp double-strand DNA. The copy number of *stx1* or *stx2* was determined by relating threshold cycle (C_T) values to standard

curves based on the following regression formula (Li et al., 2009): Y = -3.193 * log X + 35.003 (Y, C_T value; X, copy number of 16S rRNA gene) (r^2 = 0.996). The qPCR amplification efficiency was 88-98%.

2.2.4 Detection of expression of stx and host immune genes using qRT-PCR

Total RNA (0.1 µg) was further subjected to reverse transcription to synthesize cDNA using a cDNA Synthesis Kit (Bio-Rad, Hercules, CA, USA). Single-stranded cDNA was amplified using Oligo(dT)$_{12-18}$ (Life Technologies, Carlsbad, CA, USA) and SuperScript^{TM} II RT (Life Technologies, Carlsbad, CA, USA) was used to synthesize double-strand cDNA. Primers for the detection of *eae* expression are shown in Table 1.1. Quantitative RT-PCR of *stx1*, *stx2* and *eae* was then performed using the double-strand cDNA and primers (Table 1.1) with the same thermal cycling program described above in triplicates for each sample. The expression of *stx1*, *stx2* and *eae* was quantified by standard curve method described above.

In addition, four genes reported to be differentially expressed between SS and non-shedding (NS) cattle (Wang et al., 2016); chemokine (C-C motif) ligand 21 (*CCL21*), lymphotoxin beta (*LTB*), CD19 molecule (*CD19*) and 4-domains, subfamily A, member 1 (*MS4A1*) were selected to study their relationship with *stx* gene abundance and expression. The same qPCR amplification conditions were used for the four genes with their respective primers (Table 2.1). Four commonly used housekeeping genes, including bovine *GAPDH,* 18S rRNA genes*, RPLP0,* and the *β-actin* gene*,* were also quantified by qPCR (Wang et al., 2016). As β-actin exhibited the most consistent Cq value it was used as the house-keeping gene for

evaluating relative gene expression. The relative expression of each gene (*stx1*, *stx2,* and immune genes) was measured by ΔCq value, which was calculated as (Wang et al., 2016):

$$\Delta C_q = C_q target\ genes - C_q referece\ gene$$

with a higher ΔCq representing the lower expression while a lower ΔCq indicating higher expression. The qPCR amplification efficiency was 88-98%.

2.2.5 Statistical analysis

The PROC MIXED model in SAS (ver. 9.13; SAS Institute Inc., Cary, NC, USA) was used to analyze the *stx1* and *stx2* abundance as well as host gene expressions together with all potential 2- and 3-way interactions among breeds, years, and sample types. In this statistical model, breed, sample type and year were analyzed as fixed effects with steers as the random effect. Interactions were removed from the model if they were not significant ($P > 0.05$). Least square means were compared using the Bonferroni mean separation method after the removal of insignificant interactions and the significance was considered at $P < 0.05$. The difference of prevalence of *stx1* and *stx2* was analyzed using Fisher's exact tests. Non-parametric Mann-Whitney U test in R was used to assess differences in host gene expression between Stx2+ (expressed) and Stx2- (not expressed) groups, with differences considered significant at $P<0.05$. Correlation analysis was performed based on Spearman's rank correlation coefficient (R) to identify relationships between expression of *stx2* and host genes using the "ggcorrplot" package in R with significance at $P<0.05$.

Isomap, a novel method for nonlinear dimensional reduction (Tenenbaum et al., 2000), was applied to determine the effect of breed, and sampling year on the expression of immune

genes and *stx2* using the "RDRToolbox" package in R. In addition, Davis-Bouldin-Index (DBIndex) was used to compute Euclidean metrics to validate the clustering patterns of the expression of immune genes and *stx2*, with the value ≤ 1 indicating a well-separated cluster (Davies and Bouldin, 1979). Correspondence analysis (CA) was used to identify relationships among expression patterns using the "FactoMineR" package (Tekaia, 2016).

2.2.6 Identification of potential gene markers for *stx* gene expression using mathematic models

The random forest model was used to identify predictive indicators for *stx2* expression with the "RandomForest" package in R. The host gene expression data were divided into two groups: *stx2*+ (expressed) and *stx2*- (not expressed). Two-thirds of each group was used as training data, and the rest (one-third) was used for validation. The accuracy rate (number of samples recognized correctly / total number of samples) was calculated to determine the model classification performance. The mean decrease in accuracy was used to assess the importance of host genes as predictive indicators of *stx2* expression. Variables with high mean decrease in accuracy indicate the higher contribution as compared to variables with low mean decrease accuracy (Han et al., 2016). The area under the ROC curve (AUC) was calculated to assess the robustness of the prediction model with the criteria being excellent (0.9-1.0), good (0.8-0.9), fair (0.7-0.8), weak (0.6-0.7) or fail (0.5-0.6) (Zhang et al., 2016). Moreover, the Boruta method, a random forest-based feature selection with the ability to remove less informative features, was used as a supportive approach to perform this prediction using the "Boruta" package in R (Kursa and Rudnicki, 2010).

2.3 Results

2.3.1 Factors affecting the abundance and prevalence of *stx1* and *stx2*

The sampling year significantly impacted the abundance and prevalence of *stx* genes identified in RAJ samples (P<0.01), therefore, the effect of breed on the prevalence and abundance of *stx1* and *stx*2 was analyzed separately for each year. The prevalence of *stx1* and *stx2* in tissue samples was not affected by breed in either year (Table 2.2). In year 1, the prevalence of *stx1* in contents was higher (P = 0.001, Table 2.2) in AN (n=18; 78%) compared to CH (n=7; 35%) and KC (n=6; 27%), and the prevalence of *stx2* was higher (P < 0.001, Table 2.2) in AN (n=22; 96%) and CH (n=20; 100%) than in KC (n=4; 18%). However, the prevalence of *stx1* and *stx2* in content samples collected in year 2 was not affected by breed (P_{stx1}=0.069, P_{stx2}=0.272, Table 2.2) with a tendency for breed to affect the prevalence of *stx1*.

The abundance of *stx1* and *stx2* was affected (P < 0.001) by sample type (tissue vs. contents) for both years (Table 2.3). An interaction effect between breed and sample type for the abundance of *stx1* and *stx2* was detected in year 1 (P_{stx1}<0.001, P_{stx2}<0.001, Table 2.3), but not in year 2 (P_{stx1}=0.28, P_{stx2}=0.12, Table 2.3). In year 1, the abundance of *stx1* in contents was affected by breed with its abundance higher in AN> CH> KC (P<0.001, Table 2.3), while its abundance in rectal tissue was under the detection limit (Table 2.3). For *stx2*, it was detected in both tissue and content samples in year 1 with no difference in the abundance of *stx2* in tissue samples (Table 2.3), but with the higher abundance in rectal contents of AN and CH as compared to KC steers (P<0.0001, Table 2.3). For year 2, the abundance of *stx1* or *stx2* did not

differ among breeds for either tissue or contents (Table 2.3), with the abundance of *stx1* and *stx2* in tissue being higher compared to that in contents (P_{stx1}<0.001, P_{stx2}<0.001, Table 2.3), respectively.

2.3.2 Expression of *stx1* and *stx2* associated with the rectal tissue of beef steers

Expression of bacterial *stx1* was not detected, and bacterial *stx2* (defined as *stx2*+) was only detected in mucosal tissue from 13 cattle (2014: n=6, 2015: n=7, Table 2.4). The expression of *stx2* was more common in KC (n=9; 70%) than in AN (n=2; 15%) and CH (n=2; 15%). The non-parametric Kruskal-Wallis test showed that *stx2* expression did not differ among breeds (ΔCq_{AN}=5.04; ΔCq_{CH}=5.11; ΔCq_{KC}=5.04; P= 0.31), but there was a trend for difference between sampling years ($\Delta Cq_{Year\ 2014}$=4.94; $\Delta Cq_{Year\ 2015}$=5.15; P = 0.06).

2.3.3 Expression of selected immune genes in RAJ tissue from beef steers

In year 1, the expression of four selected immune genes was not affected by breed. In year 2, only expression of *CD19* and *CCL21* differed among breeds (P_{CD19}=0.02, P_{CCL21}=0.0035, Table 2.5). There was no difference (P_{MS4A1}=0.36, P_{CD19}=0.62, P_{CCL21}=0.94, P_{LTB}=0.54, Table 2.6) in the expression of the four genes between *stx2*+ and *stx2*- steers. Visually, host gene expression patterns from tissue samples were affected by year among all samples ($Value_{Year}$ =0.81, Figure 2.1A) as well as among *stx2*+ samples ($Value_{Year}$ =0.75, Figure 2.1B). However, host gene expression patterns did not differ among breeds based on DBIndex clustering value among all samples ($Value_{breed}$=9.30, Figure 2.2A) or among *stx2*+ samples ($Value_{breed}$=1.64, Figure 2.2B).

2.3.4 Association between expressions of *stx2* and host immune genes

Expression of *stx2* was negatively correlated with the expression of *MS4A1* (R=-0.56, P=0.05, Table 2.7) and positively correlated with the expression of *LTB* (R=0.60, P=0.05, Table 2.7). Neither *CD19* nor *LTB* clustered with *Stx 2+* samples but *CD19* and *LTB* were positively correlated (R=0.98, P=0.00, Table 2.7). Correspondence analysis revealed that most of the samples (12 out of 13, outlier: KC14.105) grouped together in the CA plot with *MS4A1* and *CCL21* (Figure 2.3). In the correspondence analysis (CA), Dimension 1 (Dim1) represented up to 94% of the importance with *CD19* and *LTB* contributing the most to Dim1, with Dim2 only representing 4.14% of the variation (Figure 2.3).

2.3.5 Prediction model to discover potential gene markers for *stx2* mRNA abundance

Further analysis using a random forest model classifier based on expressions of four host immune genes *MS4A1, LTB, CCL21, CD19* revealed the accuracy for predicting *stx2* mRNA abundance was 96.5% for the training data and 93.6% for the validation data. The AUC value of 0.908 for the ROC curve also revealed a high accuracy and a robust prediction (Figure 2.4A). As an indicator of *stx2* expression, the prediction accuracy of *MS4A1, LTB, CCL21, CD19* was 47.55%, 45.35%, 41.44%, 36.80%, respectively. Further Boruta analysis also revealed that all four immune genes were attributes for *stx2* expression, with the ranking *MS4A1* > *LTB* = *CD19* > *CCL21* (Figure 2.4B).

2.4 Discussion

This study characterized the abundance, prevalence, and expression of the *stx1* and *stx2* at the recto-anal junction in feedlot steers of three breeds over two consecutive years. Several studies

have quantified the copy number of *stx1* and *stx2* in cattle feces using qPCR, with estimates ranged from 0 to 5.6 $\log_{10}$(gene copies/g) (Imamovic and Muniesa, 2011; Verstraete et al., 2014). Our estimates of the copy number of *stx1* and *stx2* in contents are within these ranges, with 1.24 to 4.13 $\log_{10}$(gene copies/g) (year 1, *stx1*), 0 to 0.45 $\log_{10}$(gene copies/g) (year 2, *stx1*), 0.86 to 5.38 $\log_{10}$(gene copies/g) (year 1, *stx2*), and 4.51 to 5.09 $\log_{10}$(gene copies/g) (year 2, *stx2*). However, there was a markable difference in the copy number of *stx* in tissue samples when compared to RAJ contents. The *stx* genes associated with RAJ tissue samples ranged from 5.62 to 6.07 $\log_{10}$(gene copies/g) (year 1, *stx2*), 6.71 to 6.85 $\log_{10}$(gene copies/g) (year 2, *stx1*) and 5.61 to 5.76 $\log_{10}$(gene copies/g) (year 2, *stx2*). We speculate that the high *stx* copy numbers detected from tissues likely represent the higher possibility of the STEC colonization on RAJ mucosa. Indeed, a previous study has reported that the abundance of *E.coli* O157 strain was inconsistent between RAJ tissues and content samples (Keen et al., 2010), suggesting that *stx*-carrying bacteria were associated with the epithelium of RAJ in the steers in addition to their presence in digesta. Based on our results, digesta samples only present a proportion of the actual STEC that inhabit in the RAJ of cattle, with the higher population directly colonizing epithelial tissue. These results suggest that fecal samples together with rectal mucosa swabs or biopsies would result in a more accurate estimation of *stx* gene abundance in cattle.

Our study further revealed that the abundance and prevalence of the *stx* genes were affected by breed and sampling year, and such effects were *stx* type dependent. However, a previous study found no relationship between cattle breeds and the presence of *stx* at the RAJ (Mir et al., 2016). The inconsistency between our and previous findings may be due to

differences in breed, age (calf (Mir et al., 2016) vs. steer), and diets of the cattle. In this study, Angus, Charolais, and Kinsella Composite breeds were used to examine the abundance and prevalence of *stx* genes, while previous studies collected samples from hybrid Angus-Brahman beef calves (Mir et al., 2016). Steers in our study were fed a high grain diet and slaughtered at similar body weight, but still differed in *stx1* and *stx2* prevalence across breeds, suggesting the highly individualized response to STEC colonization. Therefore, host genetics may alter the gut environment through influences on immunity and the microbiome (Wang et al., 2018), which may influence the prevalence of STEC and the prevalence and abundance of *stx* genes in the samples. The observed differences between sampling years suggest that environmental factors together with host genetics impact the prevalence of *stx* genes in the RAJ of steers. Higher ambient temperatures have been shown to be associated with increased prevalence of both *stx1* and *stx2* in the rectal mucosa of both dairy and beef cattle (Tahamtan et al., 2010; Arumugam et al., 2011). For our study, the average ambient temperatures were similar between the two years (3.25℃ for 2014 vs 5.63℃ for 2015) and as a result it is unlikely to account for the difference in detection of *stx1* and *stx2* between years. Other ecological factors such as seasonality, water and soil sources, and factors associated with farm management may also contribute to varied STEC colonization. Future long term monitoring studies are needed to determine to what extent these environmental factors contribute to the prevalence of both *stx1* and *stx2* in the RAJ of cattle.

Although the presence of both *stx1* and *stx2* genes were detected, only expression of *stx2* was found in the RAJ tissue of beef steers. Severe STEC infections that result in HUS are

mostly associated with *stx2* as its product is 400 times more toxic (as quantified by LD_{50} in mice) than the product of *stx1* (Riley et al., 1983; Wells et al., 1983). *Stx2*-producing *E.coli* strains were reported to be in 71 % (34 out of 48) of children with HUS, while only 40% (4 out of 10) of patients were associated with *stx1*-producing *E.coli* strains (Ludwig et al., 2001). It is noticeable that the prevalence of *stx2* gene expression in steers (8.5% for year 1, 9.7% for year 2) is similar to the reported super shedder rate (~10% (Mattews et al., 2005)), suggesting the expression of *stx2* might be highly correlated with super shedding (SS) and cattle with *stx2* expression might potentially be SS. Interestingly, all *stx2*+ samples were from KC steers in 2014, suggesting KC might be more prominent carriers of STEC and further highlighting the role of breed.

We further speculate that the *stx2*+ cattle may have higher colonization of STEC. As the adherence factor intimin encoded by *eae* gene enables STEC colonization (Farfan and Torres, 2012) and the presence of *eae* is correlated with the formation of attaching and effacing (A/E) lesions (Wieler et al., 1996) by *E. coli* O157:H7 at the RAJ (Sheng et al., 2006), we also assessed the expression of it in this study. The expression of *eae* was detected in 9 out of 131 RAJ tissue samples (Data not shown). Of these, only two samples were *stx2* positive. A previous study isolated 326 *E.coli* strains from 304 fecal samples of clinically healthy wild boars, and found that 10 samples were *eae* positive (Alonso et al., 2017). Presently, only one *stx2*+ *eae*+ *E.coli* strain (*E.coli* O145:H28) associated with HUS in humans has been characterized (Alonso et al., 2017). Although the occurrence of *eae*, alone or in combination with *stx2* is sporadic, diverse *E.coli* serotypes exist in beef cattle and among them certain

serotypes could be potential human pathogens. Compared to previous studies that only reported expressions of *eae* and *stx* from fecal samples, our study is the first to report expressions of these two genes in RAJ mucosa. The detection of *stx*+, *eae*+, and *stx2*+*eae*+ cattle suggests the importance of including all serotypes instead of only *E.coli* O157:H7 for future SS research in practice to the prevention of SS transmission and the mitigation of potential human infections. Future study is needed to isolate *E. coli* serotypes who carry *stx*+, *eae*+, and *stx2*+*eae*+ genes and evaluate their abundances in RAJ and feces of beef steers to verify whether they are SS. Although the abundance of O157 strains was not quantified in this study, our study highlights the importance to use marker genes to assess all STEC populations as opposed to only E. coli O157:H7. In addition to *eae* genes, Enterohemorrhagic *E. coli* autotransporters (Eha) A and B autotransporters that can colonize bovine epithelia are vital adhesin factors in STEC and are more prevalent among STEC strains (97% and 93%, respectively) (Easton et al., 2011). Particularly, *EhaA* promoted adhesion to primary epithelial cells of bovine RAJ and should be explored to identify relationships between *EhaA* and host immunity to gain a fundamental understanding of host-STEC interactions and STEC colonization. Other adhesin factors that play a role in STEC colonization of bovine epithelia such as hemorrhagic coli pili (HCP), EspP rope-like fibers (Farfan and Torres, 2012) should also be explored to identify relationships between STEC adhesin factors and host immune gene expressions.

Previous studies have identified differences in the expression of *MS4A1, CD19, CCL21, LTB* genes at the RAJ of super-shedder vs non-shedders (Corbishley et al., 2014; Wang et al., 2016). These genes are involved in B cell proliferation (Uchida et al., 2004), B cell receptor

signaling pathway (Karnell et al., 2014), and the migration of B cells from bone narrow to lymphoid tissues (Bowman et al., 2000), as well as the induction of the inflammatory response system, respectively (Browning et al., 1995). The observed higher relative expression of *CD19* (a membrane co-receptor found on all B cells) in KC steers and the higher relative expression of *CCL21* in AN and KC than CH in 2015, suggests that expression of this gene in cattle is influenced by breed. Breed-driven gene expression against infections and biological processes have been explored in bovine tissues and cells. Examples include, the reduced expression of the *ALDOA* (Fructose-bisphosphate aldolase A) in the longissimus muscle of Wagyu- as compared to Piedmontese-sired offsprings (Lehnert et al., 2007), and the up-regulation of *CD9* (CD9 antigen) and *BoLA-DQB* (BoLa Class II histocompatibility antigen, DQB*101 beta chain) in the macrophage of Sahiwalas compared to Holstein cattle in response to *Theileria annulate* infection (Glass and Jensen, 2007). In our previous study, the variation in expression of immune genes between SS and NS, could be due to genetic variation (Wang et al., 2016), suggesting future genome-wide association studies (GWAS) are needed to identify the genotypes and/or SNPs responsible for expression of immune genes that could directly or indirectly affect STEC colonization and expression of virulence genes.

Lymphotoxin beta (*LTB*) induces an immune response and is crucial for the initiation of Lymphoid follicle (ILF) development (McDonald et al., 2005). Lymphoid follicles (ILFs) in the bovine rectum are regarded as the reservoir of secretory antibodies in the gut, serving as a frontline defensive system in the gastrointestinal (GI) tract (Tsuji et al., 2008). The positive correlation between *stx2* expression and relative expression of *LTB* suggests that cattle with

higher *stx2* expression have lower *LTB* expression, which may lead to decreased production of lymphotoxin and reduced ILF development in the RAJ. Impaired ILF has been associated with a 10 to 100-fold increase in the colonization of *Enterobacteriaceae* in the ileum of mice (Bouskra et al., 2008), and 100-fold increase in anaerobic bacteria in the small intestine of mice (Fagarasan et al., 2002). Also, a previous study indicated that super-shedders harbor a distinct fecal microbiota compared to non-shedders (Xu et al., 2014). These data suggest that changes in *LTB* expression could lead to impaired ILF function and altered microbiota, which could promote STEC colonization in cattle. Expression of *MS4A1* was negatively correlated with *stx2* expression and *MS4A1* was in the dominant position of *stx2*+ samples from the correspondence analysis, suggesting the vital role of *MS4A1* in regulating *stx2* expression and partially reflecting a strengthened adaptive immunity in *stx2*+ cattle. *MS4A1* encodes CD20 which is expressed from late pro-B cells through memory cells, with its function to enable optimal B cell immune responses against T-independent antigens (Kuijpers et al., 2010). Hence, these data suggest that *MS4A1* is the key gene in connecting *stx2* expression to host adaptive immunity, and their negative correlation suggests the establishment of host recognition mechanisms for *stx2* expression.

To our knowledge, this study is the first to explore whether host gene markers were related to *stx* expression and potential STEC colonization using artificial intelligence-based approaches (Random Forest model and Boruta method). Based on results of mean decrease accuracy in the Random Forest Model and Boruta method and the biological functions of these four immune genes, our results highlight the relationship between host immune genes and *stx2*

expression. Of the genes studied, *MS4A1* was the best predictor of the *stx2* expression and it was in the *stx2*+ sample cluster in the CA map. We used the non-parametric dimensionality reduction method, Isomap, to assess the relationship between the expression of host genes and *stx2*, and results supported the *stx2* expression is closely associated with host gene expression patterns. Isomap was initially developed for computational visual perception (Tenenbaum et al., 2000) and then used to investigate ecosystem crosstalk (Mahecha et al., 2007), human disease phenotypes and gene expression (Dawson et al., 2005). Compared to principal component analysis (PCA), this approach is less restrictive since it does not require any specific distribution (*i.e.* normal distribution) of data (Dawson et al., 2005). The clustering patterns generated by PCA were similar to Isomap results, which could be due to the limited number of genes analyzed. But the Isomap approach is suitable for mammalian studies since interactions among genetics, environment, and microbes are by nature, nonlinear (Nicholson et al., 2004). Regardless, our previous studies have reported 57 differential expressed genes between SS and NS (Wang et al., 2018) and many genes interplay in cattle to affect their immunity and microbiota, and as a consequence the complexity of gene-gene interactions should be taken into account for future studies. Further explorations to investigate more DE genes and their interactions either at the individual or whole transcriptome level could identify and verify the predictiveness of host genes as markers of *stx2* expression. In addition to the genetic background that alters the predictiveness of random forest model, mucosa attached microbes (bacteria and viruses) can also impact host immune gene expressions which should also be considered for the future construction of the prediction model. Our previous study identified

relationships between RAJ mucosa-associated bacteria and expression of 19 out of 57 DE immune genes identified from SS compared to NS (Wang et al., 2018). Although four immune genes were not part of these 19 DE genes, future studies to include the expression of these genes are needed to better understand STEC colonization and its relationship with host immune genes and model construction.

2.5 Conclusion

Taken together, our results revealed that cattle genetic background (breed) and sampling year could affect the abundance and prevalence of STEC *stx1* and *stx2* genes in the RAJ of feedlot cattle. We identified the relationships between *stx2* expression and the expression of host immune genes, and found that *stx2* expression could be driven by expression of particular host immune genes (e.g., *MS4A1*). Our study established a model to correlate host gene expression to *stx2* expression, suggesting that its expression can be driven by the host. Although *Stx* detection from feces is a more direct method, the findings from this study revealed that it may not represent the true population of STEC colonizing the RAJ, which could be influenced by epithelial immunity. Future studies are needed to elucidate mechanisms behind host-STEC interactions by applying methods including genome wide association studies (GWAS) that determine potential genetic variations related to host-STEC interactions and also explore digesta and mucosal attached microbiota variations to develop methods for the potential identification of true STEC in cattle.

2.6 References

Alonso, C. A., Mora, A., Díaz, D., Blanco, M., González-Barrio, D., Ruiz-Fons, F., et al. (2017). Occurrence and characterization of stx and/or eae-positive Escherichia coli isolated from wildlife, including a typical EPEC strain from a wild boar. *Vet. Microbiol.* 207, 69–73. doi: 10.1016/j.vetmic.2017.05.028.

Arumugam, M., Raes, J., Pelletier, E., Paslier, D. L., Yamada, T., Mende, D. R., et al. (2011). Enterotypes of the human gut microbiome. *Nature* 473, 174–180. doi: 10.1038/nature09944.

Bosilevac, J. M., and Koohmaraie, M. (2012). Predicting the Presence of Non-O157 Shiga Toxin-Producing Escherichia coli in Ground Beef by Using Molecular Tests for Shiga Toxins, Intimin, and O Serogroups. *Appl. Environ. Microbiol.* 78, 7152–7155. doi: 10.1128/aem.01508-12.

Bouskra, D., Brézillon, C., Bérard, M., Werts, C., Varona, R., Boneca, I. G., et al. (2008). Lymphoid tissue genesis induced by commensals through NOD1 regulates intestinal homeostasis. *Nature* 456, 507–510. doi: 10.1038/nature07450.

Bowman, E. P., Campbell, J. J., Soler, D., Dong, Z., Manlongat, N., Picarella, D., et al. (2000). Developmental Switches in Chemokine Response Profiles during B Cell Differentiation and Maturation. *J. Exp. Med.* 191, 1303–1318. doi: 10.1084/jem.191.8.1303.

Browning, J. L., Dougas, I., Ngam-ek, A., Bourdon, P. R., Ehrenfels, B. N., Miatkowski, K., et al. (1995). Characterization of surface lymphotoxin forms. Use of specific monoclonal antibodies and soluble receptors. *J. Immunol. (Baltim., Md : 1950)* 154, 33–46.

Cho, S., Fossler, C. P., Diez-Gonzalez, F., Wells, S. J., Hedberg, C. W., Kaneene, J. B., et al. (2007). Cattle-level risk factors associated with fecal shedding of Shiga toxin-encoding bacteria on dairy farms, Minnesota, USA. *Can. J. Vet. Res. Rev. Can. Rech. veterinaire* 73, 151–6.

Corbishley, A., Ahmad, N. I., Hughes, K., Hutchings, M. R., McAteer, S. P., Connelley, T. K., et al. (2014). Strain-Dependent Cellular Immune Responses in Cattle following Escherichia coli O157:H7 Colonization. *Infect Immun* 82, 5117–5131. doi: 10.1128/iai.02462-14.

Davies, D. L., and Bouldin, D. W. (1979). A cluster separation measure. *IEEE Trans. pattern Anal. Mach. Intell.* 1, 224–7.

Dawson, K., Rodriguez, R. L., and Malyj, W. (2005). Sample phenotype clusters in high-density oligonucleotide microarray data sets are revealed using Isomap, a nonlinear algorithm. *BMC Bioinform.* 6, 195. doi: 10.1186/1471-2105-6-195.

Easton, D. M., Totsika, M., Allsopp, L. P., Phan, M.-D., Idris, A., Wurpel, D. J., et al. (2011). Characterization of EhaJ, a New Autotransporter Protein from Enterohemorrhagic and Enteropathogenic Escherichia coli. *Front. Microbiol.* 2, 120. doi: 10.3389/fmicb.2011.00120.

Fagarasan, S., Muramatsu, M., Suzuki, K., Nagaoka, H., Hiai, H., and Honjo, T. (2002). Critical Roles of Activation-Induced Cytidine Deaminase in the Homeostasis of Gut Flora. *Science* 298, 1424–1427. doi: 10.1126/science.1077336.

Farfan, M. J., and Torres, A. G. (2012). Molecular Mechanisms That Mediate Colonization of Shiga Toxin-Producing Escherichia coli Strains. *Infect. Immun.* 80, 903–913. doi: 10.1128/iai.05907-11.

Fuller, C. A., Pellino, C. A., Flagler, M. J., Strasser, J. E., and Weiss, A. A. (2011). Shiga Toxin Subtypes Display Dramatic Differences in Potency. *Infect Immun* 79, 1329–1337. doi: 10.1128/iai.01182-10.

Glass, E. J., and Jensen, K. (2007). Resistance and susceptibility to a protozoan parasite of cattle—Gene expression differences in macrophages from different breeds of cattle. *Vet. Immunol. Immunopathol.* 120, 20–30. doi: 10.1016/j.vetimm.2007.07.013.

Han, H., Guo, X., and Yu, H. (2016). Variable Selection Using Mean Decrease Accuracy and Mean Decrease Gini Based on Random Forest. *2016 7th IEEE Int. Conf. Softw. Eng. Serv. Sci. (ICSESS)*, 219–224. doi: 10.1109/icsess.2016.7883053.

Imamovic, L., and Muniesa, M. (2011). Quantification and Evaluation of Infectivity of Shiga Toxin-Encoding Bacteriophages in Beef and Salad. *Appl. Environ. Microbiol.* 77, 3536–3540. doi: 10.1128/aem.02703-10.

Karmali, MohamedA., Petric, M., Steele, BrianT., and Lim, C. (1983). SPORADIC CASES OF HAEMOLYTIC-URAEMIC SYNDROME ASSOCIATED WITH FAECAL CYTOTOXIN AND CYTOTOXIN-PRODUCING ESCHERICHIA COLI IN STOOLS. *Lancet* 321, 619–620. doi: 10.1016/s0140-6736(83)91795-6.

Karnell, J. L., Dimasi, N., Karnell, F. G., Fleming, R., Kuta, E., Wilson, M., et al. (2014). CD19 and CD32b Differentially Regulate Human B Cell Responsiveness. *J. Immunol.* 192, 1480–1490. doi: 10.4049/jimmunol.1301361.

Keen, J. E., Laegreid, W. W., Chitko-McKown, C. G., Durso, L. M., and Bono, J. L. (2010). Distribution of Shiga-Toxigenic Escherichia coli O157 in the Gastrointestinal Tract of Naturally O157-Shedding Cattle at Necropsy. *Appl. Environ. Microbiol.* 76, 5278–5281. doi: 10.1128/aem.00400-10.

Kuijpers, T. W., Bende, R. J., Baars, P. A., Grummels, A., Derks, I. A. M., Dolman, K. M., et al. (2010). CD20 deficiency in humans results in impaired T cell–independent antibody responses. *J. Clin. Investig.* 120, 214–222. doi: 10.1172/jci40231.

Kursa, M. B., and Rudnicki, W. R. (2010). Feature Selection with the Boruta Package. *J. Stat. Softw.* 36. doi: 10.18637/jss.v036.i11.

Lee, M.-S., and Tesh, V. L. (2019). Roles of Shiga Toxins in Immunopathology. *Toxins* 11, 212. doi: 10.3390/toxins11040212.

Lehnert, S. A., Reverter, A., Byrne, K. A., Wang, Y., Nattrass, G. S., Hudson, N. J., et al. (2007). Gene expression studies of developing bovine longissimusmuscle from two different beef cattle breeds. *BMC Dev. Biol.* 7, 95. doi: 10.1186/1471-213x-7-95.

Li, M., Penner, G. B., Hernandez-Sanabria, E., Oba, M., and Guan, L. L. (2009). Effects of sampling location and time, and host animal on assessment of bacterial diversity and fermentation parameters in the bovine rumen. *J. Appl. Microbiol.* 107, 1924–1934. doi: 10.1111/j.1365-2672.2009.04376.x.

Ludwig, K., Karmali, M. A., Sarkim, V., Bobrowski, C., Petric, M., Karch, H., et al. (2001). Antibody Response to Shiga Toxins Stx2 and Stx1 in Children with Enteropathic Hemolytic-Uremic Syndrome. *J Clin Microbiol* 39, 2272–2279. doi: 10.1128/jcm.39.6.2272-2279.2001.

Mahecha, M. D., Martínez, A., Lischeid, G., and Beck, E. (2007). Nonlinear dimensionality reduction: Alternative ordination approaches for extracting and visualizing biodiversity patterns in tropical montane forest vegetation data. *Ecol. Inform.* 2, 138–149. doi: 10.1016/j.ecoinf.2007.05.002.

Majowicz, S. E., Scallan, E., Jones-Bitton, A., Sargeant, J. M., Stapleton, J., Angulo, F. J., et al. (2014). Global Incidence of Human Shiga Toxin–Producing Escherichia coli Infections and Deaths: A Systematic Review and Knowledge Synthesis. *Foodborne Pathog. Dis.* 11, 447–455. doi: 10.1089/fpd.2013.1704.

MATTHEWS, L., McKENDRICK, I. J., TERNENT, H., GUNN, G. J., SYNGE, B., and WOOLHOUSE, M. E. J. (2005). Super-shedding cattle and the transmission dynamics of Escherichia coli O157. *Epidemiology Infect.* 134, 131–142. doi: 10.1017/s0950268805004590.

McDonald, K. G., McDonough, J. S., and Newberry, R. D. (2005). Adaptive Immune Responses Are Dispensable for Isolated Lymphoid Follicle Formation: Antigen-Naive, Lymphotoxin-Sufficient B Lymphocytes Drive the Formation of Mature Isolated Lymphoid Follicles. *J. Immunol.* 174, 5720–5728. doi: 10.4049/jimmunol.174.9.5720.

Melton-Celsa, A. R. (2014). Shiga Toxin (Stx) Classification, Structure, and Function. *Microbiol Spectr* 2, EHEC-0024-2013.

Mir, R. A., Weppelmann, T. A., Elzo, M., Ahn, S., Driver, J. D., and Jeong, K. C. (2016). Colonization of Beef Cattle by Shiga Toxin-Producing Escherichia coli during the First Year of Life: A Cohort Study. *Plos One* 11, e0148518. doi: 10.1371/journal.pone.0148518.

Nicholson, J. K., Holmes, E., Lindon, J. C., and Wilson, I. D. (2004). The challenges of modeling mammalian biocomplexity. *Nat. Biotechnol.* 22, 1268–1274. doi: 10.1038/nbt1015.

Panel, E. B., Koutsoumanis, K., Allende, A., Alvarez-Ordóñez, A., Bover-Cid, S., Chemaly, M., et al. (2020). Pathogenicity assessment of Shiga toxin-producing Escherichia coli (STEC) and the public health risk posed by contamination of food with STEC. *Efsa J* 18. doi: 10.2903/j.efsa.2020.5967.

Petro, C. D., Trojnar, E., Sinclair, J., Liu, Z.-M., Smith, M., O'Brien, A. D., et al. (2019). Shiga Toxin Type 1a (Stx1a) Reduces the Toxicity of the More Potent Stx2a In Vivo and In Vitro. *Infect Immun* 87. doi: 10.1128/iai.00787-18.

Riley, L. W., Remis, R. S., Helgerson, S. D., McGee, H. B., Wells, J. G., Davis, B. R., et al. (1983). Hemorrhagic Colitis Associated with a Rare Escherichia coli Serotype. *New Engl J Medicine* 308, 681–685. doi: 10.1056/nejm198303243081203.

Scallan, E., Hoekstra, R. M., Angulo, F. J., Tauxe, R. V., Widdowson, M.-A., Roy, S. L., et al. (2011). Foodborne Illness Acquired in the United States—Major Pathogens. *Emerg Infect Dis* 17, 7–15. doi: 10.3201/eid1701.p11101.

Sheng, H., Lim, J. Y., Knecht, H. J., Li, J., and Hovde, C. J. (2006). Role of Escherichia coli O157:H7 Virulence Factors in Colonization at the Bovine Terminal Rectal Mucosa. *Infect Immun* 74, 4685–4693. doi: 10.1128/iai.00406-06.

Tahamtan, Y., Hayati, M., and Namavari, M. (2010). Prevalence and distribution of the stx, stx genes in Shiga toxin producing E. coli (STEC) isolates from cattle. *Iran. J. Microbiol.* 2, 8–13.

Tekaia, F. (2016). Genome Data Exploration Using Correspondence Analysis. *Bioinform. Biol. Insights* 10, BBI.S39614. doi: 10.4137/bbi.s39614.

Tenenbaum, J. B., Silva, V. de, and Langford, J. C. (2000). A Global Geometric Framework for Nonlinear Dimensionality Reduction. *Science* 290, 2319–2323. doi: 10.1126/science.290.5500.2319.

Tsuji, M., Suzuki, K., Kitamura, H., Maruya, M., Kinoshita, K., Ivanov, I. I., et al. (2008). Requirement for Lymphoid Tissue-Inducer Cells in Isolated Follicle Formation and T Cell-Independent Immunoglobulin A Generation in the Gut. *Immunity* 29, 261–271. doi: 10.1016/j.immuni.2008.05.014.

Uchida, J., Lee, Y., Hasegawa, M., Liang, Y., Bradney, A., Oliver, J. A., et al. (2004). Mouse CD20 expression and function. *Int. Immunol.* 16, 119–129. doi: 10.1093/intimm/dxh009.

Verstraete, K., Coillie, E. V., Werbrouck, H., Weyenberg, S. V., Herman, L., Del-Favero, J., et al. (2014). A qPCR Assay to Detect and Quantify Shiga Toxin-Producing E. coli (STEC) in Cattle and on Farms: A Potential Predictive Tool for STEC Culture-Positive Farms. *Toxins* 6, 1201–1221. doi: 10.3390/toxins6041201.

Wang, O., Liang, G., McAllister, T. A., Plastow, G., Stanford, K., Selinger, B., et al. (2016). Comparative Transcriptomic Analysis of Rectal Tissue from Beef Steers Revealed Reduced Host Immunity in Escherichia coli O157:H7 Super-Shedders. *Plos One* 11, e0151284. doi: 10.1371/journal.pone.0151284.

Wang, O., McAllister, T. A., Plastow, G., Stanford, K., Selinger, B., and Guan, L. L. (2018). Interactions of the Hindgut Mucosa-Associated Microbiome with Its Host Regulate Shedding of Escherichia coli O157:H7 by Cattle. *Appl Environ Microb* 84, e01738-17. doi: 10.1128/aem.01738-17.

Wells, J. G., Davis, B. R., Wachsmuth, I. K., Riley, L. W., Remis, R. S., Sokolow, R., et al. (1983). Laboratory investigation of hemorrhagic colitis outbreaks associated with a rare Escherichia coli serotype. *J Clin Microbiol* 18, 512–520. doi: 10.1128/jcm.18.3.512-520.1983.

Wieler, L. H., Vieler, E., Erpenstein, C., Schlapp, T., Steinrück, H., Bauerfeind, R., et al. (1996). Shiga toxin-producing Escherichia coli strains from bovines: association of adhesion with carriage of eae and other genes. *J. Clin. Microbiol.* 34, 2980–2984. doi: 10.1128/jcm.34.12.2980-2984.1996.

Xu, Y., Dugat-Bony, E., Zaheer, R., Selinger, L., Barbieri, R., Munns, K., et al. (2014). Escherichia coli O157:H7 Super-Shedder and Non-Shedder Feedlot Steers Harbour Distinct Fecal Bacterial Communities. *Plos One* 9, e98115. doi: 10.1371/journal.pone.0098115.

Yu, Z., and Morrison, M. (2004). Improved extraction of PCR-quality community DNA from digesta and fecal samples. *Biotechniques* 36, 808–812. doi: 10.2144/04365st04.

2.7 Tables and Figures

Table 2.1. Primer sequences, amplicon sizes, and annealing temperature for qPCR assays

Genes	Oligo sequence (5' to 3')	Amplicon size, bp	Reference	AnnealingTemp, °C
stx1	F: GTCACAGTAACAAACCGTAACA R: TCGTTGACTACTTCTTATCTGGA	95	Jothikumar & Griffiths, 2002	60
stx2	F: ACTCTGACACCATCCTCT R: CACTGTCTGAAACTGCTC	118	He et al, 2020	60
eae	F: TGCTGGCATTTGGTCAGGTC R: CGCTGA(AG)CCCGCACCTAAATTTGC	175	Delmas et al, 2009	60
CCL21	F: GCTATCCTGTTCTCGCCTCG R: ACTGGGCTATGGCCCTTTTG	222	Wang et al, 2016	60
LTB	F: TGGGAAGAGGAGGTCAGTCC R: TAGCTTGCCATAAGTCGGGC	215	Wang et al, 2016	62
CD19	F: CTCCCATACCTCCCTGGTCA R: GCCCATGACCCACATCTCTC	127	Wang et al, 2016	64
MS4A1	F: GCGGAGAAGAACTCCACACA R: GGGTTAGCTCGCTCACAGTT	206	Wang et al, 2016	64
β-actin	F: CTAGGCACCAGGGCGTAATG R: CCACACGGAGCTCGTTGTAG	177	Malmuthuge et al, 2012	60

Table 2.2. The prevalence analysis of *stx1* and *stx2* for samples collected from the rectal tissue and content in 2014 and 2015.

		Year 1 (2014)				Year 2 (2015)			
Sample type	**Breed**	**No. (%)Stx1-positive**	**P value**	**No.(%)Stx2-positive**	**P value**	**No.(%)Stx1-positive**	**P value**	**No.(%)Stx2-positive**	**P value**
Tissue	AN	0 (0) [a]	1	23 (100)	1	24 (100)	1	24 (100)	1
	CH	0 (0)		24 (100)		23 (100)		23 (100)	
	KC	0 (0)		24 (100)		24 (100)		24 (100)	
Content	AN	18 (78)	0.001***	22 (96)	<0.001***	1 (6)	0.069	17 (94)	0.272
	CH	7 (35)		20 (100)		0 (0)		24 (100)	
	KC	6 (27)		4 (18)		4 (17)		24 (100)	

[a] Values presented here were numbers and percentages of Stx-positive samples. Fisher's exact test was used to examine the differential prevalence of *stx1* and *stx2* among three breeds within each sample type. For comparisons, P-values were included along with the level of statistical significance (P≤0.001***).

Table 2.3. Abundance of *stx1* and *stx2* using q-PCR for samples collected from the rectal tissue and content in 2014 and 2015.

Year	Breed	AN		CH		KC		P-Value		
	Type	T	C	T	C	T	C	Breed	Type	Breed*Type
2014	*stx1*	N/D [a]	4.09 (5.20)	N/D	1.73 (5.79)	N/D	1.40 (5.47)	<0.0001 ***	<0.0001 ***	<0.0001 ***
	stx2	6.02 (0.08)	4.92 (1.01)	5.31 (0.05)	5.91 (0.22)	5.70 (0.05)	1.00 (4.65)	<0.0001 ***	<0.0001 ***	<0.0001 ***
2015	*stx1*	6.78 (0.02)	0.25 (1.11)	6.82 (0.03)	N/D	6.76 (0.03)	N/D	0.31	<0.0001 ***	0.28
	stx2	5.70 (0.02)	4.58 (1.58)	5.73 (0.03)	4.91 (0.20)	5.67 (0.03)	5.06 (0.31)	0.17	<0.0001 ***	0.12

[a] The value was presented as Mean (SE) after $\log_{10}$ transformation (gene copy numbers / g sample). T represents tissue samples, C represents contents. For content and tissue samples, the lowest abundance that can be detected corresponds to 200 (2.3 after $\log_{10}$ transformation) gene copies/g and 40 (1.5 after $\log_{10}$ transformation) gene copies/g, respectively. Therefore, *stx* gene abundance that lower than 2.3 $\log_{10}$(gene copies/g) and 1.5 $\log_{10}$(gene copies/g) for content and tissue samples was defined as 'underdetermined' ('N/D') which is assumed to be '0' in our analysis, respectively. For comparisons among different factors and among interaction effects, P-values were included along with the level of statistical significance (P≤0.001***).

Table 2.4. Profiles of positive *stx2* expression samples including sample ID, year, and breed.

Sample ID	Breed	Stx2RNA (log_{10} transformation)
2014-104	KC	5.07
2014-105	KC	5.05
2014-106	KC	5.01
2014-107	KC	4.78
2014-109	KC	4.63
2014-211	KC	5.11
2015-104	KC	5.13
2015-105	KC	5.34
2015-106	KC	5.29
2015-401	AN	5.11
2015-402	AN	5.23
2015-502	CH	5.04
2015-602	CH	4.93

Table 2.5. Quantification for relative expressions of four host gene among breeds using qRT-PCR for rectal tissue samples collected in 2014 and 2015.

Year	Immune genes	AN	CH	KC	P-Value
2014	*MS4A1*	2.80 (0.36) [a]	3.42 (0.29)	3.76 (0.44)	0.13
	CD19	-0.14 (0.38)	-0.32 (0.52)	-0.06 (0.35)	0.91
	CCL21	3.88 (0.45)	4.64 (0.35)	4.82 (0.46)	0.26
	LTB	-0.96 (0.48)	-0.97 (0.60)	-1.30 (0.44)	0.86
2015	*MS4A1*	3.76 (0.27)	3.65 (0.25)	4.26 (0.30)	0.26
	CD19	3.61 (0.28)	3.50 (0.29)	4.51 (0.24)	0.02*
	CCL21	5.94 (0.25)	4.90 (0.21)	5.87 (0.23)	0.0035***
	LTB	4.31 (0.44)	4.36 (0.40)	5.47 (0.36)	0.07

[a] The value was presented as Mean (SE) of ΔCq value that was calculated from each tissue sample under different year and breeds. For comparisons among different factors and interaction effects, P-values were included with the level of statistical significance (P<0.05*, P≤0.001***).

Table 2.6. Expression differences for four host genes between Stx2+ and Stx2- samples using non-parametric Mann-Whitney U test.

Immune genes	Mean		Z-score	P-Value
	Stx2-	Stx2+		
MS4A1	3.65	3.44	0.92	0.36
CD19	1.90	1.54	0.49	0.62
CCL21	5.02	5.04	0.08	0.94
LTB	1.90	1.30	0.61	0.54

For comparisons between Stx2+ and Stx2- group, P-values were included with the level of statistical significance ($P<0.05$*, $P\leq0.001$***).

Table 2.7. Correlation analysis among relative expressions of host genes and *stx2* expression among Stx2+ samples.

		Stx2RNA	*MS4A1*	*CD19*	*CCL21*	*LTB*
Stx2RNA	R-Value	1.00	-0.56	0.51	-0.44	0.60
	P-Value	0.00	0.05*	0.08	0.13	0.03*
MS4A1	R-Value		1.00	-0.55	0.39	-0.56
	P-Value		0.00	0.05*	0.19	0.05*
CD19	R-Value			1.00	0.19	0.98
	P-Value			0.00	0.53	0.00***
CCL21	R-Value				1.00	0.09
	P-Value				0.00	0.78
LTB	R-Value					1.00
	P-Value					0.00

R-value was defined as the correlation coefficient ranged from -1 to 1. For correlations with different genes, P-values were included along with the level of statistical significance (P<0.05*, P<0.01**, P≤0.001***).

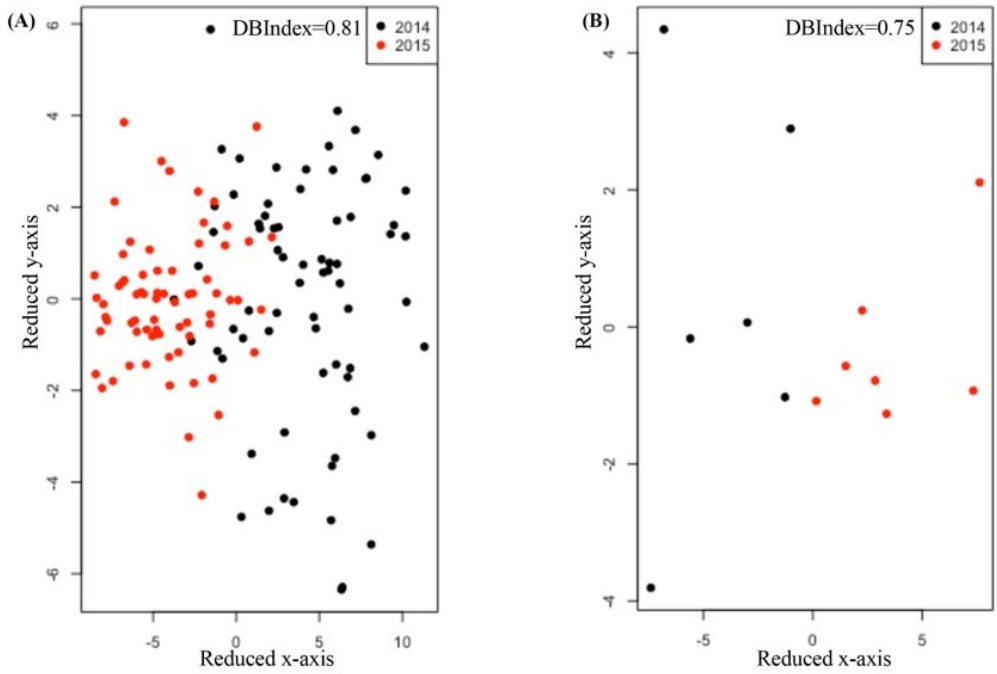


Figure 2.1. Comparisons of host gene expression patterns using non-parametric method Isomap and DBIndex value for sampling year effect.

Comparisons were performed among all samples (A) as well as among Stx2+ samples (B). Black dots and red dots refer to samples collected in 2014 and 2015, respectively. DBIndex value was shown on the right corner of each figure. The lower DBIndex value, the well-separated cluster pattern.

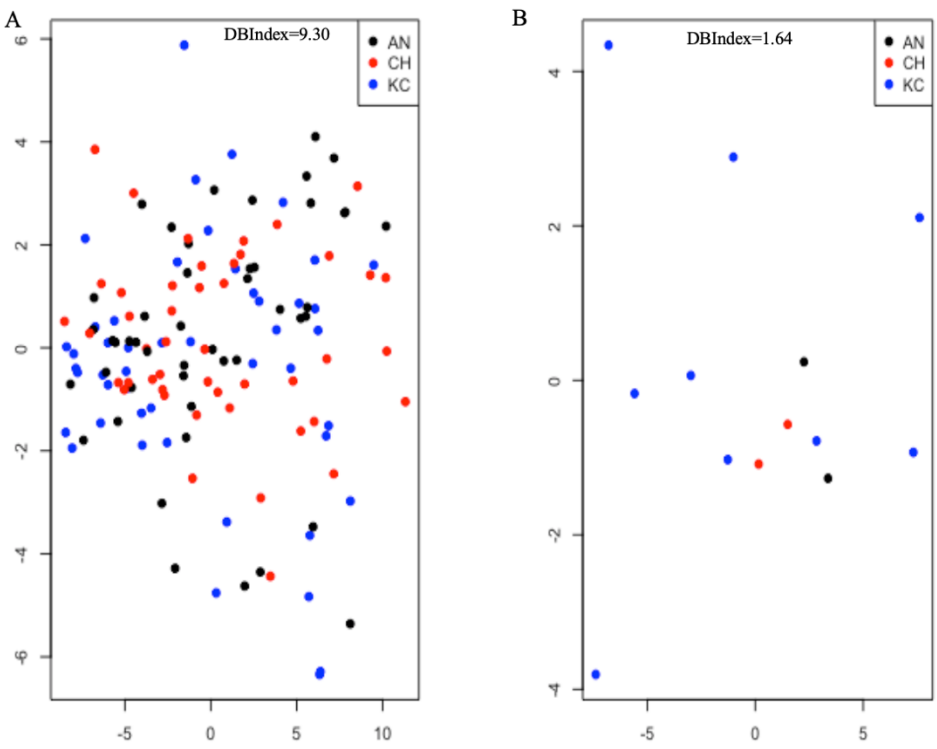


Figure 2.2. Comparisons of host gene expression patterns using non-parametric method Isomap and DBIndex value for breed effect.

Comparisons were performed among all samples (A) as well as among Stx2+ samples (B). Each dot represents a samples with black, red, and blue dots representing Angus, Charolais, Kinsella Composite breed. DBIndex value was shown on the right corner of each figure. The lower DBIndex value, the well-separated cluster pattern.

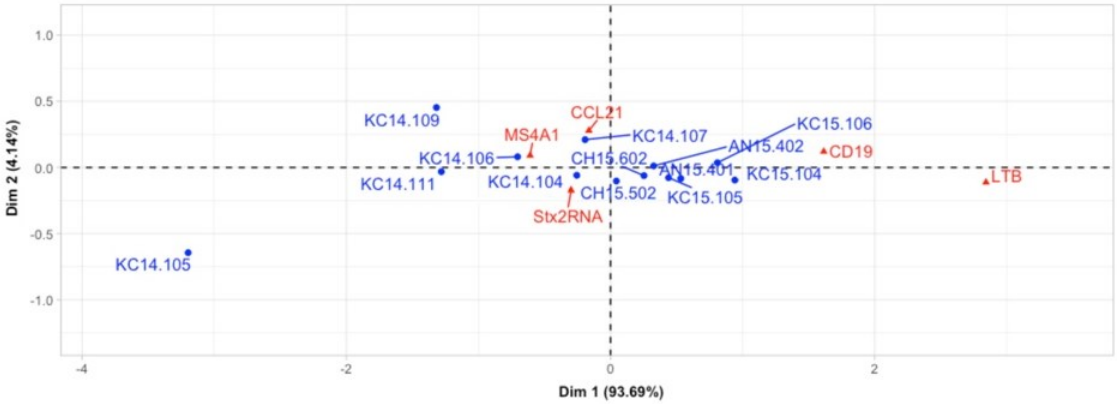


Figure 2.3. Assessment of associations between host immune gene expressions and Stx2+ samples using correspondence analysis.

Red triangles and blue dots refer to host genes and Stx2+ samples, respectively. For example, "AN14.105" means the number of this sample is 105, breed is Angus and the sample was collected in 2014.

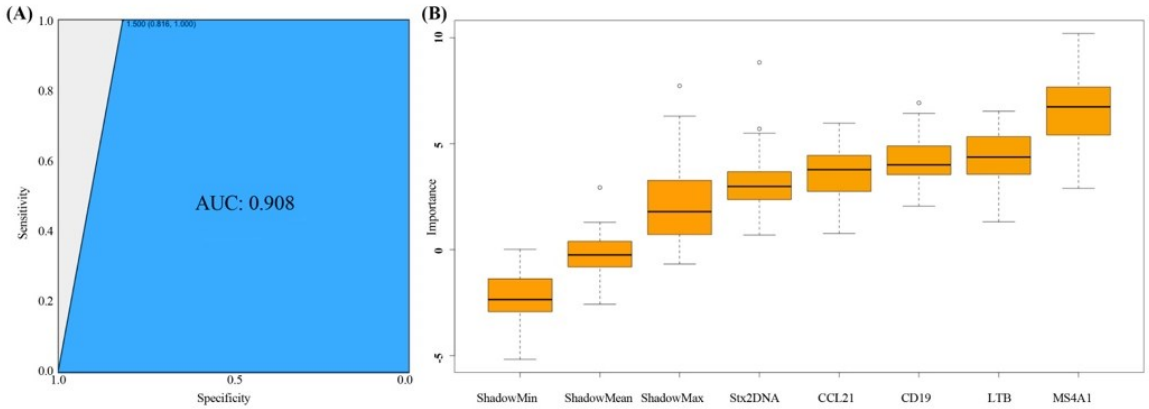


Figure 2.4. Assessment of Random Forest model using ROC curve and Burota method.

(A) Assessment of classification performance of random forest model using area under ROC (AUC). Sensitivity (y-axis) represents the fraction of samples with positive *Stx2* expression that the test correctly identifies as positive. Specificity (x-axis) represents the fraction of samples without *Stx2* expression that the test correctly identifies as negative.

(B) Rank of host immune genes as markers for *Stx2* expression prediction using Boruta method.

Chapter 3. Microbial interaction-driven community differences as revealed by network analysis[2]

3.1 Introduction

Assessing microbial profiles using diversity, composition, and abundance measurements are broadly adopted approaches that have been widely applied to determine the role of the microbiome in host health. For example, the lower richness of human fecal microbiota is associated with dyslipidemia and insulin resistance leading to obesity (Chatelier et al., 2013), and higher evenness of milk/teat microbiota is associated with dairy cow mastitis (Derakhshani et al., 2020). In addition to these quantitative measures, microbial taxa interact within ecological niches, forming "micro-communities" that may function collectively (Hirano and Takemoto, 2019). Such microbial interactions can be influenced by the host, which in turn affects the host's physiological activities. Hence, microbial interactions, together with microbial-host interactions, are critical for the establishment, maintenance, and function of microbiota (Zhou et al., 2010).

The traditional and commonly used approach to identify microbial interactions is the construction of microbial co-occurrence networks using correlation-based methods (Pearson's or Spearman's correlation coefficient) (Zhang and Horvath, 2005; Hirano and Takemoto, 2019; Derakhshani et al., 2020; Guo et al., 2020). These methods are prone to detecting spurious correlations among low abundance taxa (Xia, 2020) and can lead to an ill-defined understanding

[2] Chapter 3 was published as a part of a paper: Pan Z, Chen Y, Zhou M, McAllister TA, Guan LL. Microbial interaction-driven community differences as revealed by network analysis. Comput Struct Biotechnology J. 2021;19:6000–8.

of microbial interactions. These conventional methods also adopt subjective thresholds to define significant microbial interactions largely based on known biological information, while appropriate thresholds are hard to select, particularly for less studied and low abundant microorganisms (Zhou et al., 2010; Deng et al., 2012). Hence, effective ways to construct microbial co-occurrence networks are needed to enable an in-depth understanding of structural differences in microbial communities and how these could affect host-microbial interactions.

Shiga toxin-producing *Escherichia coli* (STEC) causes foodborne diseases that can lead to severe human infections (*i.e.* bloody diarrhea, hemolytic uremic syndrome) (Karmali et al., 1983). Cattle are the main asymptomatic carriers of STEC with the rectal-anal junction (RAJ) being the main colonization site (Sheng et al., 2006). As a result, Cattle can shed STEC to the surrounding environment and therefore promote cattle-cattle/human transmission (Donkersgoed et al., 2001; Munns et al., 2015). Our recent research has reported that host (*i.e.* host genetics, immunity, microRNAs (Wang et al., 2016, 2017, 2018, 2021)) and microbial interactions may play a role in affecting STEC colonization in cattle. In addition, Shiga toxins (*stx*) are major virulence factors in STEC with prototype toxins being designated as Shiga toxin 1 (*stx1*) and Shiga toxin 2 (*stx2*) (Fraser et al., 2004). Our recent study revealed that *stx2* expression was associated with host immune gene expression and potential STEC colonization (Pan et al., 2021). Although distinctive fecal microbial communities have been reported in super-shedder (cattle shed > 10^4 STEC in feces per gram, SS) compared to non-shedders (Xu et al., 2014), little knowledge is derived from *stx2* in STEC and its relationships with RAJ microbiota and microbial interactions. We speculated that microbial interactions together with

the diversity and composition of RAJ microbiota are associated with *stx2* expression and STEC colonization in cattle. Hence, this study aimed to assess the rectal microbial communities and interactions in response to *stx2* expression in STEC at the RAJ, and the role of microbiota divergence in abundance among microbial interactions using the (Zhou et al., 2011) (RMT)-based method and within-/among- module connectivity (Guimerà and Amaral, 2005). We aimed to identify keystone taxa and low abundant taxa contributing to microbial interaction networks and structural stability to better understand the role of rectal microbiota in *stx2* expression and potential STEC colonization in beef cattle.

3.2 Materials and methods

3.2.1 Animal study and sample collection

The animal trial and identification of *stx2* gene abundance and expression were described in Chapter 2. Briefly, ten cm^2 recto-anal junction (RAJ) tissue and 10 mL rectal contents were collected from a total of 143 feedlot cattle (585.84 $\pm$ 64.99 kg) within 30 min after slaughter at a federally approved abattoir.

Extraction of DNA and RNA from powdered tissue, and the assessment of DNA and RNA quality were described in Chapter 2. The detection of copy numbers of *stx2* gene from DNA using PCR and *stx2* transcript from RNA using RT-qPCR are also described in Chapter 2. Both PCR and RT-qPCR were conducted in triplicate for each sample and followed the same thermal program on a StepOnePlus™ Real-Time PCR System (Applied Biosystems, Foster City, CA, USA): one cycle at 95°C for 20 s followed by 40 cycles of 3 s at 95°C, and 30 s at 60°C. Twelve

rectal digesta samples collected from steers whose mucosal samples were confirmed to possess *stx2* gene without expression (defined as Stx2- group, n=6) in the bacteria, and those with *stx2* that was expressed (defined as Stx2+ group, n=6) in STEC were selected with minimized differences in body weight and age at slaughter between two groups based on one-way ANOVA ($P_{body\ weight}$ =0.07, $P_{slaughter\ age}$ =0.30, Table 3.1).

3.2.2 Amplicon sequencing and microbial community analysis

Total genomic DNA was extracted from frozen rectal digesta samples using repeated bead beating and a column (RBBC) method, and purified using the QIAmp Stool Mini Kit (Qiagen, Germany) following the manufacturer's protocols. The concentration and quality of DNA were further determined using the NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific, USA).

To generate the rectum bacterial compositional profiles, the bacterial V1-V3 region of the 16S rRNA gene was amplified using bacterial primers Bac9F (5'-GAGTTTGATCMTGGCTCAG-3') and Ba515Rmod1 (5'-CCGCGGCKGCTGGCAC-3') (Wang et al., 2018). The PCR amplification products were verified using agarose gel (2%) electrophoresis. Two-step PCR was performed for PCR amplicon generation and barcoding. In detail, the PCR was conducted with the following thermal program of an initial denaturation at 94°C for 2 mins, followed by 33 cycles of 94°C for 30s, annealing at 58°C for 30s, elongation at 72°C for 30s, and at last a final elongation at 72°C for 7 mins. Furthermore, the second PCR was performed using amplicons produced in the first step for barcoding with the following

program: initial denaturation at 95°C for 10 mins, followed by 15 cycles of 95°C for 30s, 60°C for 30s, 72°C for 60s, and a final elongation at 72°C for 3 mins.

All amplicon libraries were sequenced using an Illumina MiSeq pair-end 300 bp platform at Centre d'expertise et de services Génome Québec (Quebec, Canada). The raw sequence data were assigned to each sample according to the corresponding barcode and were processed using QIIME2 (Version 2019.10) (Bolyen et al., 2019). Quality control, denoising, removal of chimeric sequences, and generation of amplicon sequencing variants (ASVs) were performed using the QIIME2 plugin DADA2 (Callahan et al., 2016). Taxonomic classification was performed in QIIME2 using a taxonomic classifier with the SILVA database (version 132) as the reference. The Good's coverage index was used to evaluate the adequacy of sequencing depth to generate bacterial profiles in each sample. For diversity analyses, alpha diversity was estimated using Shannon (evenness) and Chao1 (richness) indices. Beta diversity was evaluated based on Weighted Unifrac distance using phylogenetic distances across identified taxa in a phylogenetic tree and the abundance of each feature to determine the similarity between Stx2- vs. Stx2+ groups. All diversity metrics were calculated using scripts implemented in QIIME2. The visualization of alpha- and beta- diversity was performed using the ggplot2 package in R. Differentially abundant (DA) genera between Stx2- vs. Stx2+ groups were identified using the DESeq2 package in R. The cut-off of DA genera and group-specific taxa is a relative abundance >0.1% and presence in at least two out of six animals in each group.

3.2.3 Construction of microbial co-occurrence networks

Random matrix theory (RMT)-based method was employed to construct the microbial co-occurrence network to identify microbial interactions using molecular ecological network analysis (MENA) (http://ieg4.rccc.ou.edu/mena) (Deng et al., 2012). Briefly, the absolute abundance of microbial genera data was uploaded. The absolute abundance microbial genera dataset was appropriately standardized and the pairwise Pearson correlation suggested by the author's manual was employed to generate correlation coefficients and a similarity matrix (Deng et al., 2012). The similarity matrix was subsequently transformed into an adjacency matrix by applying the automatic generated threshold to the correlation values based on the RMT approach (Deng et al., 2012). In this study, a connection stands for a strong (Pearson's $r > 0.85$) and significant ($P < 0.01$) correlation. The visualization of the co-occurrence network was performed using gephi (0.9.2).

3.2.4 Characterization of topological properties of co-occurrence networks

The modularity of each network was estimated using MENA along with the evaluation of other topological properties, including the clustering coefficient (Watts and Strogatz, 1998; Ravasz et al., 2002), average path length (Albert and Barabási, 2002), graph density (Wolfe, 1997), and average degree (Albert and Barabási, 2002). Random networks and power-law distribution assessments were generated to evaluate whether empirical networks were prone to error and to identify microbiota interactions that were due to non-random patterns that represent the empirical structure of microbial communities (Horner-Devine et al., 2007; Barberán et al., 2012). Random networks were constructed based on the Maslov-Sneppen method (Maslov and

Sneppen, 2002) using MENA, which kept number of nodes (microbial taxa) and edges (connections) unchanged, but rewired positions of all links in the network.

From the network modularity perspective, taxa could be classified into network hubs, module hubs, connectors that represent generalists in the community, and peripherals that represent specialists in the community. Generalists refer to taxa that are highly connected with others both within and among modules (network hubs), within a module (module hubs), and among different modules within a network (connectors). Specialists represent peripheral taxa that interact less with other taxa (including a node that is only connected within a module or at least 60% links within the module) (Guimerà and Amaral, 2005). Within-module connectivity (Zi) and among-module connectivity (Pi) were computed based on the following algorithm (Guimerà and Amaral, 2005):

$$Z_i = \frac{k_i - \bar{k}_{Si}}{\sigma_{ksi}}$$

where k_i is the number of links of node i to other nodes in its module; s_i, $\bar{k}_{Si}$ is the average of k across all nodes in s_i; and σ_{ksi} represents the standard deviation of k in s_i. Hence, Zi (Within-module connectivity) characterizes to what extent node i is connected to others in its module.

$$P_i = 1 - \sum_{s=1}^{N_M} \left(\frac{k_{is}}{k_i}\right)^2$$

where k_{is} is the number of links of nodes i to nodes in module s; and k_i is the total degree of node i within a network. Therefore, Pi (among-module connectivity) of a node stands for evenly distribution of links among all modules if its value is close to 1, and 0 if all of its links are within its own module.

Zi and Pi were characterized using MENA with the classification as follows: network hubs (Zi>2.5; Pi>0.62), module hubs (Zi >2.5; Pi<0.62), connectors (Zi<2.5; Pi>0.62) and peripherals (Zi<2.5; Pi<0.62) (Guimerà and Amaral, 2005). The thresholds for classifying nodes into the aforementioned four roles in the network were determined by both heuristic determinations and the concept of 'basin of attraction' (Guimerà and Amaral, 2005). Briefly, Zi and Pi were both computed for each node in the network, and density plots were adopted to visualize the gradient of the value of Zi and Pi that can 'flow' to the local minimum (termed as 'basin of attraction'). In other words, the region of the space that 'flow' toward a certain minimum value is therefore regarded as the optimal threshold for Zi and Pi, being 2.5 and 0.62, respectively (Guimerà and Amaral, 2005).

Network stability is a critical component that tests if a network is resilient to perturbations sourced from external factors other than interactions (Hicklin, 2004; Csermely, 2009). Here, natural connectivity (Jun et al., 2010) was introduced to describe network stability differences in response to *stx2* expression in STEC. The estimation of natural connectivity was based on the following algorithm:

$$\mathrm{ave}(\lambda) = \ln\left(\frac{1}{N}\sum_{i=1}^{N} e^{\lambda_i}\right)$$

where $\mathrm{ave}(\lambda)$ is the natural connectivity, N is the number of nodes in the network, λ_i is the eigenvalue of the adjacency matrix. A total of 137 nodes representing 80% of total nodes were randomly removed from the adjacency matrix and λ_i and $\mathrm{ave}(\lambda)$ were re-calculated after each removal. The visualization of the natural connectivity was performed using the ggplot2 package in R.

The identification of potential hub microbial taxa in the network was validated using Lasso regression in R package 'glmnet'. Twelve samples were regarded as the training data with 9 samples (6 Stx2+; 3 Stx2-) as the external test data. The accuracy rate (the number of samples recognized correctly / total number of samples) was estimated to determine the model classification performance.

3.2.5 Data availability

All the sequencing data used in the current study have been submitted to NCBI Sequence Read Archive (SRA) under the accession numbers from SRR14769039 to SRR14769050 (BioProject ID PRJNA736180).

3.3 Results

3.3.1 Taxonomic assessment of the RAJ content-associated microbiota

A total of 179,917 filtered pair-end reads were generated with 14,993 ± 437 (mean ± SE) sequences per sample, and a total of 2,798 amplicon sequence variants (ASVs) were identified ranging from 174 to 275 (Table 3.2). The Good's coverage was > 99.9 % for all samples, indicating adequate sequencing depth.

From all samples, more than 99% of reads were classified into 13 phyla, with Firmicutes (72.7 ± 2.0 %) and Bacteroidetes (24.6 ± 1.9 %) being the most predominant (relative abundance > 10%, Table 3.3). At the family level, 59 families were identified with 7 unclassified and 52 classified, of which *Ruminococcaceae* (47.2 ± 1.5 %), *Lachnospiraceae* (9.2 ± 1.2%), and *Prevotellaceae* (8.8 ± 1.1%) were the most predominant families. At the

genus level, 154 taxa were identified (20 unclassified; 134 classified), of which *Ruminococcaceae* UCG-005 (23.1 ± 2.1 %) and *coprostanoligenes* group (9.1 ± 0.7%) from *Ruminococcaceae* family and *Christensenellaceae* R-7 group (7.3 ± 0.7%) from *Christensenellaceae* family were the most abundant genera. The most frequently detected genera (average relative abundance >0.5%) belonged to Firmicutes (19 out of 27, 4 unclassified, 15 classified) and Bacteroidetes (8 out of 27, 1 unclassified, 8 classified), respectively.

There were 52 to 75 genera identified from each sample with 15 core genera shared by all 12 samples (Figure 3.1A, Table 3.4). Specifically, 30.7% (4 out of 13 phyla, Firmicutes, Bacteroidetes, Actinobacteria, Proteobacteria), 18.6% (11 out of 59 families, 6 out of 11 belonged to Firmicutes; 5 out of 11 belonged to Bacteroidetes;1) and 9.7% (15 out of 154 genera, 9 out of 15 belonged to Firmicutes; 6 out of 15 genera belonged to Bacteroidetes; Figure 3.1A, Table 3.4) were present in all samples. In addition, twenty-four genera were Stx2- specific and 13 genera were Stx2+ specific with 66 genera shared by both groups (Figure 3.1B). Genera belonging to Firmicutes were the most predominant taxa in Stx2- (10 out of 24) and Stx2+ (4 out of 13) groups (Table 3.5).

3.3.2 Comparable diversity and composition of RAJ content-associated microbiota between Stx2- and Stx2+ groups

Neither Shannon nor Chao1 indices differed (Kruskal Wallis test, $P > 0.05$, Figure 3.2A) between the bacterial communities from Stx2- and Stx2+ groups. Further comparison of the similarity of microbial communities between the two groups using ANOSIM (Analysis of similarities) revealed no clustering patterns ($P = 0.52$, Figure 3.2B) at the phylum ($P > 0.5$,

Figure 3.3), family, or genus (Both P > 0.1) levels. No differentially abundant taxa at the phylum, family, or genus level were identified between Stx2- and Stx2+ groups.

3.3.3 General co-occurrence patterns in each co-occurrence network

The non-random co-occurrence patterns were observed based on the significant power-law distribution in each group ($R^2_{Stx2+} = 0.98$, $P<0.05$; $R^2_{Stx2-} = 0.97$, $P < 0.01$) as well as the greater value of structural properties in the empirical as compared to the random networks (Table 3.6). Hence, non-random empirical co-occurrence networks were established to uncover real-world microbial interactions. The empirical network consisted of 86 nodes (genera) with 322 edges (a mean of 7.49 edges per node) for Stx2- (Table 3.6, Figure 3.4A), and 77 nodes with 243 edges for Stx2+ (Table 3.6, Figure 3.4B). The average network distance between all paired nodes (average path length, APL) was 2.84 edges with a diameter (longest distance) of 8 edges in the Stx2- network and 2.86 edges with a diameter of 7 edges in the Stx2+ network (Table 3.6). The clustering coefficient (the degree to which nodes tend to cluster together) was 0.32 for the Stx2- network and 0.33 for the Stx2+ network (Table 3.6). The modularity index was 0.44 for the Stx2- network and 0.46 for the Stx2+ network (values >0.4 suggest that the network has a modular structure (Newman, 2006)) (Table 3.6).

3.3.4 Identification of keystone taxa and their associations with different co-occurrence patterns

High modularity was observed for both networks with 7 and 6 observed modules in the Stx2- and Stx2+ co-occurrence networks, respectively (Figure 3.4A,B) with no network hubs being identified. More than 80% of nodes (Stx2-: 74 out of 86, 86.0%; Stx2+: 62 out of 77, 80.6%)

with an abundance > 0.2% were classified as peripherals (Figure 3.5A). Only 14 % (connectors: 11 out of 86; module hubs: 1 out of 86; Stx2- group) and 19.4% (connectors: 14 out of 77; module hubs: 1 out of 77; Stx2+ group) of taxa from Stx2- and Stx2+, respectively, were designated as specialists (Figure 3.5A). Specifically, more than 85% (23 out of 27) of connectors had a lower abundance (< 0.2%) with a high Pi value (0.62 ~ 0.77) (Figure 3.5B).

Furthermore, generalists formed differential clustering patterns between Stx2- and Stx2+ networks. Generalists were evenly distributed across identified ecological clusters in the Stx2- network. Four out of fourteen generalists belonged to module 3, whereas other generalists were in module 1 (3 out of 11), module 2 (2 out of 11), module 4 (1 out of 11), module 5 (1 out of 11) and module 6 (1 out of 11). However, in the Stx2+ network, generalists were not evenly distributed with 7, 4, 3, and 1 out of 14 belonging to modules 1, 3, 2, and 4, respectively.

Moreover, decreased stability of the network was evidenced by reduced natural connectivity in the Stx2+ network in comparison to the Stx2- network (Figure 3.6). The natural connectivity that supported network stability in each group gradually decreased with increasing removal of nodes, while natural connectivity in the Stx2+ network always being lower than that in the Stx2- network, regardless of the number of nodes removed from each network (Figure 3.6).

3.3.5 Group-specific taxa as keystone taxa in microbial interactions

More than 50% of the group-specific genera (8 out of 12 in Stx2-; 6 out of 12 in Stx2+, Table 3.7) belonged to generalists with a lower abundance (relative abundance < 0.2%). Among generalists in the Stx2- network included; *Parvibacter*, *Candidatus saccharimonas*,

Acetitomaculum, as well as unknown genera within the Bacteroidetes, Proteobacteria, Tenericutes, Cyanobacteria, and *Veillonellaceae*. Among generalists in the Stx2+ network, *Prevotellaceae* Ga6A1 group, *Flexilinea, Ruminiclostridium* 9, *Ruminococcus* 1, *Acetobacter*, and *Streptomyces* were group-specific genera in microbial communities in the Stx2+ group. All the core microbes shared by two groups were regarded as specialists that were poorly connected with other taxa within networks.

3.4 Discussion

This study assessed microbial interactions in response to *stx2* expression in STEC at RAJ in beef steers, shedding light on microbial mechanisms regulating STEC colonization as well as providing novel approaches to understanding differences in microbial communities. The compositional profiles of microbial communities identified in our study were comparable to those identified from the rectum content of dairy cattle (Mao et al., 2015), and fecal microbiota of beef cattle (Shanks et al., 2011; Xu et al., 2014) with the most dominant phyla being Firmicutes and Bacteroidetes with an accumulative relative abundance of up to 94.1%, 89.5%, and 80.6%, respectively. The proportion of three main families **(***Ruminococcaceae* (47.2 ± 1.5 %), *Lachnospiraceae* (9.2 ± 1.2%), and *Prevotellaceae* (8.8 ± 1.1%)) were also similar to the bacterial taxa identified in a study where cattle were shedding > 10^4 cfu/g of *E. coli* O157 in that 31.8%, 10.5%, 9.0% of total reads from fecal microbiota were assigned to *Ruminococcaceae, Prevotellaceae,* and *Lachnospiraceae,* respectively (Xu et al., 2014). Although all cattle were raised under the same high-grain diet and management conditions and

were similar in age and body weight, individual variations among microbial communities were still observed including the relative abundance of each microbial community and the proportion of predominant taxa shared by each microbial community. Variation in the microbial fecal profile of individual cattle has been previously reported (Xu et al., 2014), a finding that has been attributed to differences in age, weight, and diet. As these variables were relatively consistent in our study, other host factors along with microbial crosstalk within the microbiota might be the main drivers of individualized microbial composition.

Despite highly similar microbial profiles including comparable alpha and beta diversity between Stx2- and Stx2+ groups, microbial interactions within the microbiota varied. Microbes can interact with each other for the purpose of co-evolution (Braga et al., 2016), leading to adaptation and specialization (Braga et al., 2016) of certain microbial taxa, which promote future alterations in the microbial community. To date, the widely used approaches to study the microbial interactions are based on network construction mainly using correlation-based (Pearson's and Spearman correlation coefficient) and maximal information coefficient (MIC) methods that have proven to be less useful in inferring microbial ecological networks assessed by area under the precision-recall curves (AUPR) (Hirano and Takemoto, 2019). These metrics (*i.e.,* correlation-based approaches, MIC) are less applicable to microbial compositional data as the assumption of independent variables can be satisfied, leading to the generation of spurious correlations (Aitchison, 1981; Hirano and Takemoto, 2019), and therefore are less powerful for inferring real microbial interactions. To overcome such limitations in this study, the similarity matrix was first established using Pearson's correlations and then the RMT- based approach

was employed to determine the reference point which enabled automatic threshold selection and minimized noise. The RMT-based approach was established based on two universal roles of random matrix theory: the distribution of two nearest eigenvalues which follow Gaussian orthogonal ensemble (GOE) statistics if correlations exist (a true correlation will follow GOE distribution in RMT theory), while it follows Poisson distribution if there is no correlation (Luo et al., 2007). Particularly the transition between GOE and Poisson distribution serves as the reference point to distinguish non-random relationships (that is true correlations) in the data matrix from background noise (Luo et al., 2007). In other words, a correlation refers to the GOE distribution while non-correlation represents the Poisson distribution. The transition point from GOE to Poisson distribution is the reference point for the automatic generated threshold used for the construction of the microbial network and the identification of meaningful correlations from the noise (*i.e.* certain correlations might not follow GOE and therefore could be noise or spurious correlations). Hence, compared to studies that only adopt correlations (Pearson's and Spearman correlation coefficient) for constructing microbial networks, the RMT-based approach is more effective at identifying true interactions within the response networks to *stx2* expression.

Comparisons of random and empirical networks based on topological properties, confirmed constructed networks were effective for investigating interactions between *stx2* expression and microbial communities. Particularly, modules (clusters) were powerful topological features to reflect network differences, referring to the fundamental units whose constituent elements (nodes) are functionally similar in terms of specific chemical, and

biological processes (Lecca and Re, 2015). Previous explorations in soil-microbial interactions revealed that clusters in the network have specific and different functions that enable microbes to respond to different soil conditions (Jiao et al., 2016; Shi et al., 2020). For instance, three major modules with diverse functions (electron-transfer, biogeochemical C- and N- cycles, organic contaminant degradation) were characterized for microbial communities from soil contaminated with oil (Jiao et al., 2016), Clusters were shown to be crucial components of microbial communities in the network and necessary to develop an understanding of modularity in networks and microbial interactions. The Stx2- and Stx2+ groups formed comparable numbers of clusters while clusters in each group included different taxa, suggesting functions of clusters in each group differed in their interactions in response to *stx2* expression in STEC at the RAJ. Regardless, the highly connected genera among densely connected clusters of nodes (that is, modules) were observed in both Stx2- and Stx2+ groups resulting in the formation of 'small world' topologies, indicative of empirical networks that are more clustered than random networks.

Identifications of generalists also furthered the understanding of microbial community structure and differential microbial interactions, which play an empirical rather than theoretical role among microbial interactions. Particularly, identified generalists in our study are group-specific taxa, which are at low relative abundance but might play unique roles such as regulating microbial interactions within each network through nutritional supplementation or competition. *E. coli* O157:H7 can utilize ethanolamine as the free nitrogen source in the bovine small intestine, thus the presence or absence of ethanolamine utilizing bacteria could be a contributing

factor to diverse microbial communities (Bertin et al., 2011; Xu et al., 2014). In our study, *Streptomyces* identified as Stx2+ specific genera were capable of metabolizing ethanolamine (Krysenko et al., 2019), which might generate a niche for the survival of bacteria that express Stx2. However, not all group-specific generalists in microbial communities play a role in nutrient supplementation for STEC. For instance, *Acetobacter* (a group-specific generalist from the Stx2+ group) is capable of converting ethanol to acetic acid which can inhibit STEC (Han et al., 1992). A similar case is also observed in Stx2- group that *Acetitomaculum* (a group-specific generalist from Stx2- group), an acetogenic bacteria (Van et al., 1998), may also suppress STEC. It is noticeable that more than 85% (23 out of 27) of generalists were low abundant taxa with an average abundance <0.2%, highlighting the irreplaceable role of rare taxa in microbial interactions. For example, rare *Methanotrophs* (*i.e. Methylocaldum*) act as 'primary producers' in methane-driven food webs (Lu et al., 2021) and rare symbionts (*i.e. Symbiodinium*) increase coral-algal assemblies in the face of environmental alterations (Ziegler et al., 2018). Hence, these results raise the possibility that less abundant group-specific generalists contribute to the differential degradation of organic matter and play a practical role in microbial interactions by mediating nutrient availability in a manner that may positively or negatively affect STEC colonization. The lack of well-defined approaches for determining keystone taxa in networks impedes our understanding of microbial interactions. Previous studies adopted global topological properties (*i.e.* betweenness centrality) for inferring keystone taxa (Poudel et al., 2016), which neglects the fact that networks tend to be modular. Identification of generalists (or keystone taxa) following the concept of among/within- module

connectivity provides the added advantage of further understanding microbial interactions. Thus, the approach in our study of using the concept of modularity to compute the role of each node is more representative of network modularity and the role of each node with/among the overall structure.

As a supportive approach, lasso regression revealed group-specific microbes that contributed to mucosa-associated *stx2* expression. Among the 7 selected genera based on lasso regression, 2 out of 7 (*Eubacterium* and *Alistipes*) were core and 3 out of 7 (*Prevotella 7,* the genera from family *Paludibacteraceae, Sutterella*) were group-specific genera. However, all of the core microbes were identified as peripherals that were poorly connected with other taxa and were not expected to affect the robustness of the network (Memmott et al., 2004). Compared to studies that emphasize the value of core microbiota in the maintenance of microbial communities, group-specific genera exhibiting diverse functions could potentially interact with the host and Stx2-carrying/Stx2-expressed bacteria, leading to the establishment of a different co-occurrence network. Hence, our study highlights the valuable role of group-specific taxa in microbial co-occurrence networks which advance the understanding of microbial interactions in terms of *stx2* expression and potential STEC colonization. It is noticeable that the identified microbial networks should be validated in the future. In addition to the mathematical algorithm that reduces spurious correlations, adding stable isotope probes in the DNA-derived microbial community and analyzing heavy-isotope enriched DNA could be a complementary approach, which could differentiate inactive microbial interactions from ecological meaningful microbial crosstalk (Kaupper et al., 2021a, 2021b). For instance, a DNA-based stable isotope probing

(SIP) approach using [^{13}C]CH_4 has been employed to identify real microbial interactions by co-occurrence analysis in ombrotrophic peatlands (Kaupper et al., 2021a). However, using heavy isotope probes to examine true microbial interactions within the dense microbial communities of the gut needs further evaluation.

3.5 Conclusions

Overall, our results revealed comparable diversity and composition of RAJ microbiota were observed between Stx2- and Stx2+ cattle with more than 60% genera recognized as members of the core microbiome. However, RMT-based network analysis revealed varied microbial interactions, keystone taxa, and stability of microbial communities in response to *stx2* expression. Group-specific taxa play an important role in the network which might drive microbiota-*stx2* interactions. This study also constitutes an in-depth understanding of host-STEC interactions and highlights the possibility of altering the gut environment to mitigate *stx2* expression through modifying the gut microbiota. However, future validations using a larger sample size are needed to verify the proposed methods for deciphering microbial community differences as well as the principles of microbe coexistence that determine microbial interactions. Regardless, our findings highlight the critical role of group-specific genera among microbial interactions related to the expression of *stx2* in bacteria and shed light on using an approach that integrates group-specific genera with network analysis to identify and characterize differences in microbial communities with comparable microbial profiles.

3.6 References

Aitchison, J. (1981). A new approach to null correlations of proportions. *J Int Ass Math Geol* 13, 175–189. doi: 10.1007/bf01031393.

Albert, R., and Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Rev Mod Phys* 74, 47–97. doi: 10.1103/revmodphys.74.47.

Barberán, A., Bates, S. T., Casamayor, E. O., and Fierer, N. (2012). Using network analysis to explore co-occurrence patterns in soil microbial communities. *Isme J* 6, 343–351. doi: 10.1038/ismej.2011.119.

Bertin, Y., Girardeau, J. P., Chaucheyras-Durand, F., Lyan, B., Pujos-Guillot, E., Harel, J., et al. (2011). Enterohaemorrhagic Escherichia coli gains a competitive advantage by using ethanolamine as a nitrogen source in the bovine intestinal content. *Environ Microbiol* 13, 365–377. doi: 10.1111/j.1462-2920.2010.02334.x.

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., et al. (2019). Author Correction: Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* 37, 1091–1091. doi: 10.1038/s41587-019-0252-6.

Braga, R. M., Dourado, M. N., and Araújo, W. L. (2016). Microbial interactions: ecology in a molecular perspective. *Braz J Microbiol* 47, 86–98. doi: 10.1016/j.bjm.2016.10.005.

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* 13, 581–583. doi: 10.1038/nmeth.3869.

Chatelier, E. L., Nielsen, T., Qin, J., Prifti, E., Hildebrand, F., Falony, G., et al. (2013). Richness of human gut microbiome correlates with metabolic markers. *Nature* 500, 541–546. doi: 10.1038/nature12506.

Csermely, P. (2009). Weak Links, The Universal Key to the Stability of Networks and Complex Systems. *Frontiers Collect*, 53–100. doi: 10.1007/978-3-540-31157-7_3.

Deng, Y., Jiang, Y.-H., Yang, Y., He, Z., Luo, F., and Zhou, J. (2012). Molecular ecological network analyses. *Bmc Bioinformatics* 13, 113. doi: 10.1186/1471-2105-13-113.

Derakhshani, H., Plaizier, J. C., Buck, J. D., Barkema, H. W., and Khafipour, E. (2020). Composition and co-occurrence patterns of the microbiota of different niches of the bovine mammary gland: potential associations with mastitis susceptibility, udder inflammation, and teat-end hyperkeratosis. *Animal Microbiome* 2, 11. doi: 10.1186/s42523-020-00028-6.

Donkersgoed, J. V., Berg, J., Potter, A., Hancock, D., Besser, T., Rice, D., et al. (2001). Environmental sources and transmission of Escherichia coli O157 in feedlot cattle. *Can Vet J La Revue Vétérinaire Can* 42, 714–20.

Fraser, M. E., Fujinaga, M., Cherney, M. M., Melton-Celsa, A. R., Twiddy, E. M., O'Brien, A. D., et al. (2004). Structure of Shiga Toxin Type 2 (Stx2) from Escherichia coli O157:H7*. *J Biol Chem* 279, 27511–27517. doi: 10.1074/jbc.m401939200.

Guimerà, R., and Amaral, L. A. N. (2005). Functional cartography of complex metabolic networks. *Nature* 433, 895–900. doi: 10.1038/nature03288.

Guo, W., Zhou, M., Ma, T., Bi, S., Wang, W., Zhang, Y., et al. (2020). Survey of rumen microbiota of domestic grazing yak during different growth stages revealed novel maturation patterns of four key microbial groups and their dynamic interactions. *Animal Microbiome* 2, 23. doi: 10.1186/s42523-020-00042-8.

Han, K., Lim, H. C., and Hong, J. (1992). Acetic acid formation in escherichia coli fermentation. *Biotechnol Bioeng* 39, 663–671. doi: 10.1002/bit.260390611.

Hicklin, A. (2004). Network Stability: Opportunity or Obstacles? *Public Organization Rev* 4, 121–133. doi: 10.1023/b:porj.0000031625.78226.bc.

Hirano, H., and Takemoto, K. (2019). Difficulty in inferring microbial community structure based on co-occurrence network approaches. *Bmc Bioinformatics* 20, 329. doi: 10.1186/s12859-019-2915-1.

Horner-Devine, M. C., Silver, J. M., Leibold, M. A., Bohannan, B. J. M., Colwell, R. K., Fuhrman, J. A., et al. (2007). A COMPARISON OF TAXON CO-OCCURRENCE

PATTERNS FOR MACRO- AND MICROORGANISMS. *Ecology* 88, 1345–1353. doi: 10.1890/06-0286.

Jiao, S., Liu, Z., Lin, Y., Yang, J., Chen, W., and Wei, G. (2016). Bacterial communities in oil contaminated soils: Biogeography and co-occurrence patterns. *Soil Biology Biochem* 98, 64–73. doi: 10.1016/j.soilbio.2016.04.005.

Jun, W., Barahona, M., Yue-Jin, T., and Hong-Zhong, D. (2010). Natural Connectivity of Complex Networks. *Chinese Phys Lett* 27, 078902. doi: 10.1088/0256-307x/27/7/078902.

Karmali, MohamedA., Petric, M., Steele, BrianT., and Lim, C. (1983). SPORADIC CASES OF HAEMOLYTIC-URAEMIC SYNDROME ASSOCIATED WITH FAECAL CYTOTOXIN AND CYTOTOXIN-PRODUCING ESCHERICHIA COLI IN STOOLS. *Lancet* 321, 619–620. doi: 10.1016/s0140-6736(83)91795-6.

Kaupper, T., Mendes, L. W., Harnisz, M., Krause, S. M. B., Horn, M. A., and Ho, A. (2021a). Recovery of Methanotrophic Activity Is Not Reflected in the Methane-Driven Interaction Network after Peat Mining. *Appl Environ Microb* 87. doi: 10.1128/aem.02355-20.

Kaupper, T., Mendes, L. W., Lee, H. J., Mo, Y., Poehlein, A., Jia, Z., et al. (2021b). When the going gets tough: Emergence of a complex methane-driven interaction network during recovery from desiccation-rewetting. *Soil Biology Biochem* 153, 108109. doi: 10.1016/j.soilbio.2020.108109.

Krysenko, S., Matthews, A., Okoniewski, N., Kulik, A., Girbas, M. G., Tsypik, O., et al. (2019). Initial Metabolic Step of a Novel Ethanolamine Utilization Pathway and Its Regulation in Streptomyces coelicolor M145. *Mbio* 10. doi: 10.1128/mbio.00326-19.

Lecca, P., and Re, A. (2015). Detecting modules in biological networks by edge weight clustering and entropy significance. *Frontiers Genetics* 6, 265. doi: 10.3389/fgene.2015.00265.

Lu, L., Li, X., Li, Z., Chen, Y., García, C. A. S. y, Yang, J., et al. (2021). Aerobic methanotrophs in an urban water cycle system: Community structure and network interaction pattern. *Sci Total Environ* 772, 145045. doi: 10.1016/j.scitotenv.2021.145045.

Luo, F., Yang, Y., Zhong, J., Gao, H., Khan, L., Thompson, D. K., et al. (2007). Constructing gene co-expression networks and predicting functions of unknown genes by random matrix theory. *Bmc Bioinformatics* 8, 299. doi: 10.1186/1471-2105-8-299.

Mao, S., Zhang, M., Liu, J., and Zhu, W. (2015). Characterising the bacterial microbiota across the gastrointestinal tracts of dairy cattle: membership and potential function. *Sci Rep-uk* 5, 16116. doi: 10.1038/srep16116.

Maslov, S., and Sneppen, K. (2002). Specificity and stability in topology of protein networks. *Arxiv*. doi: 10.1126/science.1065103.

Memmott, J., Waser, N. M., and Price, M. V. (2004). Tolerance of pollination networks to species extinctions. *Proc Royal Soc Lond Ser B Biological Sci* 271, 2605–2611. doi: 10.1098/rspb.2004.2909.

Munns, K. D., Selinger, L. B., Stanford, K., Guan, L., Callaway, T. R., and McAllister, T. A. (2015). Perspectives on Super-Shedding of Escherichia coli O157:H7 by Cattle. *Foodborne Pathog Dis* 12, 89–103. doi: 10.1089/fpd.2014.1829.

Newman, M. E. J. (2006). Modularity and community structure in networks. *Proc National Acad Sci* 103, 8577–8582. doi: 10.1073/pnas.0601602103.

Pan, Z., Chen, Y., McAllister, T. A., Gänzle, M., Plastow, G., and Guan, L. L. (2021). Abundance and Expression of Shiga Toxin Genes in Escherichia coli at the Recto-Anal Junction Relates to Host Immune Genes. *Front Cell Infect Mi* 11, 633573. doi: 10.3389/fcimb.2021.633573.

Poudel, R., Jumpponen, A., Schlatter, D. C., Paulitz, T. C., Gardener, B. B. M., Kinkel, L. L., et al. (2016). Microbiome Networks: A Systems Framework for Identifying Candidate Microbial Assemblages for Disease Management. *Phytopathology* 106, 1083–1096. doi: 10.1094/phyto-02-16-0058-fi.

Ravasz, E., Somera, A. L., Mongru, D. A., Oltvai, Z. N., and Barabási, A.-L. (2002). Hierarchical Organization of Modularity in Metabolic Networks. *Science* 297, 1551–1555. doi: 10.1126/science.1073374.

Shanks, O. C., Kelty, C. A., Archibeque, S., Jenkins, M., Newton, R. J., McLellan, S. L., et al. (2011). Community Structures of Fecal Bacteria in Cattle from Different Animal Feeding Operations†. *Appl Environ Microb* 77, 2992–3001. doi: 10.1128/aem.02988-10.

Sheng, H., Lim, J. Y., Knecht, H. J., Li, J., and Hovde, C. J. (2006). Role of Escherichia coli O157:H7 Virulence Factors in Colonization at the Bovine Terminal Rectal Mucosa. *Infect Immun* 74, 4685–4693. doi: 10.1128/iai.00406-06.

Shi, Y., Delgado-Baquerizo, M., Li, Y., Yang, Y., Zhu, Y.-G., Peñuelas, J., et al. (2020). Abundance of kinless hubs within soil microbial networks are associated with high functional potential in agricultural ecosystems. *Environ Int* 142, 105869. doi: 10.1016/j.envint.2020.105869.

Van, T. D. le, Robinson, J. A., Ralph, J., Greening, R. C., Smolenski, W. J., Leedle, J. A. Z., et al. (1998). Assessment of Reductive Acetogenesis with Indigenous Ruminal Bacterium Populations and Acetitomaculum ruminis. *Appl Environ Microb* 64, 3429–3436. doi: 10.1128/aem.64.9.3429-3436.1998.

Wang, O., Liang, G., McAllister, T. A., Plastow, G., Stanford, K., Selinger, B., et al. (2016). Comparative Transcriptomic Analysis of Rectal Tissue from Beef Steers Revealed Reduced Host Immunity in Escherichia coli O157:H7 Super-Shedders. *Plos One* 11, e0151284. doi: 10.1371/journal.pone.0151284.

Wang, O., McAllister, T. A., Plastow, G., Stanford, K., Selinger, B., and Guan, L. L. (2017). Host mechanisms involved in cattle Escherichia coli O157 shedding: a fundamental understanding for reducing foodborne pathogen in food animal production. *Sci Rep-uk* 7, 7630. doi: 10.1038/s41598-017-06737-4.

Wang, O., McAllister, T. A., Plastow, G., Stanford, K., Selinger, B., and Guan, L. L. (2018). Interactions of the Hindgut Mucosa-Associated Microbiome with Its Host Regulate Shedding of Escherichia coli O157:H7 by Cattle. *Appl Environ Microb* 84, e01738-17. doi: 10.1128/aem.01738-17.

Wang, O., Zhou, M., Chen, Y., McAllister, T. A., Plastow, G., Stanford, K., et al. (2021). MicroRNAomes of Cattle Intestinal Tissues Revealed Possible miRNA Regulated Mechanisms Involved in Escherichia coli O157 Fecal Shedding. *Front Cell Infect Mi* 11, 634505. doi: 10.3389/fcimb.2021.634505.

Watts, D. J., and Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature* 393, 440–442. doi: 10.1038/30918.

Wolfe, A. W. (1997). Social Network Analysis: Methods and Applications. *Am Ethnol* 24, 219–220. doi: 10.1525/ae.1997.24.1.219.

Xia, Y. (2020). Chapter Eleven Correlation and association analyses in microbiome study integrating multiomics in health and disease. *Prog Mol Biol Transl* 171, 309–491. doi: 10.1016/bs.pmbts.2020.04.003.

Xu, Y., Dugat-Bony, E., Zaheer, R., Selinger, L., Barbieri, R., Munns, K., et al. (2014). Escherichia coli O157:H7 Super-Shedder and Non-Shedder Feedlot Steers Harbour Distinct Fecal Bacterial Communities. *Plos One* 9, e98115. doi: 10.1371/journal.pone.0098115.

Zhang, B., and Horvath, S. (2005). A General Framework for Weighted Gene Co-Expression Network Analysis. *Stat Appl Genet Mol* 4, Article17. doi: 10.2202/1544-6115.1128.

Zhou, J., Deng, Y., Luo, F., He, Z., Tu, Q., and Zhi, X. (2010). Functional Molecular Ecological Networks. *Mbio* 1, e00169-10. doi: 10.1128/mbio.00169-10.

Zhou, J., Deng, Y., Luo, F., He, Z., and Yang, Y. (2011). Phylogenetic Molecular Ecological Network of Soil Microbial Communities in Response to Elevated CO2. *Mbio* 2, e00122-11. doi: 10.1128/mbio.00122-11.

Ziegler, M., Eguíluz, V. M., Duarte, C. M., and Voolstra, C. R. (2018). Rare symbionts may contribute to the resilience of coral–algal assemblages. *Isme J* 12, 161–172. doi: 10.1038/ismej.2017.151.

3.7 Tables and figures

Table 3.1. Sample demographics

Group	ID	Body Weight (kg)	Mean Value(SD)	Kill Age (days)	Mean Value(SD)
Stx2-	14-101	566		463	
	14-103	584		451	
	14-112	479	539 (41.70)	429	461(18.38)
	14-205	557		474	
	14-206	496		476	
	14-207	552		474	
Stx2+	14-104	698		460	
	14-105	555		455	
	14-106	616	597 (54.61)	465	451 (9.82)
	14-107	555		448	
	15-104	569		440	
	15-106	591		443	
One Way ANOVA		F-Value= 4.32; P-Value= 0.07		F-Value= 1.22; P-Value= 0.30	

Table 3.2. Summary of amplicon sequencing results

Group	ID	No. of raw reads	No. of filtered reads	No. of features	No. of ASVs	Good's Coverage
Stx2-	14-101	30482	13612	5438	275	>99%
	14-103	33013	14809	4758	215	>99%
	14-112	31604	14390	7269	261	>99%
	14-205	34445	15846	5731	234	>99%
	14-206	31578	14100	4858	238	>99%
	14-207	32687	15111	5039	174	>99%
Stx2+	14-104	38798	17486	6142	256	>99%
	14-105	33962	15584	5381	228	>99%
	14-106	31882	14301	5927	250	>99%
	14-107	31248	12693	6067	214	>99%
	15-104	39153	17863	6142	253	>99%
	15-106	31328	14122	4288	200	>99%

Table 3.3. Taxonomic profiles at phylum level between Stx2- and Stx2+ group

Phylum/Group	Stx2- Mean	SD	Stx2+ Mean	SD	P-Value (FDR adjusted)
Actinobacteria	0.0096	0.0101	0.0120	0.0072	0.66
Bacteroidetes	0.2720	0.0533	0.2207	0.0714	0.47
Cyanobacteria	0.0030	0.0024	0.0036	0.0045	0.74
Firmicutes	0.6977	0.0595	0.7562	0.0685	0.47
Proteobacteria	0.0122	0.0114	0.0043	0.0023	0.47
Tenericutes	0.0010	0.0007	0.0008	0.0007	0.87
Others	0.0046	0.0060	0.0024	0.0023	0.73

α value was computed by the power analysis with $P<0.21$ considered as a significant when power=0.8.

Table 3.4. Summary of core bacterial genera shared by all samples and phylum and family to which these taxa belong

Phylum	Family	Genus
Firmicutes	*Ruminococcaceae*	*Ruminococcaceae UCG-004*
		Ruminococcaceae UCG-005
		Ruminococcaceae UCG-010
		Ruminococcaceae UCG-013
		[Eubacterium] coprostanoligenes group
		Candidatus Soleaferrea
	Christensenellaceae	*Christensenellaceae R-7 group*
	Family XIII	*Family XIII AD3011 group*
	Lachnospiraceae	__
Bacteroidetes	*Bacteroidaceae*	*Bacteroides*
	Muribaculaceae	*uncultured bacterium*
	Prevotellaceae	*Prevotellaceae UCG-003*
	Rikenellaceae	*Alistipes*
		Rikenellaceae RC9 gut group
	Uncultured	*uncultured Bacteroidales bacterium*

Table 3.5. Summary of Stx2- specific and Stx2+ specific microbial genera, and phylum and family to which these taxa belong.

Phylum	Family	Genus
Stx2- specific		
Actinobacteria	*Eggerthellaceae*	*DNF00809*
		Parvibacter
		Prevotella 1
	Prevotellaceae	*Prevotella 7*
		Prevotella 9
Bacteroidetes		*uncultured*
	O_Bacteroidales	__
	uncultured	__
Cyanobacteria	*uncultured rumen bacterium*	*uncultured rumen bacterium*
Firmicutes	*Erysipelotrichaceae*	__
	Lachnospiraceae	*[Eubacterium] hallii group*
		Acetitomaculum
	Family XIII	*[Eubacterium] nodatum group*
	Ruminococcaceae	*Ruminococcaceae UCG-011*
		Ruminiclostridium 1
	Erysipelotrichaceae	*Candidatus Stoquefichus*
		Erysipelotrichaceae UCG-004
	Veillonellaceae	__
		Selenomonas 1
Patescibacteria	*Saccharimonadaceae*	*Candidatus Saccharimonas*
Proteobacteria	*Uncultured*	*unidentified rumen bacterium RF32*
	Uncultured	*gut metagenome*
Tenericutes	*uncultured bacterium*	*uncultured bacterium*
	gut metagenome	*gut metagenome*
Stx2+ specific		
Actinobacteria	*Streptomycetaceae*	*Streptomyces*
Bacteroidetes	*Paludibacteraceae*	*uncultured*
	Muribaculaceae	__
	Bacteroidales RF16 group	*uncultured Parabacteroides*
Chloroflexi	*Anaerolineaceae*	*Flexilinea*
Cyanobacteria	*O_Gastranaerophilales*	__
Firmicutes	*Ruminococcaceae*	*Ruminococcus 1*
		Ruminiclostridium 9
	Lachnospiraceae	*Tyzzerella*
		Tyzzerella 4
Proteobacteria	*Burkholderiaceae*	*Sutterella*
	Acetobacteraceae	*Acetobacter*
Tenericutes	*O_Mollicutes RF39*	__

Table 3.6. Topological properties of empirical and random network between Stx2- and Stx2+ groups.

	Empirical network		Random network	
Group	Stx2-	Stx2+	Stx2-	Stx2+
Nodes	86	77	86	77
Edges	322	243	322	243
Modularity (MD)	0.44	0.46	0.29 (± 0.011)	0.31 (± 0.012)
Clustering coefficient (CC)	0.32	0.30	0.12 (± 0.014)	0.12 (± 0.016)
Average path length (APL)	2.84	2.86	2.47 (± 0.030)	2.58 (± 0.033)
Graph density (GD)	0.088	0.083	0.09	0.08
Average degree (AD)	7.49	6.31	7.49	6.31

Table 3.7. Stx2- and Stx2+ generalist genera and their associated bacterial phyla.

Phylum	Genus	Group-specific
Stx2- group		
Actinobacteria	*Parvibacter*	Stx2- specific
Bacteroidetes	f_*Bacteroidales* RF16 group	
	f_ *Prevotellaceae*	Stx2- specific
Cyanobacteria	o_ *Gastranaerophilales*	Stx2- specific
	f_ *Veillonellaceae*	Stx2- specific
Firmicutes	f_ *Veillonellaceae*	
	Acetitomaculum	Stx2- specific
Patescibacteria	*Candidatus saccharimonas*	Stx2- specific
Planctomycetes	p-1088-a5 gut group	
Proteobacteria	*Parasutterella*	
	o_ *Rhodospirillales*	Stx2- specific
Tenericutes	o_ *Izimaplasmatales*	Stx2- specific
Stx2+ group		
Actinobacteria	*Saccharopolyspora rectivirgula*	
	Streptomyces	Stx2+ specific
Bacteroidetes	*Prevotellaceae* Ga6A1 group	Stx2+ specific
	Prevotellaceae UCG-001	
	Parabacteroides	
Chloroflexi	*Flexilinea*	Stx2+ specific
	Oscillibacter	
	Ruminiclostridium 9	Stx2+ specific
Firmicutes	*Ruminococcus* 1	Stx2+ specific
	f_ *Erysipelotrichaceae*	
	Cellulosilyticum	
Planctomycetes	p-1088-a5 gut group	
Proteobacteria	*Acetobacter*	Stx2+ specific
	Parasutterella	

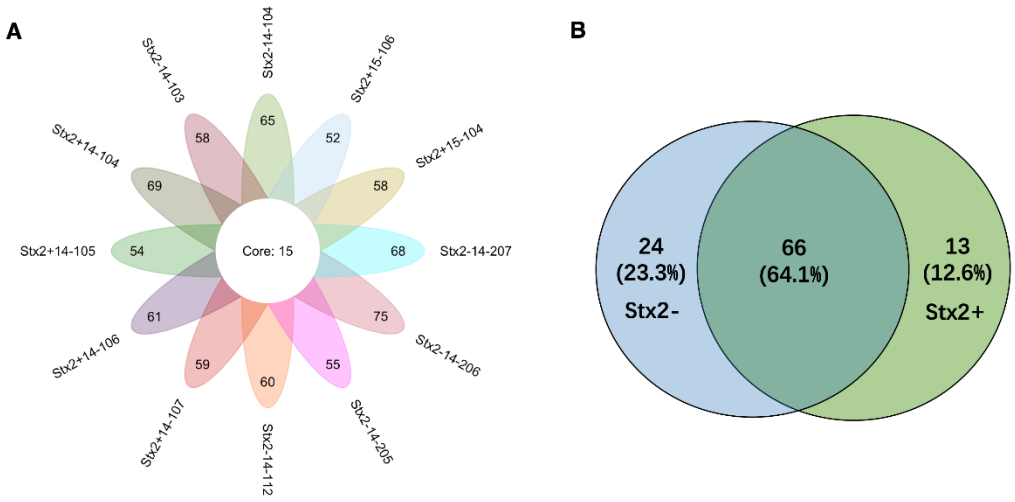


Figure 3.1. Shared and specific genera between Stx2- and Stx2+ groups.

A. A flower plot visualization of the number of core genera shared by each sample (in the center) and the number of specific genera found in each sample (in the petals). B. Genera detected in Stx2+ and Stx2- group. Detected genera, total relative abundance of 0.1% within at least two samples in each group.

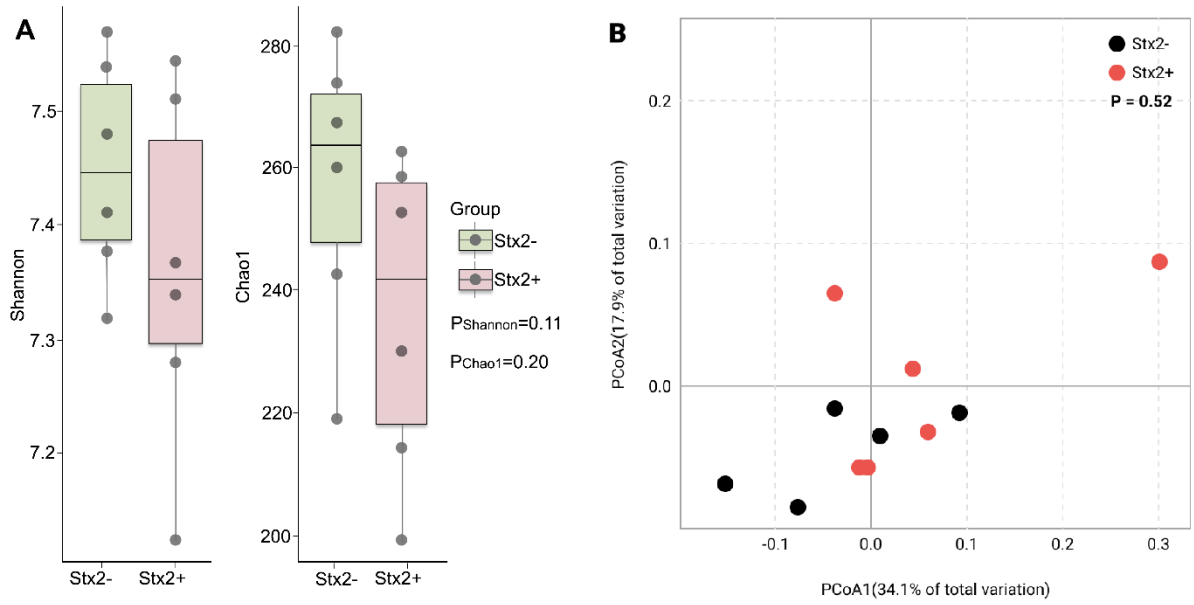


Figure 3.2. Comparison of diversity metrics between Stx2- and Stx2+ groups.

A. Shannon and Chao1 indices were used to estimate the evenness and richness between Stx2- and Stx2+ groups, respectively. The horizontal bars within boxes represent medians. The top and bottom of each box represent the 75th and 25th percentiles, respectively. The upper and lower whiskers extend to data no more than 1.5x the interquartile range from the upper edge and lower edge of the box, respectively. The Kruskal-Wallis test was used to determine whether indices between the two groups were significant. ($P \leq 0.05$). B. Principal coordinate analysis (PCoA) was used for visualization of Weighted Unifrac distance. The PERMANOVA was used to test for the similarity of clustering patterns between Stx2- and Stx2+ groups. Differences were considered significant at $P \leq 0.05$.

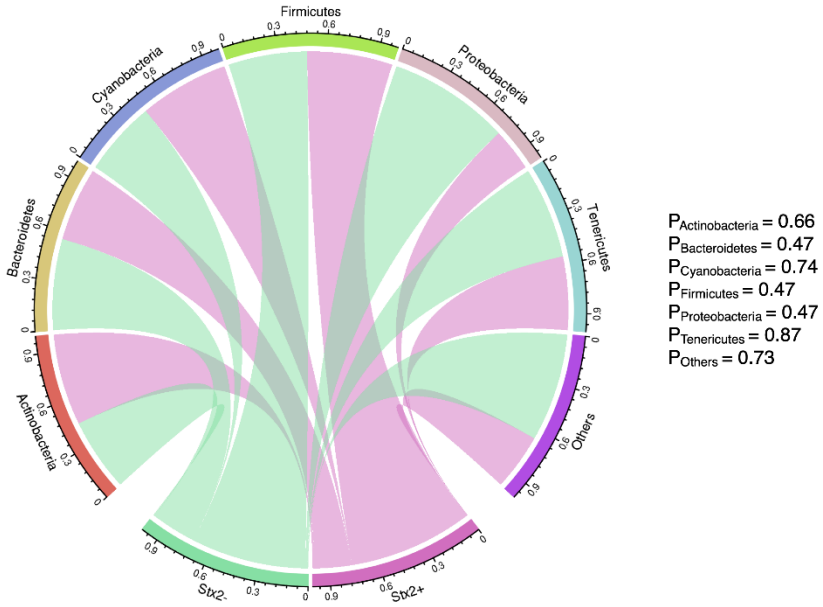


Figure 3.3. Comparison of average relative abundance of phyla between Stx2- and Stx2+ groups.

Circos plots were used for visualization of average relative abundance of phyla between Stx2- and Stx2+ groups. Phyla in each group with a total relative abundance of >0.1% in at least three samples were included. The Kruskal-Willas test showed that the average relative abundance of phyla was comparable (P>0.05) between Stx2- and Stx2+ groups. P-values were false discovery rate (FDR, q=0.05) adjusted.

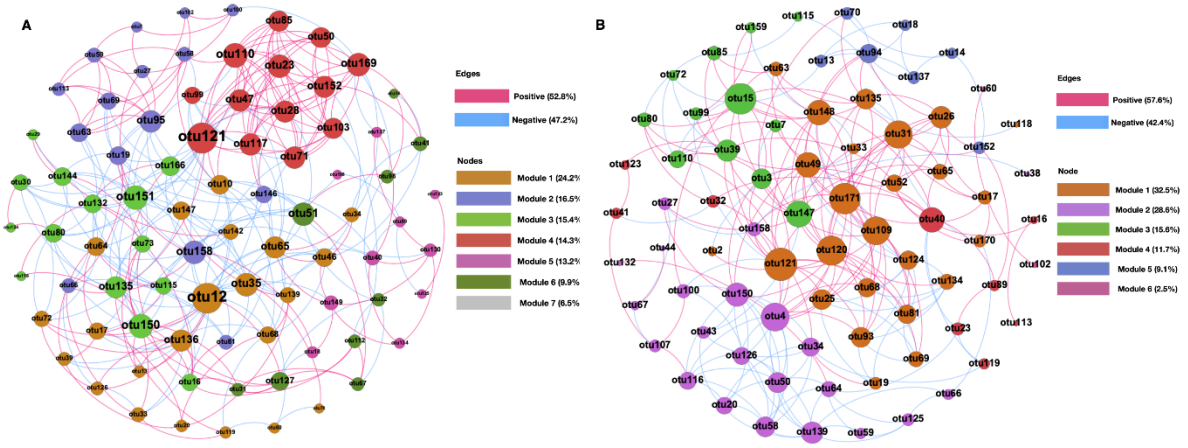


Figure 3.4. Co-occurrence networks of bacterial genera in (A) Stx2- and (B) Stx2+ groups

The size of each node is proportional to the number of connections (that is, degree); the color of connections between two nodes represents a positive (red) or a negative correlation (blue). Each module is presented as a specific color.

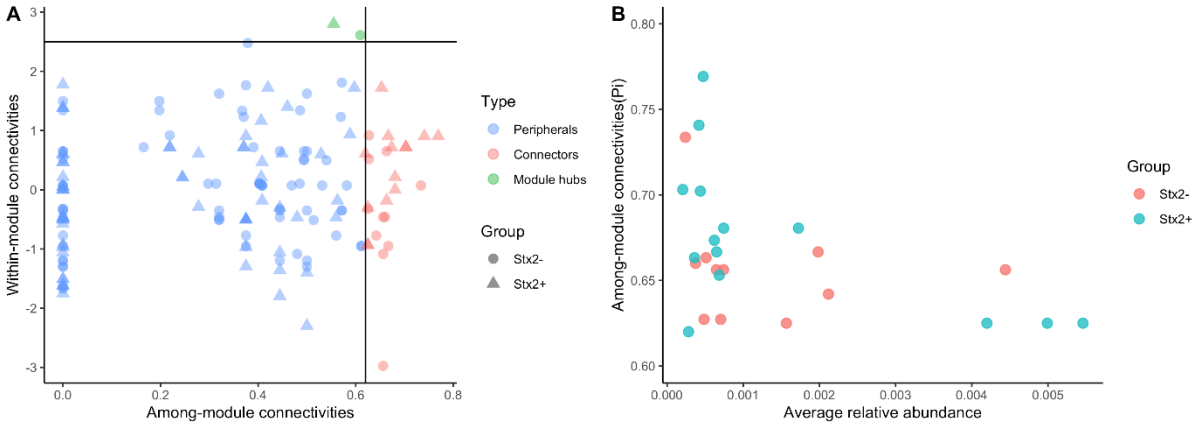


Figure 3.5. Distribution and average relative abundance of genera based on their network roles.

A. Distribution of bacterial genera based on their network roles. Nodes in the network were classified as peripherals, modular hubs, or connectors. No network hubs were identified in networks from both groups. B. Average relative abundance of connectors.

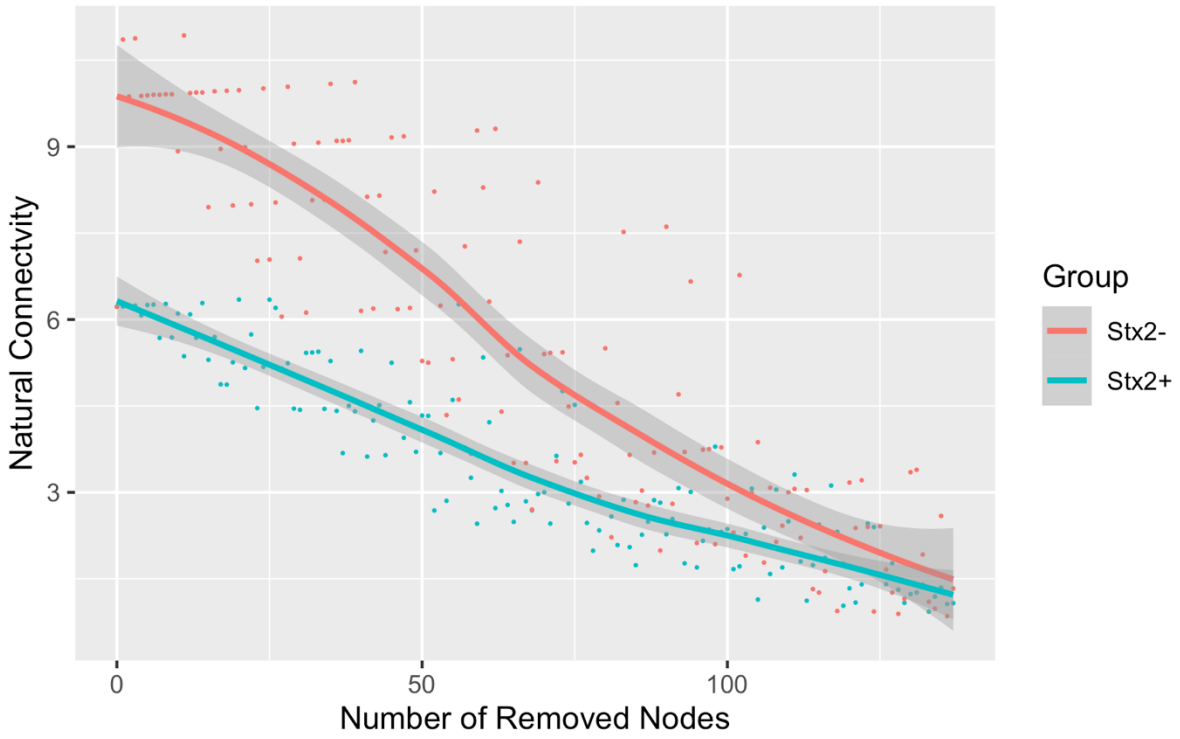


Figure 3.6. The natural connectivity representing the stability of co-occurrence networks in both Stx2- and Stx2+ groups.

Chapter 4. Assessment of host transcriptome variation in response to strain-specific Shiga toxin-*Escherichia coli* O157 colonization at the rectal-anal junction in veal calves

4.1 Introduction

Shiga toxin-*Escherichia coli* (STEC) possesses Shiga toxins as the major virulence factors, with prophages being designated to Shiga toxin 1 (stx1) and Shiga toxin 2 (stx2). The stx2 is 400 times more toxic than stx1 and is associated with human illnesses such as bloody diarrhea and, hemolytic uremic syndrome (HUS) (Chase-Topping et al., 2008). The stx2 gene has a number of variants including stx2 a, b, c, d, e, f, and g. Although the acquisition of stx2a prophage has been reported to be most often associated with STEC O157 pathogenicity (Fraser et al., 2004; Melton-Celsa, 2014; Castro et al., 2017; Ryu et al., 2020), a recent study revealed that phage type 21/28 which contains both stx2a and stx2c encoding prophages but does not express functional stx2a was also commonly linked to STEC disease in humans (Chase-Topping et al., 2008). For instance, 61% of HUS cases in children were caused by PT 21/28$^{stx2a\text{-}stx2c+}$ in Scotland from 1997 to 2001 (Chase-Topping et al., 2008).

Ruminants, especially cattle are the major reservoir for STEC O157 with the rectal-anal junction (RAJ) being the major colonization site (Cobbold and Desmarchelier, 2000; Wang et al., 2017, 2018; McCabe et al., 2019). Cattle who shed more than 10^4 CFU STEC per gram of feces are defined as 'super shedders' (SS), which are considered the primary source of STEC transmission on farms (Chase-Topping et al., 2008; Xu et al., 2014; Munns et al., 2015; Wang et al., 2016). Downregulated differentially expressed genes (DEGs) identified in SS compared

to non-shedders have been shown to be related to host immunity (both humoral and cell-mediated immune responses) and metabolism (*i.e.,* cholesterol absorption) (Wang et al., 2016, 2017). Another study revealed that the production of stx2a in STEC O157 in veal calves could induce increased levels of local (rectal lymph node) and systemic (serum) H7- (flagella antigen 7) specific antibodies (Fitzgerald et al., 2019), suggesting stx2a production in STEC could affect host responses to STEC. However, STEC O157 colonization is likely a complicated process that is influenced by multi-biological pathways whose interrelationships are largely undefined.

Using transcript abundance to characterize host functional variation is a common approach in both human and animal studies, particularly, studies focused on the function of high-and low-abundance mRNAs and their roles in host physiology and health. For instance, high abundance transcripts in 21-24 year olds were involved in glucose metabolism and electron transport, while these transcripts were 30% less abundant in the vastus lateralis muscles of 66-77 year olds. This suggests that high abundance transcripts are related to host phenotypic changes in senescent muscles (Welle et al., 2000). Previous studies suggested that low abundance transcripts with the most fold changes serve as important regulators in various biological processes (Araki et al., 2006; Lomana et al., 2020). For example, the low abundant transcripts Laminin Subunit Alpha 3 (*LAMA3*) and Titin (*TTN*) were identified as novel biomarkers for the detection of human gastric cancer (Bizama et al., 2014). However, the function of high- and low-abundant transcripts in the mucosa RAJ and their functional shifts in repones to STEC O157 colonization has not been investigated. As the major site of STEC O157

colonization, the knowledge of the fundamental function of RAJ mucosa in controlling bacterial colonization in young ruminants is limited. Furthermore, the extent to which mucosal function varies in terms of stx2 subtype-specific STEC O157 colonization is unclear. Therefore, we hypothesize that the rectal mucosal transcriptome and its function are altered as a result of STEC O157 colonization in veal calves and such variation depends on the production of stx2a. We also speculate that abundance and diversity of transcripts is influenced by STEC O157 colonization. We aimed to adopt high-throughput RNA sequencing to explore gene expression in RAJ mucosa as it responds to a strain-specific STEC O157 challenge of young calves.

4.2 Materials and methods

4.2.1 Animal study and sample collection

The bacterial strains (PT 21/28 *stx2a-stx2c+* and RE 21/28 *stx2a+stx2c+* STEC O157) and challenge trials with young calves have previously been by Fitzgerald et al (Fitzgerald et al., 2019). All animal work was carried out at the Moredun Research Institute (MRI) following the Animal Care Protocol approved by the MRI animal experiments committee and ethical review committee. Briefly, male Holstein-Friesian dairy calves were purchased by MRI at three weeks of age. Prior to the oral challenge, calves were pre-screened five times per week using both plate culture and RT-qPCR tof fecal samples to ensure that calves were STEC O157 negative. Calves were randomly assigned into three groups: PT 21/28 *stx2a-stx2c+* (WT group, n=6), RE 21/28 *stx2a+ stx2c+* (RE group, n=6) and control (CT group, n=11). Calves were orally challenged by orogastric intubation with ~10^9 CFU of each strain in 10 mL of Lysogeny broth. Fecal

samples and STEC were enumerated as described by Fitzgerald et al (., 2019). Briefly, fecal samples were collected from the rectum daily for the first 18 days post oral challenge and subsequently every other day until the completion of the experiment after 26 days. For sentinel animals, fecal samples were taken every other day. Ten grams of feces were placed into 90 ml of sterile PBS and 10-fold serial dilutions of the feces samples were then prepared in PBS and 100 µl from three dilutions were plated across a 1,000-fold range of dilutions onsorbitol MacConkey agar containing 15 µg nalidixic acid (NAL-SMAC). Plates were incubated at 37 °C overnight and colonies were counted in plates that contained 30 to 300 colonies. Randomly selected colonies from each plate were confirmed as O157 positive by latex agglutination (Oxoid, Basingstoke, United Kingdom) following the manufacturer's instructions. Samples from the control animals were plated directly onto cefixime –tellurite (CT)-SMAC plates containing 0.05 mg/L cefixime, 2.5 mg/L tellurite. Plates were incubated overnight at 37 °C and then enumerated the next day.

The recto-anal junction (RAJ) tissue was biopsied using non-evasive tools at week 1 (before challenge, T1), week 2 (the highest fecal shedding level for challenge groups, T2) and week 5 (recovered from STEC O157 colonization, T5) of the trial. Tissue samples were stored at -80 °C before use.

4.2.2 RNA extraction and sequencing

RNA was isolated from 0.1 g biopsy tissue using trizol reagent (Invitrogen Corporation, Carlsbad, CA, USA), and purified using a RNeasy MinElute Cleanup kit (Qiagen, Valencia, CA, USA), and assessed using Agilent 2200 TapeStation (Agilent Technologies, Santa Clara,

CA, USA) and Qubit 3.0 Fluorometer (Invitrogen, Carlsbad, CA, USA). RNA with an integrity number (RIN) greater than 7.0 and ratio of A260/A280 from 1.7 to 2.4 was used for RNA-sequencing.

Extracted total RNA (1ug) was used for library construction using the Truseq Stranded total RNA sample preparation kit (Illumina, San Diego, CA, USA) following the manufacturer's manual. The rRNAs were first depleted using biotinylated, target-specific oligos combined rRNA removal beads. The remaining RNA was fragmented and then subject to first-strand cDNA synthesis with reverse transcription using random primers. The DNA polymerase I and Rnase H were used for the second strand cDNA synthesis followed by ligation of the indexed-adapters and PCR enrichment (98°C for 30 sec, followed by 15 cycles of: 98°C for 10 sec, 60°C for 30 sec, 72°C for 30 sec, and 72°C for 5 min). The quality of constructed libraries was assessed using Agilent 2200 TapeStation and a Qubit 2.0 Fluorometer. RNA sequencing was performed using a HiSeq 4000 sequencing system (Illumina, San Diego, CA, USA), with paired-end (100 bp) sequencing at Genome Quebec Innovation Centre, Montreal, Quebec, Canada.

4.2.3 Transcriptome profiling for calf rectum

Sequencing reads were first subjected to the quality filter and adapter trimming using FastQC and bbDuk. The filter high-quality reads were then mapped against the reference bovine reference assembly ARS-UCD 1.2 using STAR (Version 2.7.1a). Feature counts of each ensembl ID were then generated using subread (Version 2.0.0) and were then normalized into TPM (transcripts per million) using the formula: (exon reads in genes) / (total exon reads) x 1

million. Transcripts with an average TPM>0.2 were considered as detected transcripts that could be used in further analysis (Sriram et al., 2019). Host gene expression patterns were identified using principal component analysis (PCA). The similarities of gene expression patterns from pre- (T1) to post-challenge (T5) in CT, WT, and RE groups were identified using pairwise principal coordinate analysis (PCoA) with ANOSIM (analysis of similarities) testing the significance of similarities ($P<0.05$ was considered significant, $P \geq 0.05$ & <0.1 as a trend of being significant).

4.2.4 Identification of abundance-specific genes and their functional annotations

The core transcriptome was defined as genes expressed in all samples in one group at a certain time point (e.g. genes expressed in all samples collected from the CT-T1 group). About 90% of reads from the expressed genes were designated as highly abundant transcripts in each group. The identified core transcripts with accumulated expression of about 1% of reads were considered to be low abundance transcripts. Since the expression of low abundant genes can be individual effects or random expression (Chess, 2013), the unsupervised self-organizing map (SOM) in R ('Kohonen' package) was further adopted to organize low abundant genes into biological meaningful clusters and select feature low abundant transcripts for functional annotation. The SOM is a non-parametric unsupervised machine learning approach, which enables the transformation of high to low dimension transcriptomic data, while preserving the topological structure of the data (Reusch et al., 2005; Astel et al., 2007; Das et al., 2015). The major pathways of identified high- and flow-abundant transcripts were annotated using

Ingenuity pathway analysis (IPA, QIAGEN, Redwood City, CA, United States www.qiagen.com/ingenuity).

4.2.5 Identification of differentially expressed genes in response to STEC O157 challenge and its functional annotations

Differentially expressed genes (DEGs) were identified using pairwise comparisons: WT vs. CT, RE vs. CT, and WT vs. RE post-challenge (T2 and T5) using the DeSeq2 package in R. The package DESeq2 tests differential expression of genes using negative binomial generalized linear models and estimates the dispersion and logarithmic fold changes of genes. Log2 fold change of each DEG was calculated using the equation: for comparisons between WT vs. CT or RE vs. CT: log2 fold change = log2 (average TPM in challenged calves / average TPM in sentinel calves); for the comparison between WT vs. RE: log2 fold change = log2(average TPM in the WT group / average TPM in the RE group). The log2 fold change for DEGs was set to $\leq$ -1 or $\geq$ 1 with negative values indicating down regulation and positive values indicating up regulation. The DEGs were defined as genes with Benjamini-Hochberg adjusted $P<0.05$ and absolute log2 fold change >1. The functions of DEGs were enriched using the Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) ($P<0.05$ as the cut-off).

4.3 Results

4.3.1 Transcriptome profiling of bovine rectal mucosal biopsies

In total, an average of 33.7 ± 0.88 M high-quality mapped reads were generated from 49 ± 12 M pair-ended reads with 68.6 ± 0.06% mapping rates. A range of 18,574 to 20,445 transcripts

was identified (at least one read mapped to the gene in at least one sample). There were no remarkable differences in the quantity of identified genes from pre- to post-challenge in CT, WT, or RE groups (All P>0.05, Figure 4.1).

Rectal transcriptome profiling patterns varied from pre- to post-challenge in the WT group, while profiling patterns were similar from pre- to post-challenge in both control and RE groups (Figure 4.2). The PCoA analysis revealed that similarities of transcript patterns in the CT group were not affected by calf age from T1 to T2 (T2 vs. T1, P=0.66, Figure 4.3A), although there was a trend for significance from T2 to T5 (T5 vs. T2, P=0.06, Figure 4.3A) and it was affected by calf age from T1 to T5 (T5 vs. T1, P=0.05, Figure 4.3E). Individual responses could be a significant factor affecting the similarity in gene expression patterns in the CT group in comparison of T1 and T2 (T2 vs. T1, P=0.01, Figure 4.3B), T2 and T5 (T5 vs. T2, P=0.05, Figure 4.3D), while it did not affect transcript patterns between T1 and T5 (T1 vs. T5, P=0.18, Figure 4.3E). For the WT group, transcript patterns were remarkably different between T1 and T2 (T2 vs. T1, P=0.05, Figure 4.4A), and T2 and T5 (T5 vs. T2, P=0.03, Figure 4.4C), while they did not differ in T5 compared to T1 (T5 vs. T1, P=0.25, Figure 4.4E). The similarities of transcript patterns from pre- to post-challenge was not affected by individual responses in the WT group (All P>0.1, Figure 4.4B, D, F). For the RE group, similarities of transcript patterns were not affected by calf age from pre- to post-challenge (All P>0.1, Figure 4.5A, C, E). Transcript patterns at T1 compared to T2 (T1 vs. T2, P=0.02, Figure 4.5B) and T5 (T1 vs. T5, P=0.01, Figure 4.5F) were affected, while this response was not observed between T2 and T5 (T2 vs. T5, P=0.23, Figure 4.5D).

4.3.2 Transcripts divergent in abundance were associated with different functions

The pan-transcriptome analysis revealed 17,082 to 18,110 core transcripts, with 247 to 363 unique transcripts identified among all groups from pre- to post- challenge (Table 4.1). A range of 5,360 to 5,634, 5,049 to 5,577, and 5,055 to 5,354 highly abundant transcripts were identified from pre- to post-challenge in the CT, WT, and RE groups, respectively (Table 4.1). The identified highly abundant transcripts shared similar functions revealed by IPA with the top 5 being designated to the BAG2 signaling pathway, EIF2 signaling pathway, Inhibition of ARE-Mediated mRNA Degradation Pathway, Mitochondrial Dysfunction, and Oxidative Phosphorylation.

Moreover, an average of 5,947 ± 377 low abundant transcripts were identified (low abundant transcripts refer to core transcripts that accounted for up to 1% of total reads, Table 4.2). The self-organizing map further revealed that an average of 2,166 ± 162 ranging from 1,952 to 2,502 were considered as featured low abundant transcripts (Table 4.2, Figure 4.6). The IPA enrichment analysis using featured low abundant transcripts revealed that 37 canonical pathways were shared in the CT group from T1 to T5 (Figure 4.7A). Among the top 10 enriched pathways from T1 to T5 in the CT group, three out of ten pathways including axonal guidance signaling, CREB signaling in neurons, and cardiac hypertrophy signaling (enhanced) overlapped from T1 to T5 (Figure 4.7B). For the WT group, 36 canonical pathways were shared from pre- to post-challenge (Figure 4.8A). Among the top 10 enriched pathways from pre- to post-challenge in the WT group, four out of ten pathways including CREB signaling in neurons, cardiac hypertrophy signaling, axonal guidance signaling, and histidine degradation III shared

from T1 to T5 (Figure 4.8B). Similarly, for the RE group, 36 canonical pathways were shared from pre- to post-challenge (Figure 4.9A). And five out of ten pathways including cardiac hypertrophy signaling (enhanced), G protein signaling mediated by tubby, CREB signaling in neurons, apelin muscle signaling pathway, and axonal guidance signaling were shared from pre- to post-challenge among the top 10 enriched pathways in the RE group (Figure 4.9B). It was noted that three pathways including Axonal Guidance Signaling, CREB Signaling in Neurons, and Cardiac Hypertrophy Signaling (Enhanced) overlapped from T1 to T5 in all CT, WT, and RE groups.

4.3.3 STEC O157 *stxa2-stx2c+* challenge inhibited host responses at the peak O157 fecal shedding stage compared to STEC O157 *stxa2+stx2c+* challenge

Compared to genes identified in the CT-T2 group, a total of 302 DEGs comprising 26 GO terms were downregulated in the WT-T2 group, while no DEGs were downregulated in the RE-T2 group (Table 4.3). No GO terms were enriched for upregulated DEGs identified in calves from both WT (n=39) and RE (n=1) groups compared to unchallenged calves at T2 (Table 4.3). A total of 214 downregulated DEGs comprising 45 GO terms were identified in the WT-T2 group compared to the RE-T2 group (Table 4.3). Particularly, 150 out of 411 (36.5%) downregulated DEGs and 24 GO terms were shared between WT vs. CT at T2 and WT vs. RE at T2 (Figure 4.10A). Twenty-three GO terms were WT-T2 specific, among which 21 terms were observed in comparison of WT vs. RE at T2 and 2 terms (protein-glutamine gamma-glutamyltransferase activity and detoxification) were observed for WT vs. CT at T2 (Figure 4.10B). Among these WT-T2 group-specific GO terms, response to bacterium, humoral immune response, and

antimicrobial humoral response were directly related to host immunity. In detail, 18 downregulated DEGs were involved in these host immune-related functions: *ENSBTAG00000046482, TMEM229B, ENSBTAG00000054561, ENSBTAG00000052579, UPK1B, IL36G, S100A8, IL36A, LPO* in the functional response to the bacterium; *KRT1* in the functional humoral immune response; *S100A12* in the functional antimicrobial humoral response and humoral immune response; *PGLYRP4, IL36RN, ENSBTAG00000014329, S100A9, SPINK5, PGLYRP3, WFDC5* in functional responses to the bacterium, antimicrobial humoral responses, and humoral immune responses (Table 4.4). Six DEGs were upregulated in the WT-T2 group compared to the RE-T2 group, while no GO terms were enriched (Table 4.3).

4.3.4 Comparable but enhanced host responses were induced in both STEC O157 *stx2a-stx2c+* and O157 *stx2a+stx2c+* challenged calves when STEC O157 fecal shedding levels returned to normal

Compared to the CT-T5 group, a total of 195 upregulated DEGs comprising 51 GO terms were identified in the WT-T5 group, and 200 upregulated DEGs comprising 54 GO terms were identified in the RE-T5 group (Table 4.3). The Venn plot further revealed 152 upregulated DEGs (Fig 11A) and 45 GO terms (Figure 4.11B) were shared between comparisons of WT vs. CT and RE vs. CT at T5. No GO terms were enriched using either upregulated (n=2) or downregulated (n=1) DEGs in the WT-T5 group compared to the RE-T5 group (Table 4.3). A total of 123 DEGs (Figure 4.12A) and 28 GO terms (Figure 4.12B) overlapped among downregulated DEGs at WT-T2 and upregulated DEGs at WT-T5 groups. Two GO terms were

WT-T2 specific (protein-glutamine gamma-glutamyltransferase activity and detoxification) and 28 GO terms were WT-T5 specific.

4.4 Discussion

A higher number of core transcripts (17,082 to 18,110) were identified compared to our prior study using mucosa samples from adult beef cattle (an average of 11,773 core transcripts). This difference may be due to the different physiological stages in calves and adult beef cattle, as young calves may express more genes involved in the development of organs and tissues (Werner et al., 2018). Also, this difference could be due to different experimental models; our study used a challenge model with STEC O157 strains, while the previous study collected mucosal samples from adult beef cattle that were naturally exposed to STEC. We noticed that calf age could be one of co-founding factors that influenced transcription patterns. The animal-to-animal variation in gene expression has been studied in mice (Vedell et al., 2011) and veal calves (He et al., 2018). Both studies suggested that environmental and host factors were responsible for differences in gene expression among individuals. Although we did not assess the genotypes of calves in this study, it is possible that environmental perturbations together with genetic differences (e.g. Single Nucleotide Polymorphisms, SNPs) affect transcriptome profiles. Taken together, our study is the first to explore the mucosal rectal transcriptome profiling, shedding light on the fundamental knowledge of factors potentially affecting the rectal mucosal transcriptome.

Furthermore, we found that rectal changes in the transcriptome profile were influenced by the production of stx2a. The wild type STEC O157 (stx2a-) which has an insertion sequence in the stx2a prophage has been reported to result in lower rectal antibody responses to *E. coli* O157 antigens (*i.e.,* H7, Tir, Intimin) as compared to challenge with a stx2a^{+} strain of *E. coli* O157 (Fitzgerald et al., 2019). However, we found that stxa2^{-} *E. coli* O157 colonization inhibited host responses at the transcriptome level when calves become SS and these decreased responses were then elevated when shedding levels returned to prechallenge levels. This shift in host responses in relation to shedding phenotype, suggests that the host transcriptome varies with the fecal shedding of stxa2^{-} *E. coli* O157. Among these pathways, we found serine activity-related functions were inhibited at peak fecal shedding of STEC in stx2a- O157 challenged calves (*i.e.* serine-type endopeptidase inhibitor activity). Serine biosynthesis is a major metabolic pathway in *E.coli,* which can be used for protein synthesis and as a precursor for the biosynthesis of glycine (Stover et al., 1992). However, the inhibitor activity of serine-type endopeptidase was inhibited based on functional annotations in the host, suggesting that serine production should be elevated and therefore could be beneficial for bacterial survival and proliferation.

Additionally, only one KEGG pathway (based on downregulated DEGs), the IL-17 signaling pathway, was enriched in WT at T2. The IL-17 (the interleukin 17 family) is a family of pro-inflammatory cystine knot cytokines, that can promote chemokine production so as to recruit monocytes and neutrophils to the site of inflammation (Korn et al., 2009; Cua and Tato, 2010). As this pathway was downregulated in WT-T2, it suggests that stx2a- STEC O157

colonization in RAJ epithelium could inhibit the host anti-inflammatory effect. However, this pathway was not enriched in RE-T2, suggesting that the production of stx2a could inhibit host IL-17 signaling. Summarization of the network of overlapping GO terms in the WT- T2 group revealed that functions were clustered into two major modules, one related to the extracellular surface of host epithelium where *E.coli* O157 can colonize and interact with host cells (Figure 4.13). The other cluster is related to tissue integrity. Previous findings revealed that the integrity of tissue barrier and epithelial regeneration were impacted by *E.coli* O157 colonization in calves and mice (Roxas et al., 2010; Fitzgerald et al., 2019). *E.coli* O157 can induce localized effacement of microvilli causing attach/effacement lesions (A/E lesion), which are detrimental to the epithelium and cause a negative effect on tissue barrier integrity (Roxas et al., 2010). This finding suggests that tissue damage is necessary for stx2a- *E.coli* O157 colonization of the epithelium.

The S100 protein family (*S100A8, S100A9*) was downregulated in the IL-17 signaling pathway and influenced tissue barrier integrity in WT-T2. The S100 protein family is a family of calcium-binding cytosolic proteins serving as important regulators of host immunity in mammals (Xia et al., 2018). It interacts with membrane proteins to enable transcriptional regulation and DNA repair and can also serve as a signal for regulating immune homeostasis and inflammation (Xia et al., 2018). Previous research revealed that *S100A8* and *S100A9* involved in the IL-17 signaling pathway responded to the relative abundance of hindgut mucosa microbiota in SS. That is, *S100A8* was positively related to *Pseudomonadaceae*, and *S100A9* was negatively related to *Ruminococcaceae* (Wang et al., 2018). In addition to these two genes,

S100A12 which has also been shown to be influenced by hindgut mucosa microbiota in SS (positive interaction with *Lachnospiraceae*) (Wang et al., 2018), was downregulated with regard to its role in RAGE receptor binding, extracellular region, and antimicrobial humoral response in WT-T2. In addition to the antimicrobial functions of *S100* family genes, these genes are also involved in tissue repair (Xia et al., 2018). A mouse model of angiotensin-induced cardiac damage revealed that *S100A8/A9* released by granulocytes could upregulate pro-inflammatory genes and induce the release of cytokines and chemokines, resulting in myocardial tissue inflammation and barrier damage (Wu et al., 2014). Therefore, *S100* family genes (*i.e., S100A8/9/12*) are vital genes for stx2a- *E.coli* O157 colonization and host responses, which contribute to both inflammatory responses and barrier integrity, with the potential to act as biomarkers of *E.coli O157* colonization and host immunity.

The tug of war between stx2a-*E.coli* O157 and the host exists, where the host defends stx2a-*E.coli* O157 colonization and attempts to eliminate this bacterium. Particularly, the expression of peptidase inhibitor was downregulated in WT-T2, resulting in increased production of peptidase. Peptidases are enzymes capable of cleaving proteins and are widely distributed on the cell surface and function in immune responses such as peptide-mediated inflammation and T-cell activation (Velden and Hulsmann 1999). The inhibition of peptidase inhibitors suggests the host may proactively produce peptidase in an effort to kill stx2a- *E.coli* O157. It was noticeable that the tug between the host and *E.coli* O157 continued until the calves returned to NS from SS. We found that at WT-T2, most of the inhibited functions were subsequently upregulated, suggesting that host functions were enhanced that led to a reduction

in stx2a-*E.coli* O157 shedding. The functions enriched by upregulated DEGs clustered into two major modules, similar to the clustering results at T2: an extracellular region and barrier integrity (Figure 4.14). While one KEGG pathway, the Wnt signaling pathway and one Reactome pathway, the IL-36 signaling pathway were enriched as a result of upregulated DEGs in WT-T5. The Wnt signaling pathway has been shown to have anti-inflammatory activity that can limit the survival of bacteria (Ma and Hottiger, 2016). While the IL-36 signaling pathway is critical for integrating host innate and adaptive immune responses for host protection of enteropathogenic bacteria colonization (Ngo et al., 2020). A study challenged mice deficient in IL-36R (receptor for IL-36) with *Citrobacter rodentium* (*C. rodentium*, a good *in vivo* model for A/E lesion causing bacteria) and found that expression of IL-22 and other antimicrobial proteins (AMPs) was decreased, while intestinal barrier integrity and resistance to bacterial colonization were impaired (Ngo et al., 2020). These responses did not negatively impact the health of mice. These findings suggest that the host protective role of IL-36R provides barrier protection from intestinal damage and enteric bacterial colonization as a result of the induction of IL-22 and AMPs.

Unlike responses to stx2a- *E.coli* O157 challenged calves at T2, stx2a+ *E.coli* O157 challenged calves (RE) did not respond to bacterial colonization during the peak fecal shedding stage, rather host responses were not altered and were similar to unchallenged calves. However, transcriptional responses elicited by stx2a- and stx2a+ *E.coli* O157 infection were similar at T5. Although the molecular mechanisms as to how the production of stx2a alters host responses are unclear, several studies suggest that stx2a plays a critical role in regulating host immunities

(Hoffman et al., 2006; Fitzgerald et al., 2019). A previous study in calves challenged with stx2+ or stx2- O157 for the first time and stx2+ O157 in a second inoculation, found that calves initially inoculated with stx2- O157 developed a higher magnitude of lymphoproliferative responses, suggesting that stx2 may suppress RAJ epithelial immune responses in cattle (Hoffman et al., 2006). However, the role of stx2 in host immune responses has not been consistent, as another study suggested that it did not suppress host immune responses in calves (Fitzgerald et al., 2019). The researchers tested antibody responses to stx2a+ O157 colonization in calves and found increased antibody levels of H7, Tir, EspA, and intimin-specific IgA compared to stx2a- O157 challenged calves (Fitzgerald et al., 2019). The role of stx2a production in O157 on host responses remains unclear but based on our results we developed a model of host defensive responses after inoculation with STEC O157 (Figure 4.15). For calves challenged with stx2a- O157, host responses were extensively inhibited at T2 and then enhanced at T5. While for calves challenged with stx2a+ O157, host responses remained unchanged at T2, subsequently decreased and then increased before reaching a similar level compared to stx2a+ O157 calves at T5.

In addition, we observed transcripts divergent in abundance (high and low) are functionally different. The types of highly abundant transcripts were both calf age- and challenge-independent. These functions were closely related to the maintenance of host homeostasis and host housekeeping functions. For instance, functions were related to stress (BAG2 signaling pathway, (Qin et al., 2016)), mRNA translation (EIF2 signaling pathway, (Kimball, 1999; Shrestha et al., 2012)) and anti-mRNA Degradation (Inhibition of ARE-

Mediated mRNA Degradation Pathway, (Toeuf et al., 2018)) and cell metabolism (Mitochondrial Dysfunction and Oxidative Phosphorylation (Wilson, 2017; Zhunina et al., 2021)). Since our challenge trial occurred over 29 days from pre- to post-challenge, such short-term time differences may not significantly affect host RAJ function. However, the function of highly abundant transcripts reflected the active metabolism of the RAJ mucosa in young calves and that challenge of strain-specific O157 did not alter host housekeeping functions. Interestingly, these functions were not identified as the top physiological functions at the RAJ mucosa of adult cattle (Wang et al., 2016), suggesting that expression of genes at the RAJ maybe age dependent.

Furthermore, our study revealed that transcripts with low abundance in all CT, WT, RE from T1 to T5 shared similar functions associated with neuronal processes and were challenge independent. Both Axonal Guidance signaling and CREB signaling in neurons were found to be in common from pre- to post-challenge in CT, WT, and RE groups, which are functionally related to complex neuronal development and synaptogenesis (Morris et al., 2006; Sakamoto et al., 2011; Russell and Bashaw, 2018). The rectum supports bowel movements that can eliminate stool from the host body (termed as defecation), a process that is controlled by parasympathetic activities (Shafik et al., 2002, 2003; Browning and Travagli, 2021; Nakata et al., 2022). The function of synaptogenesis ensures the formation of synapses, which anecessary component for defecation (Shafik et al., 2003; Russell and Bashaw, 2018), highlighting that low abundant transcripts are related to host housekeeping functions. Interestingly, we found that Cardiac Hypertrophy Signaling (Enhanced) is also associated with CT, WT, and RE groups and was

enriched by featured low abundant transcripts from pre to post challenge. This function relates to cardiac disease and is presumably not related to rectal function. While a recent study suggested the occurrence of a potential gut-heart axis, where products produced by intestinal microbiota (*i.e.,* trimethylamine N-oxide, short-chain fatty acids) could be resorbed in the intestine and distributed via the circulation and influence cardiovascular disease (Kamo et al., 2017; Bartolomaeus et al., 2020). Through such a process, intestinal microbiota and host expression in the intestinal tract could play an important role in heart function (Kamo et al., 2017; Bartolomaeus et al., 2020). Further studies are needed to identify how host rectum-hindgut microbiota interactions may influence the gut-heart axis. Through the functional analysis of high- and low-abundant transcripts in the calf rectum, we found that these functions are not sensitive to STEC O157 colonization, but rather related to fundamental biochemistry processes and housekeeping functions.

The current definition of "lower" abundance of transcripts may largely reflect our ability for transcript detection rather than the biological significance of the detected genes. However, we adopted an advanced machine learning based approach through a self-organizing map (SOM) to select important low abundance genes in our study. A common computational approach for clustering is hierarchical clustering, where data are forced into a strict hierarchy of nested subsets and the closest pair of points is grouped. However, strict phylogenetic trees used in hierarchical clustering are not suitable for similar expression levels and can cluster data points based on local decisions with no means to evaluate clustering patterns (Tamayo et al., 1999). The SOM provides a thorough summary of a massive data set by extracting the most prominent

patterns while retaining structural information and therefore is more efficient at selecting functionally important genes that are similar in terms of their expression levels (Kohonen, 1990; Tamayo et al., 1999; Reusch et al., 2005; Astel et al., 2007). For example, by using SOM, researchers were able to group more than 6,000 yeast genes into several clusters, involved in a diversity of functions, providing a fundamental understanding of yeast functional genomes (Törönen et al., 1999). Although our research did not identify different functions of low-abundant transcripts, we still highlighted the possibility of adopting SOM to filter out biological-importance transcripts from high-throughput RNA-seq reads.

In the current study, we did not observe host responses directly related to innate and adaptive immunities. Possibly, the approach we used for pathway identification limited our understanding of variations in the host transcriptome. Identification of DEGs and functional annotation using DEGs belong to the overall representation analysis, which determines whether genes from pre-defined gene sets (DEGs) are present more than expected (Maleki et al., 2020; Karp et al., 2021). However, such an approach has several drawbacks: the choice of thresholds for significant pathways is artificial and might affect the result of downstream analysis (Maleki et al., 2020) and it is incapable of detecting biologically important but low expression signals (Maleki et al., 2020). While another method, gene set enrichment analysis (GSEA) considers all genes, not only those that are DEGs (Subramanian et al., 2005; Reimand et al., 2019). For instance, by using GSEA, researchers found that LINC00263 (one of the most dysregulated lncRNAs in lung adenocarcinomas) was closely related to the NF-κB signaling pathway, while such relations were not directly observed in GO and KEGG enrichments (Liu et al., 2020).

Regardless, our study suggested that the dynamic interactions between host and STEC O157 colonization, of which host functions involved the ability to combat O157 colonization and maintenance of tissue barrier integrity were altered.

4.5 Conclusion

Taken together, our study identified that rectal mucosal transcriptomic and functional responses vary in terms of the production of stx2a in STEC. We also identified functional shifts of transcripts in terms of their abundance. We found that high- and low-abundant transcripts functioned in the host's biochemical processes and its metabolism with little relation to calf age or challenge with O157. Calves challenged with stx2a- O157 exhibited inhibited responses 7 days post challenge, while these functions were enhanced 26 days post challenge when calves recovered from SS. Varied functions were grouped into two clusters: the extracellular region where host cells combat O157 colonization and alter host signaling pathways, and tissue barrier integrity. We also identified that genes involved in the *S100A* family were significantly related to variation in host responses and may serve as biomarkers of host responses. For calves challenged with stx2a+ O157, we did not observe any host responses 7 days post challenge while they exhibited enhanced responses 26 days post challenge in a manner similar to that in stx2a- challenged calves. Future studies are needed to elucidate host-STEC interactions by analyzing host transcriptomes in samples collected from T3 and T4 and by applying methods of including genome wide association studies (GWAS) that determine potential genetic variations related to host-STEC interactions. By doing so, we could determine the mechanisms

of STEC O157 colonization and stx2a expression on host responses from a comprehensive perspective.

4.6 References

Araki, R., R. Fukumura, N. Sasaki, Y. Kasama, N. Suzuki, H. Takahashi, Y. Tabata, T. Saito, and M. Abe. 2006. More Than 40,000 Transcripts, Including Novel and Noncoding Transcripts, in Mouse Embryonic Stem Cells. STEM CELLS 24:2522–2528. doi:10.1634/stemcells.2006-0005.

Astel, A., S. Tsakovski, P. Barbieri, and V. Simeonov. 2007. Comparison of self-organizing maps classification approach with cluster and principal components analysis for large environmental data sets. Water Res 41:4566–4578. doi:10.1016/j.watres.2007.06.030.

Bartolomaeus, H., V. McParland, and N. Wilck. 2020. Darm-Herz-Achse. Herz 45:134–141. doi:10.1007/s00059-020-04897-0.

Bizama, C., F. Benavente, E. Salvatierra, A. Gutiérrez-Moraga, J.A. Espinoza, E.A. Fernández, I. Roa, G. Mazzolini, E.A. Sagredo, M. Gidekel, and O.L. Podhajcer. 2014. The low-abundance transcriptome reveals novel biomarkers, specific intracellular pathways and targetable genes associated with advanced gastric cancer. Int. J. Cancer 134:755–764. doi:10.1002/ijc.28405.

Browning, K.N., and R.A. Travagli. 2021. Comprehensive Physiology. Compr. Physiol. 4:1339–1368. doi:10.1002/cphy.c130055.

Castro, V.S., R.C.T. Carvalho, C.A. Conte-Junior, and E.E.S. Figuiredo. 2017. Shiga-toxin Producing Escherichia coli: Pathogenicity, Supershedding, Diagnostic Methods, Occurrence,

and Foodborne Outbreaks. Compr Rev Food Sci F 16:1269–1280. doi:10.1111/1541-4337.12302.

Chase-Topping, M., D. Gally, C. Low, L. Matthews, and M. Woolhouse. 2008. Super-shedding and the link between human infection and livestock carriage of Escherichia coli O157. Nat Rev Microbiol 6:904–912. doi:10.1038/nrmicro2029.

Chess, A. 2013. Random and Non-Random Monoallelic Expression. Neuropsychopharmacology 38:55–61. doi:10.1038/npp.2012.85.

Cobbold, R., and P. Desmarchelier. 2000. A longitudinal study of Shiga-toxigenic Escherichia coli (STEC) prevalence in three Australian dairy herds. Vet Microbiol 71:125–137. doi:10.1016/s0378-1135(99)00173-x.

Cua, D.J., and C.M. Tato. 2010. Innate IL-17-producing cells: the sentinels of the immune system. Nat Rev Immunol 10:479–489. doi:10.1038/nri2800.

Das, G., M. Chattopadhyay, and S. Gupta. 2015. A Comparison of Self-organising Maps and Principal Components Analysis. Int J Market Res 58:815–834. doi:10.2501/ijmr-2016-039.

Fitzgerald, S.F., A.E. Beckett, J. Palarea-Albaladejo, S. McAteer, S. Shaaban, J. Morgan, N.I. Ahmad, R. Young, N.A. Mabbott, L. Morrison, J.L. Bono, D.L. Gally, and T.N. McNeilly. 2019. Shiga toxin sub-type 2a increases the efficiency of Escherichia coli O157 transmission between animals and restricts epithelial regeneration in bovine enteroids. Plos Pathog 15:e1008003. doi:10.1371/journal.ppat.1008003.

Fraser, M.E., M. Fujinaga, M.M. Cherney, A.R. Melton-Celsa, E.M. Twiddy, A.D. O'Brien, and M.N.G. James. 2004. Structure of Shiga Toxin Type 2 (Stx2) from Escherichia coli O157:H7*. J Biol Chem 279:27511–27517. doi:10.1074/jbc.m401939200.

He, Z., A. Fischer, Y. Song, M. Steele, and L.L. Guan. 2018. Genome wide transcriptome analysis provides bases on colonic mucosal immune system development affected by colostrum feeding strategies in neonatal calves. BMC Genom. 19:635. doi:10.1186/s12864-018-5017-y.

Hoffman, M.A., C. Menge, T.A. Casey, W. Laegreid, B.T. Bosworth, and E.A. Dean-Nystrom. 2006. Bovine Immune Response to Shiga-Toxigenic Escherichia coli O157:H7. Clin Vaccine Immunol 13:1322–1327. doi:10.1128/cvi.00205-06.

Kamo, T., H. Akazawa, J. Suzuki, and I. Komuro. 2017. Novel Concept of a Heart-Gut Axis in the Pathophysiology of Heart Failure. Korean Circ. J. 47:663–669. doi:10.4070/kcj.2017.0028.

Karp, P.D., P.E. Midford, R. Caspi, and A. Khodursky. 2021. Pathway size matters: the influence of pathway granularity on over-representation (enrichment analysis) statistics. BMC Genom. 22:191. doi:10.1186/s12864-021-07502-8.

Kim, Y.C., Y.-C. Jung, Z. Xuan, H. Dong, M.Q. Zhang, and S.M. Wang. 2006. Pan-genome isolation of low abundance transcripts using SAGE tag. FEBS Lett. 580:6721–6729. doi:10.1016/j.febslet.2006.11.013.

Kimball, S.R. 1999. Eukaryotic initiation factor eIF2. The Int. J. Biochem. Cell Biology 31:25–29. doi:10.1016/s1357-2725(98)00128-9.

Kohonen, T. 1990. The self-organizing map. P Ieee 78:1464–1480. doi:10.1109/5.58325.

Korn, T., E. Bettelli, M. Oukka, and V.K. Kuchroo. 2009. IL-17 and Th17 Cells. Immunology 27:485–517. doi:10.1146/annurev.immunol.021908.132710.

Liu, S., W. Lai, Y. Shi, N. Liu, L. Ouyang, Z. Zhang, L. Chen, X. Wang, B. Qian, D. Xiao, Q. Yan, Y. Cao, S. Liu, and Y. Tao. 2020. Annotation and cluster analysis of long noncoding RNA linked to male sex and estrogen in cancers. npj Precis. Oncol. 4:5. doi:10.1038/s41698-020-0110-5.

Lomana, A.L.G. de, U. Kusebauch, A.V. Raman, M. Pan, S. Turkarslan, A.P.R. Lorenzetti, R.L. Moritz, and N.S. Baliga. 2020. Selective Translation of Low Abundance and Upregulated Transcripts in Halobacterium salinarum. mSystems 5:e00329-20. doi:10.1128/msystems.00329-20.

Ma, B., and M.O. Hottiger. 2016. Crosstalk between Wnt/β-Catenin and NF-κB Signaling Pathway during Inflammation. Frontiers Immunol. 7:378. doi:10.3389/fimmu.2016.00378.

Maleki, F., K. Ovens, D.J. Hogan, and A.J. Kusalik. 2020. Gene Set Analysis: Challenges, Opportunities, and Future Research. Frontiers Genet. 11:654. doi:10.3389/fgene.2020.00654.

McCabe, E., C.M. Burgess, D. Lawal, P. Whyte, and G. Duffy. 2019. An investigation of shedding and super-shedding of Shiga toxigenic Escherichia coli O157 and E. coli O26 in cattle presented for slaughter in the Republic of Ireland. Zoonoses Public Hlth 66:83–91. doi:10.1111/zph.12531.

Melton-Celsa, A.R. 2014. Shiga Toxin (Stx) Classification, Structure, and Function. Microbiol Spectr 2. doi:10.1128/microbiolspec.ehec-0024-2013.

Mills, K.H.G. 2023. IL-17 and IL-17-producing cells in protection versus pathology. Nat. Rev. Immunol. 23:38–54. doi:10.1038/s41577-022-00746-9.

Morris, J.S., T. Stein, M. Pringle, C.R. Davies, S. Weber-Hall, R.K. Ferrier, A.K. Bell, V.J. Heath, and B.A. Gusterson. 2006. Involvement of axonal guidance proteins and their signaling partners in the developing mouse mammary gland. J. Cell. Physiol. 206:16–24. doi:10.1002/jcp.20427.

Munns, K.D., L.B. Selinger, K. Stanford, L. Guan, T.R. Callaway, and T.A. McAllister. 2015. Perspectives on Super-Shedding of Escherichia coli O157:H7 by Cattle. Foodborne Pathog Dis 12:89–103. doi:10.1089/fpd.2014.1829.

Nakata, R., F. Tanaka, N. Sugawara, Y. Kojima, T. Takeuchi, M. Shiba, K. Higuchi, and Y. Fujiwara. 2022. Analysis of autonomic function during natural defecation in patients with irritable bowel syndrome using real-time recording with a wearable device. PLOS ONE 17:e0278922. doi:10.1371/journal.pone.0278922.

Ngo, V.L., H. Abo, M. Kuczma, E. Szurek, N. Moore, O. Medina-Contreras, A. Nusrat, D. Merlin, A.T. Gewirtz, L. Ignatowicz, and T.L. Denning. 2020. IL-36R signaling integrates innate and adaptive immune-mediated protection against enteropathogenic bacteria. Proc. National Acad. Sci. 117:27540–27548. doi:10.1073/pnas.2004484117.

Qin, L., J. Guo, Q. Zheng, and H. Zhang. 2016. BAG2 structure, function and involvement in disease. Cell. Mol. Biology Lett. 21:18. doi:10.1186/s11658-016-0020-2.

Reimand, J., R. Isserlin, V. Voisin, M. Kucera, C. Tannus-Lopes, A. Rostamianfar, L. Wadi, M. Meyer, J. Wong, C. Xu, D. Merico, and G.D. Bader. 2019. Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. Nat Protoc 14:482–517. doi:10.1038/s41596-018-0103-9.

Reusch, D.B., R.B. Alley, and B.C. Hewitson. 2005. Relative Performance of Self-Organizing Maps and Principal Component Analysis in Pattern Extraction from Synthetic Climatological Data. Polar Geogr 29:188–212. doi:10.1080/789610199.

Roxas, J.L., A. Koutsouris, A. Bellmeyer, S. Tesfay, S. Royan, K. Falzari, A. Harris, H. Cheng, K.J. Rhee, and G. Hecht. 2010. Enterohemorrhagic E. coli alters murine intestinal epithelial tight junction protein expression and barrier function in a Shiga toxin independent manner. Lab. Investig. 90:1152–1168. doi:10.1038/labinvest.2010.91.

Russell, S.A., and G.J. Bashaw. 2018. Axon guidance pathways and the control of gene expression. Dev. Dyn. 247:571–580. doi:10.1002/dvdy.24609.

Ryu, J.-H., S. Kim, J. Park, and K.-S. Choi. 2020. Characterization of virulence genes in Escherichia coli strains isolated from pre-weaned calves in the Republic of Korea. Acta Vet Scand 62:45. doi:10.1186/s13028-020-00543-1.

Sakamoto, K., K. Karelina, and K. Obrietan. 2011. CREB: a multifaceted regulator of neuronal plasticity and protection. J. Neurochem. 116:1–9. doi:10.1111/j.1471-4159.2010.07080.x.

Shafik, A., O. El-Sibai, and I. Ahmed. 2002. Parasympathetic extrinsic reflex: Role in defecation mechanism. World J. Surg. 26:737–740. doi:10.1007/s00268-002-6285-9.

Shafik, A.E., A. Shafik, O. EI-Sibai, and I. Ahmed. 2003. Role Of Sympathetic Innervation In The Defecation Mechanism: A Novel Concept Of Its Function. The J. Spinal Cord Medicine 26:150–154. doi:10.1080/10790268.2003.11753676.

Shrestha, N., W. Bahnan, D.J. Wiley, G. Barber, K.A. Fields, and K. Schesser. 2012. Eukaryotic Initiation Factor 2 (eIF2) Signaling Regulates Proinflammatory Cytokine Expression and Bacterial Invasion*. J. Biological Chem. 287:28738–28744. doi:10.1074/jbc.m112.375915.

Sriram, K., S.Z. Wiley, K. Moyung, M.W. Gorr, C. Salmerón, J. Marucut, R.P. French, A.M. Lowy, and P.A. Insel. 2019. Detection and Quantification of GPCR mRNA: An Assessment and Implications of Data from High-Content Methods. ACS Omega 4:17048–17059. doi:10.1021/acsomega.9b02811.

Stover, P., M. Zamora, K. Shostak, M. Gautam-Basak, and V. Schirch. 1992. Escherichia coli serine hydroxymethyltransferase. The role of histidine 228 in determining reaction specificity.. J. Biological Chem. 267:17679–17687. doi:10.1016/s0021-9258(19)37096-6.

Subramanian, A., P. Tamayo, V.K. Mootha, S. Mukherjee, B.L. Ebert, M.A. Gillette, A. Paulovich, S.L. Pomeroy, T.R. Golub, E.S. Lander, and J.P. Mesirov. 2005. Gene set

enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc National Acad Sci 102:15545–15550. doi:10.1073/pnas.0506580102.

Tamayo, P., D. Slonim, J. Mesirov, Q. Zhu, S. Kitareewan, E. Dmitrovsky, E.S. Lander, and T.R. Golub. 1999. Interpreting patterns of gene expression with self-organizing maps: Methods and application to hematopoietic differentiation. Proc. National Acad. Sci. 96:2907–2912. doi:10.1073/pnas.96.6.2907.

Toeuf, B. de, R. Soin, A. Nazih, M. Dragojevic, D. Jurėnas, N. Delacourt, L.V. Ngoc, A. Garcia-Pino, V. Kruys, and C. Gueydan. 2018. ARE-mediated decay controls gene expression and cellular metabolism upon oxygen variations. Sci. Reports 8:5211. doi:10.1038/s41598-018-23551-8.

Törönen, P., M. Kolehmainen, G. Wong, and E. Castrén. 1999. Analysis of gene expression data using self-organizing maps. FEBS Lett. 451:142–146. doi:10.1016/s0014-5793(99)00524-4.

Vedell, P.T., K.L. Svenson, and G.A. Churchill. 2011. Stochastic variation of transcript abundance in C57BL/6J mice. BMC Genom. 12:167. doi:10.1186/1471-2164-12-167.

VELDEN, V.D., and A.R. HULSMANN. 1999. Peptidases: structure, function and modulation of peptide-mediated effects in the human lung. Clin. Exp. Allergy 29:445–456. doi:10.1046/j.1365-2222.1999.00462.x.

Wang, O., G. Liang, T.A. McAllister, G. Plastow, K. Stanford, B. Selinger, and L.L. Guan. 2016. Comparative Transcriptomic Analysis of Rectal Tissue from Beef Steers Revealed Reduced Host Immunity in Escherichia coli O157:H7 Super-Shedders. Plos One 11:e0151284. doi:10.1371/journal.pone.0151284.

Wang, O., T.A. McAllister, G. Plastow, K. Stanford, B. Selinger, and L.L. Guan. 2017. Host mechanisms involved in cattle Escherichia coli O157 shedding: a fundamental understanding for reducing foodborne pathogen in food animal production. Sci Rep-uk 7:7630. doi:10.1038/s41598-017-06737-4.

Wang, O., T.A. McAllister, G. Plastow, K. Stanford, B. Selinger, and L.L. Guan. 2018. Interactions of the Hindgut Mucosa-Associated Microbiome with Its Host Regulate Shedding of Escherichia coli O157:H7 by Cattle. Appl Environ Microb 84:e01738-17. doi:10.1128/aem.01738-17.

Welle, S., K. Bhatt, and C.A. Thornton. 2000. High-abundance mRNAs in human muscle: comparison between young and old. J. Appl. Physiol. 89:297–304. doi:10.1152/jappl.2000.89.1.297.

Werner, M.S., B. Sieriebriennikov, N. Prabh, T. Loschko, C. Lanz, and R.J. Sommer. 2018. Young genes have distinct gene structure, epigenetic profiles, and transcriptional regulation. Genome Res. 28:1675–1687. doi:10.1101/gr.234872.118.

Wilson, D.F. 2017. Oxidative phosphorylation: regulation and role in cellular and tissue metabolism. The J. Physiol. 595:7023–7038. doi:10.1113/jp273839.

Wu, Y., Y. Li, C. Zhang, X. A, Y. Wang, W. Cui, H. Li, and J. Du. 2014. S100a8/a9 Released by CD11b+Gr1+ Neutrophils Activates Cardiac Fibroblasts to Initiate Angiotensin II–Induced Cardiac Inflammation and Injury. Hypertension 63:1241–1250. doi:10.1161/hypertensionaha.113.02843.

Xia, C., Z. Braunstein, A.C. Toomey, J. Zhong, and X. Rao. 2018. S100 Proteins As an Important Regulator of Macrophage Inflammation. Frontiers Immunol. 8:1908. doi:10.3389/fimmu.2017.01908.

Xu, Y., E. Dugat-Bony, R. Zaheer, L. Selinger, R. Barbieri, K. Munns, T.A. McAllister, and L.B. Selinger. 2014. Escherichia coli O157:H7 Super-Shedder and Non-Shedder Feedlot Steers Harbour Distinct Fecal Bacterial Communities. Plos One 9:e98115. doi:10.1371/journal.pone.0098115.

Zhang, Q.-Y., Z.-B. Yan, Y.-M. Meng, X.-Y. Hong, G. Shao, J.-J. Ma, X.-R. Cheng, J. Liu, J. Kang, and C.-Y. Fu. 2021. Antimicrobial peptides: mechanism of action, activity and clinical potential. Mil. Méd. Res. 8:48. doi:10.1186/s40779-021-00343-2.

Zhunina, O.A., N.G. Yabbarov, A.V. Grechko, A.V. Starodubova, E. Ivanova, N.G. Nikiforov, and A.N. Orekhov. 2021. The Role of Mitochondrial Dysfunction in Vascular Disease,

Tumorigenesis, and Diabetes. Frontiers Mol. Biosci. 8:671908. doi:10.3389/fmolb.2021.671908.

4.7 Tables and figures

Table 4.1. Number of transcripts identified in CT, WT[1], and RE[2] group from pre- (T1) to post- (T5) challenge.

	CT			WT			RE		
	T1	T2	T5	T1	T2	T5	T1	T2	T5
Total[3]	21819	21835	21732	21695	21515	21577	21628	21837	21783
Core[4]	17853	17906	17742	18229	17804	17613	17082	17482	18110
Unique	304	316	275	362	307	286	247	343	317
Highly abundant[5]	5360	5634	5534	5049	5577	5417	5125	5354	5055

1 WT refers to calves challenged with STEC O157$^{stx2a\text{-}stx2c+}$ strain pre-challenge.

2 RE refers to calves challenged with STEC O157$^{stx2a+stx2c+}$ strain pre-challenge.

3 Total represents total transcripts identified in each group.

4 Core represents genes that were universally expressed in all samples in CT, or WT or RE group at certain time points

5 Highly abundant represents genes that accumulated to >90% total reads.

Table 4.2. Number of low abundant and featured low abundant transcripts[1] identified in CT, WT, and RE group from pre- (T1) to post- (T5) challenge.

Group	Time	No. low abundant transcripts	No. featured low abundant transcripts
CT	T1	5911	2236
	T2	5934	2024
	T5	5476	1952
WT	T1	6598	2502
	T2	6014	2101
	T5	5821	2114
RE	T1	5504	2079
	T2	5717	2134
	T5	6548	2355

1 Featured low abundant transcripts were identified using the self-organising map approach.

Table 4.3. The number of up- and down- regulated DEGs and GO terms enriched by DEGs in comparison of WT and CT, RE and CT, WT and RE groups at T2 and T5.

Group	Time	Up-regulated DEGs ; No. of GO terms [1]	Down-regulated DEGs; No. of GO terms [2]	Total DEGs
WT vs. CT	T2	39; 0	302; 26	341
	T5	195; 51	5; 0	200
RE vs. CT	T2	1; 0	0; 0	1
	T5	200; 54	6; 0	206
WT vs. RE	T2	6; 0	214; 45	220
	T5	2; 0	1; 0	3

1 The value before semicolon and after semicolon represent up-regulated DEGs and GO terms enriched by up-regulated DEGs, respectively.

2 The value before semicolon and after semicolon represent down-regulated DEGs and GO terms enriched by down-regulated DEGs, respectively.

Table 4.4. Downregulated DEGs involved in host-immune related functions

Genes	Log2 fold change	P-value	Terms
ENSBTAG00000046482	-1.6	0.04	
TMEM229B	-2.0	<0.01	
ENSBTAG00000054561	-2.1	<0.01	
ENSBTAG00000052579	-2.7	0.01	
UPK1B	-2.8	<0.01	response to bacterium
IL36G	-5.3	0.02	
S100A8	-6.2	<0.01	
IL36A	-6.5	<0.01	
LPO	-7.3	<0.01	
PGLYRP4	-2.7	0.01	
IL36RN	-2.9	0.04	
ENSBTAG00000014329	-2.9	<0.01	response to bacterium
S100A9	-5.6	<0.01	antimicrobial humoral response
SPINK5	-6.7	<0.01	humoral immune response
PGLYRP3	-6.7	0.01	
WFDC5	-7.0	<0.01	
S100A12	-3.5	<0.01	antimicrobial humoral response & humoral immune response
KRT1	-6.8	<0.01	humoral immune response

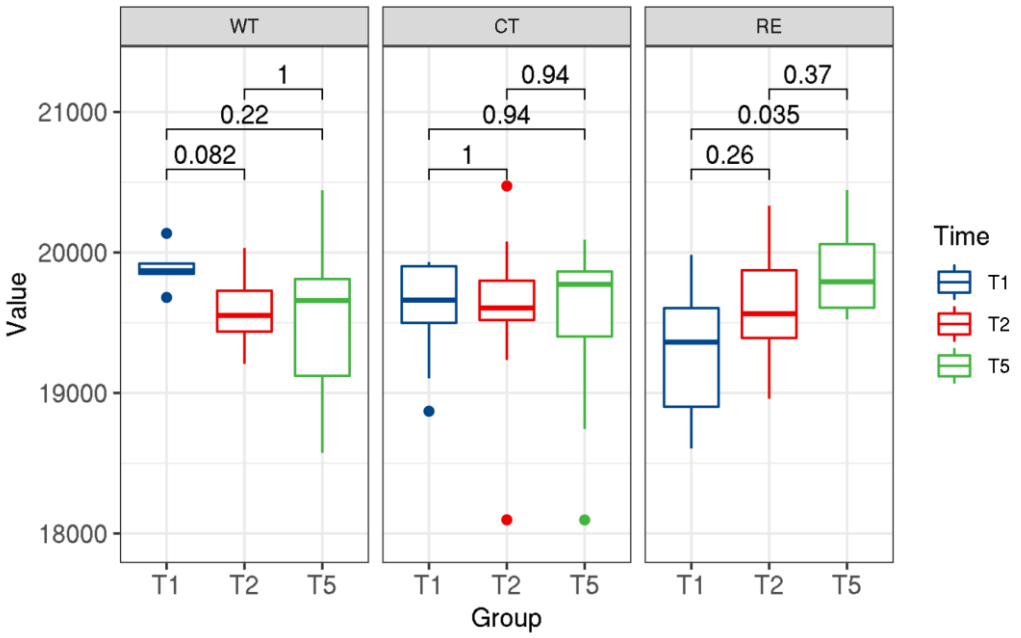


Figure 4.1. Number of paiedr-end RNA-seq reads for WT, CT, and RE groups from pre- (T1) to post- (T5) challenge.

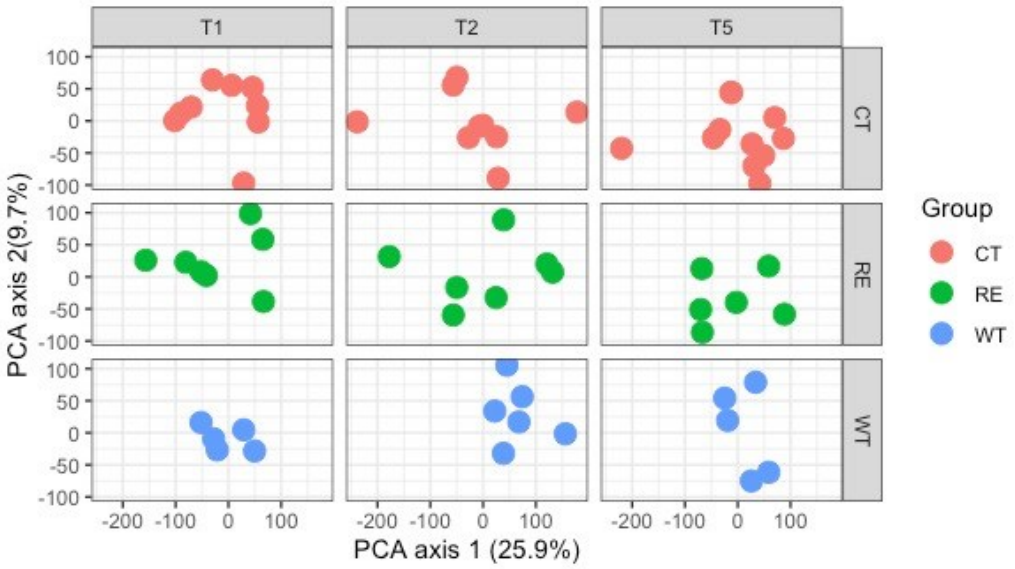
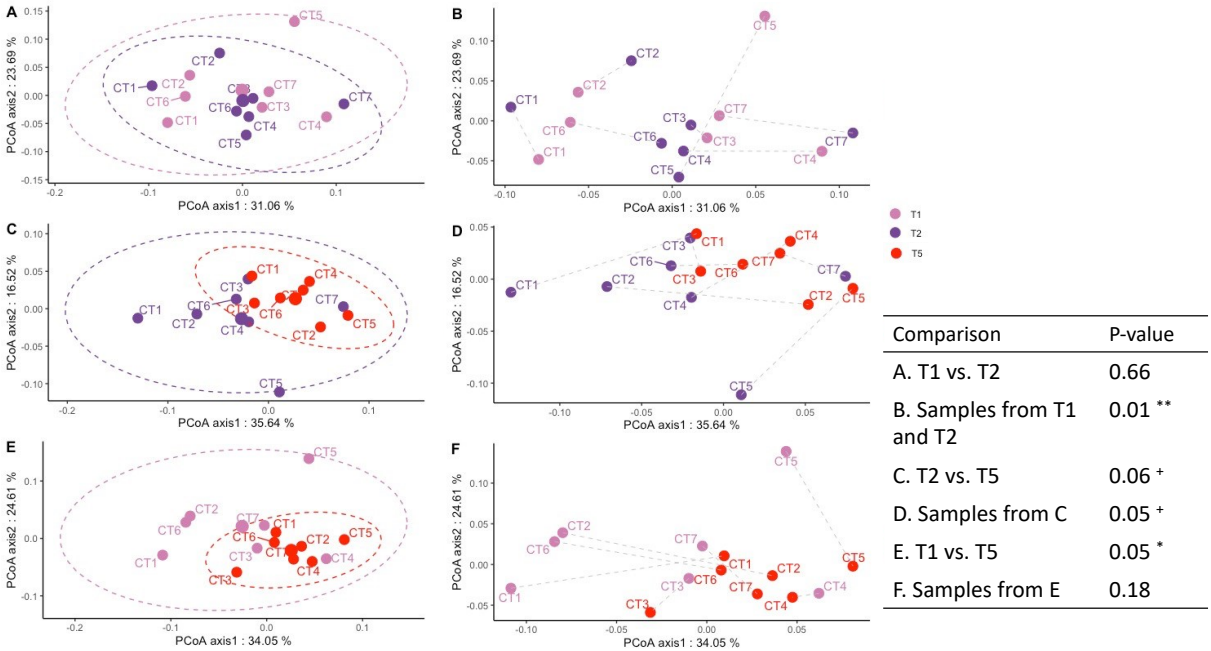


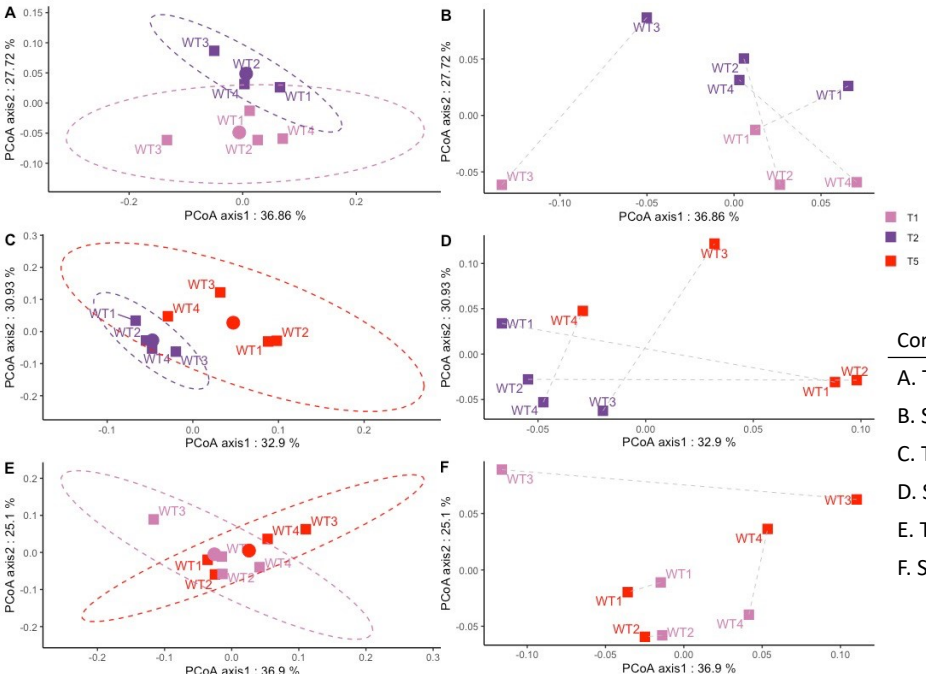
Figure 4.2. Principal component analysis (PCA) of host transcripts for CT, WT, and RE groups from pre- (T1) to post- (T5) challenge.



Comparison	P-value
A. T1 vs. T2	0.66
B. Samples from T1 and T2	0.01 **
C. T2 vs. T5	0.06 +
D. Samples from C	0.05 +
E. T1 vs. T5	0.05 *
F. Samples from E	0.18

Figure 4.3. Principal coordinate analyses (PCoA) of Bray-Curtis distances of host transcript similarities to the CT group.

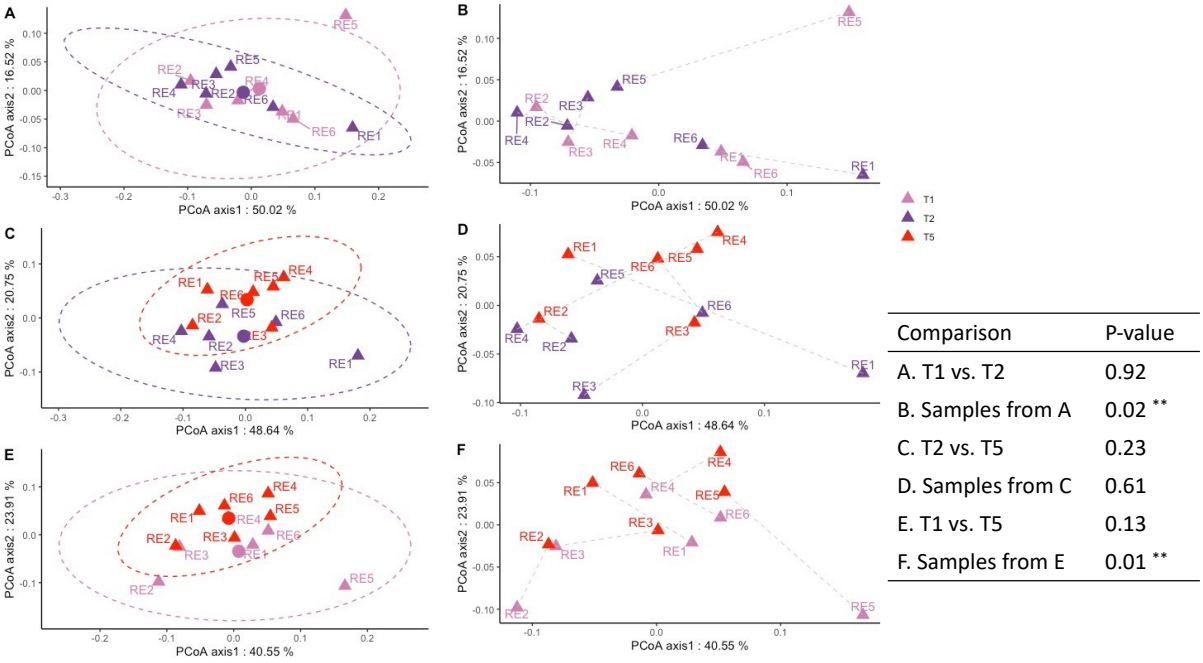
The distance was calculated between T1 and T2 (A), T2 and T5 (C), T1 and T5 (E) in the CT group. Each dot represents a sample and the dashed lines (B, D, F) indicate the shift of host transcripts representing individual responses in a sample from T1 to T2 (B), T2 to T5 (D), T1 to T5 (F). The PERMANOVA analysis assessed the significance of time-to time differnces (A,C,E) and pairwise individual variations (B,D,F) on microbial structural similarities with $P<0.05$ as a significance and $P \geq 0.05$ & <0.1 as a tendceny being significance.



Comparison	P-value
A. T1 vs. T2	0.05 +
B. Samples from A	0.29
C. T2 vs. T5	0.03 **
D. Samples from C	0.96
E. T1 vs. T5	0.25
F. Samples from E	0.41

Figure 4.4. Principal coordinate analyses (PCoA) of Bray-Curtis distances of host transcript similarities to the WT group.

The distance was calculated between T1 and T2 (A), T2 and T5 (C), T1 and T5 (E) in the WT group. Each dot represents a sample and the dashed lines (B, D, F) indicate that the shift of host transcripts representing individual responses in a sample from T1 to T2 (B), T2 to T5 (D), T1 to T5 (F). The PERMANOVA analysis assessed the significance of time-to time differences (A,C,E) and pairwise individual variations (B,D,F) on microbial structural similarities with $P<0.05$ as a significance and $P \geq 0.05$ & <0.1 as a tendceny being significance.



Comparison	P-value
A. T1 vs. T2	0.92
B. Samples from A	0.02 **
C. T2 vs. T5	0.23
D. Samples from C	0.61
E. T1 vs. T5	0.13
F. Samples from E	0.01 **

Figure 4.5. Principal coordinate analyses (PCoA) of Bray-Curtis distances of host transcript similarities to the RE group.

The distance was calculated between T1 and T2 (A), T2 and T5 (C), T1 and T5 (E) in the RE group. Each dot represents a sample and the dashed lines (B, D, F) indicate that the shift of host transcripts representing individual responses in a sample from T1 to T2 (B), T2 to T5 (D), T1 to T5 (F) ($P<0.05$ as a significance and $P \geq 0.05$ & <0.1 as a tend being significance). The PERMANOVA analysis assessed the significance of time-to time differences (A,C,E) and pairwise individual variations (B,D,F) on microbial structural similarities with $P<0.05$ as a significance and $P \geq 0.05$ & <0.1 as a tendceny being significance.

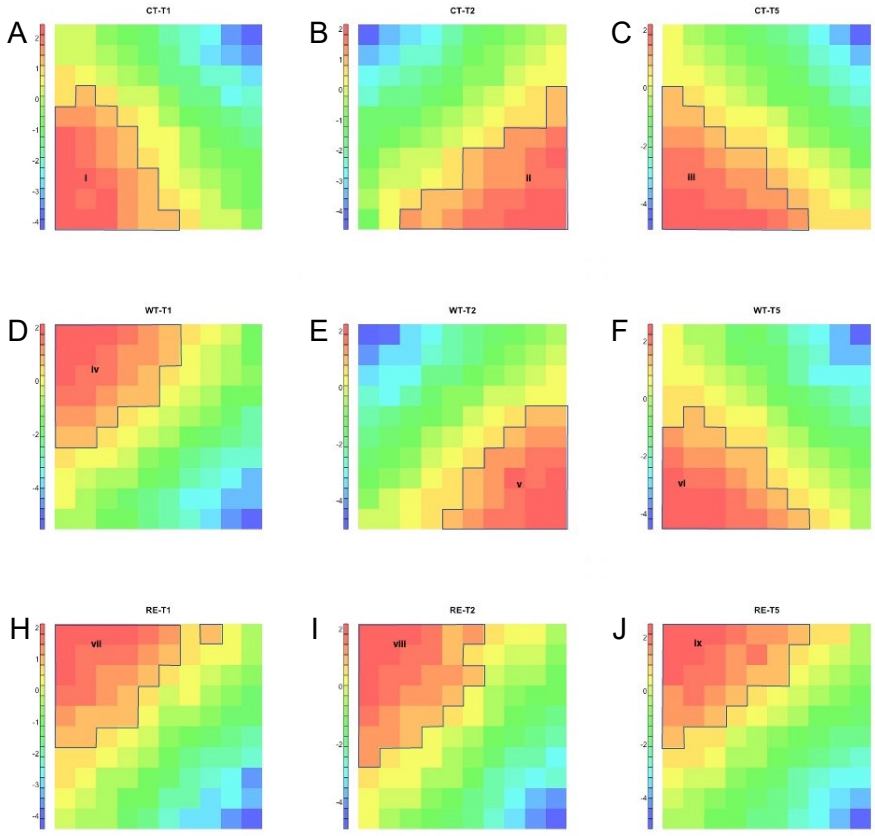


Figure 4.6. The self-organizing map-based clustering of low abundant transcripts.

The self-organizing map was performed for CT-T1 (A), CT-T2 (B), CT-T5 (C), WT-T1 (D), WT-T2 (E), WT-T5 (F), RE-T1 (G), RE-T2 (H), RE-T5 (I) with featured low-abundant transcripts being clusters I to IX.

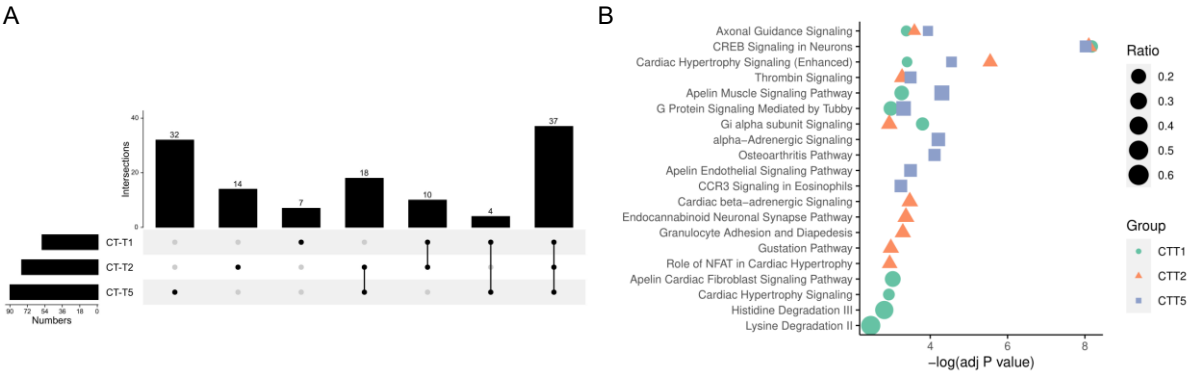


Figure 4.7. The functional analysis of featured low abundant transcripts idenfied from the self-organizing map in the CT group.

A. The upset plot showing overlapped GO functions enriched using featured low abundant transcripts idenfied from self-organizing map in the CT group from T1 to T5.

B. The top 10 most enriched GO functions enriched using featured low abundant transcripts idenfied from self-organizing map in the CT group at T1, T2, and T5, respectively.

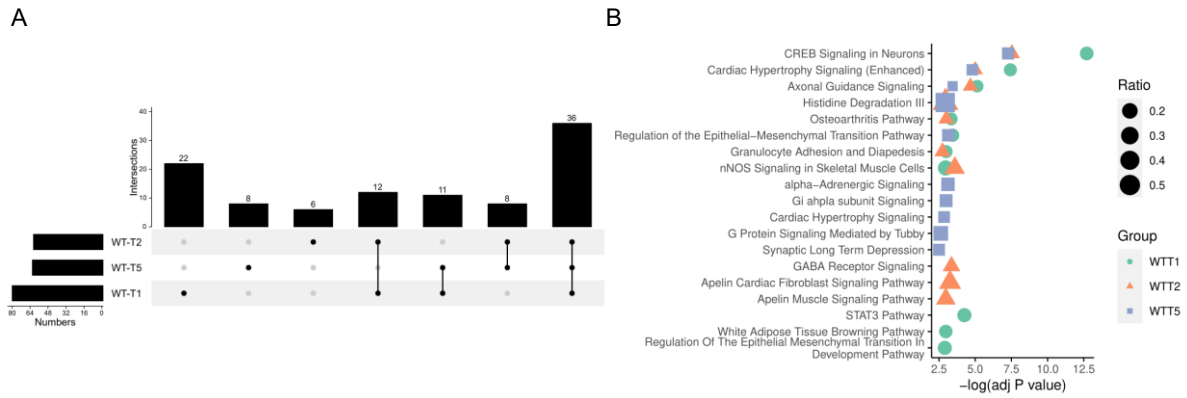


Figure 4.8. The functional analysis of featured low abundant transcripts idenfied from the self-organizing map in the WT group.

A. The upset plot showing overlapped GO functions enriched using featured low abundant transcripts idenfied from self-organizing map in the WT group from T1 to T5.

B. The top 10 most enriched GO functions enriched using featured low abundant transcripts idenfied from self-organizing map in the WT group at T1, T2, and T5, respectively.

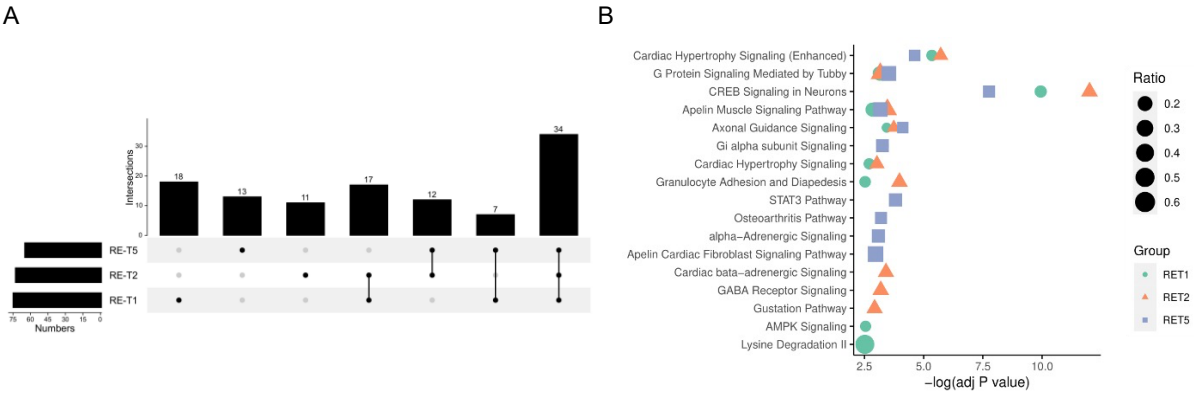


Figure 4.9. The functional analysis of featured low abundant transcripts idenfied from the self-organizing map in the RE group.

A. The upset plot showing overlapped GO functions enriched using featured low abundant transcripts idenfied from self-organizing map in the CT group from T1 to T5.

B. The top 10 most enriched GO functions enriched using featured low abundant transcripts idenfied from self-organizing map in the CT group at T1, T2, and T5, respectively.

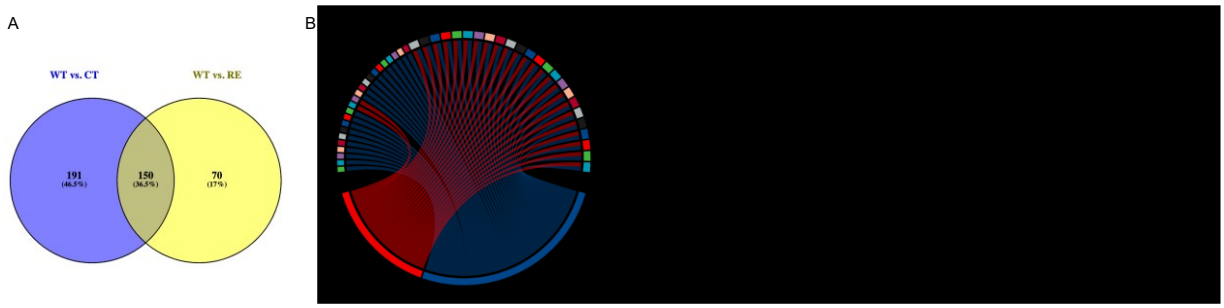


Figure 4.10. The overview of downregulated DEGs with enriched GO functions in the WT- T2 group.

A. The Venn plot showing overlapped downregulated DEGs in the WT- T2 group compared to CT-T2 and RE-T2 groups, respectively.

B. The circus plot showing overlapped GO functions enriched by downregulated DEGs in the WT- T2 group compared to CT-T2 and RE-T2 groups. No. 1 to 10 and 13 to 23 represent GO functions that were specific to the WT- T2 group compared to the RE-T2 group. No. 11 to 12 represent GO functions that were specific to the WT- T2 group compared to the CT-T2 group. No. 24 to 47 represent overlapped GO functions in the WT- T2 group compared to both CT-T2 and RE-T2 groups.

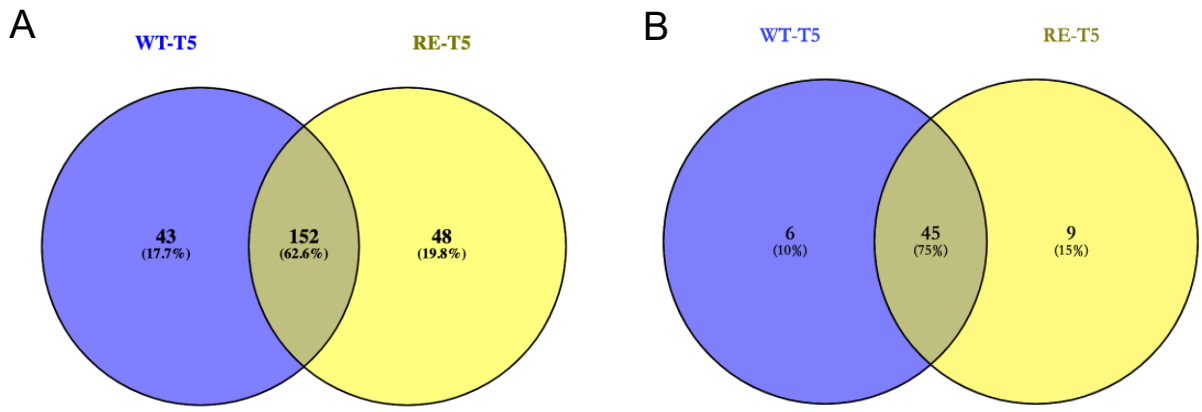


Figure 4.11. The Venn plot showing overlapped upregulated DEGs (A) with enriched GO functions (B) in comparison of WT-T5 and RE-T5.

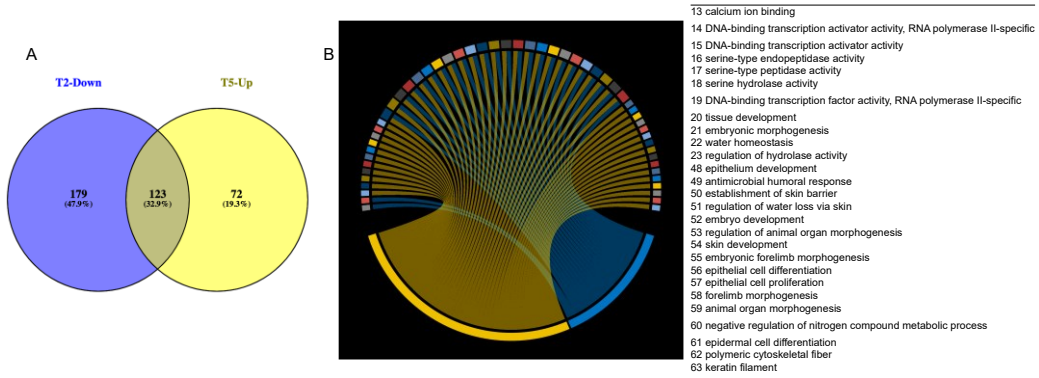


Figure 4.12. The upregulated DEGs with enriched GO functions at WT- T5.

A. The venn plot showing overlapped DEGs between the WT- T5 group (upregulated) and the WT-T2 (downregulated) group.

B. The circus plot showing overlapped GO functions enriched by upregulated DEGs in the WT-T5 group and downregulated DEGs in the WT-T2 group. No. 13 to 23 and 48 to 63 represent GO functions that were specific to the WT- T5 group enriched by upregulated DEGs. No. 11 to 12 represent GO functions that were specific to the WT- T2 group enriched by downregulated DEGs. No. 24 to 47 represent overlapped GO functions between the WT- T2 group and the WT-T5 group.

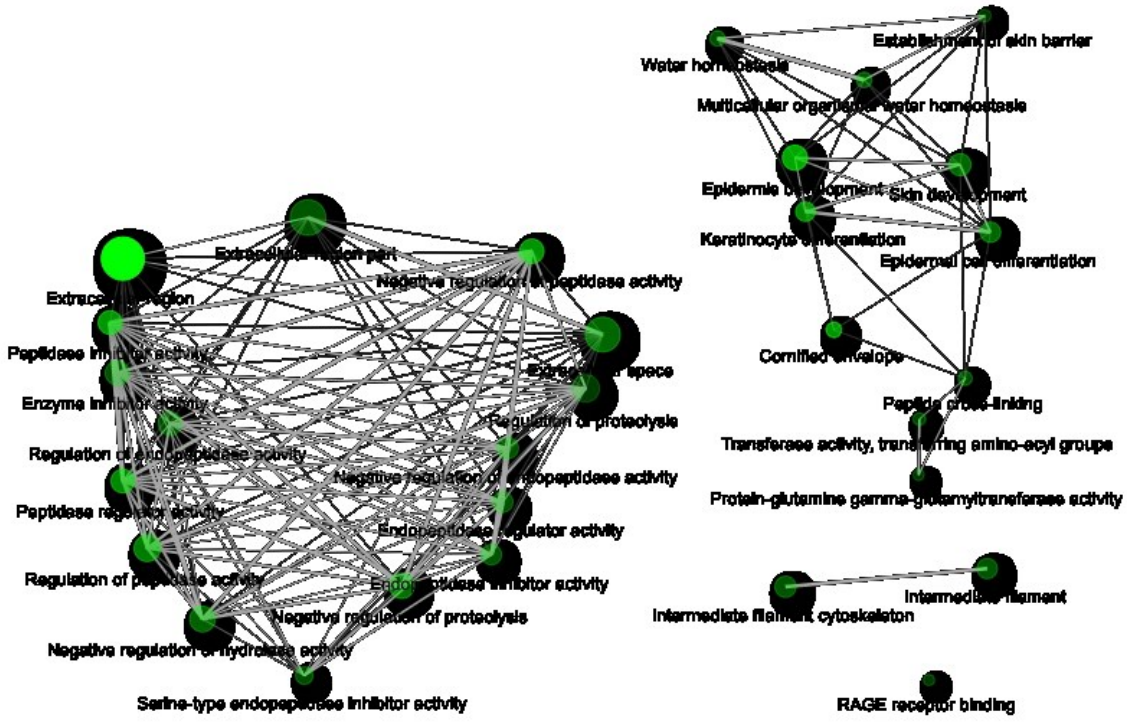


Figure 4.13. A proposed network summarizing GO terms in the WT-T2 group.

Two pathways (nodes) are connected if they share 20% or more genes. Darker nodes are more significantly enriched gene sets. Bigger nodes represent larger gene sets. Thicker edges represent a higher proportion of transcripts were shared among each term.

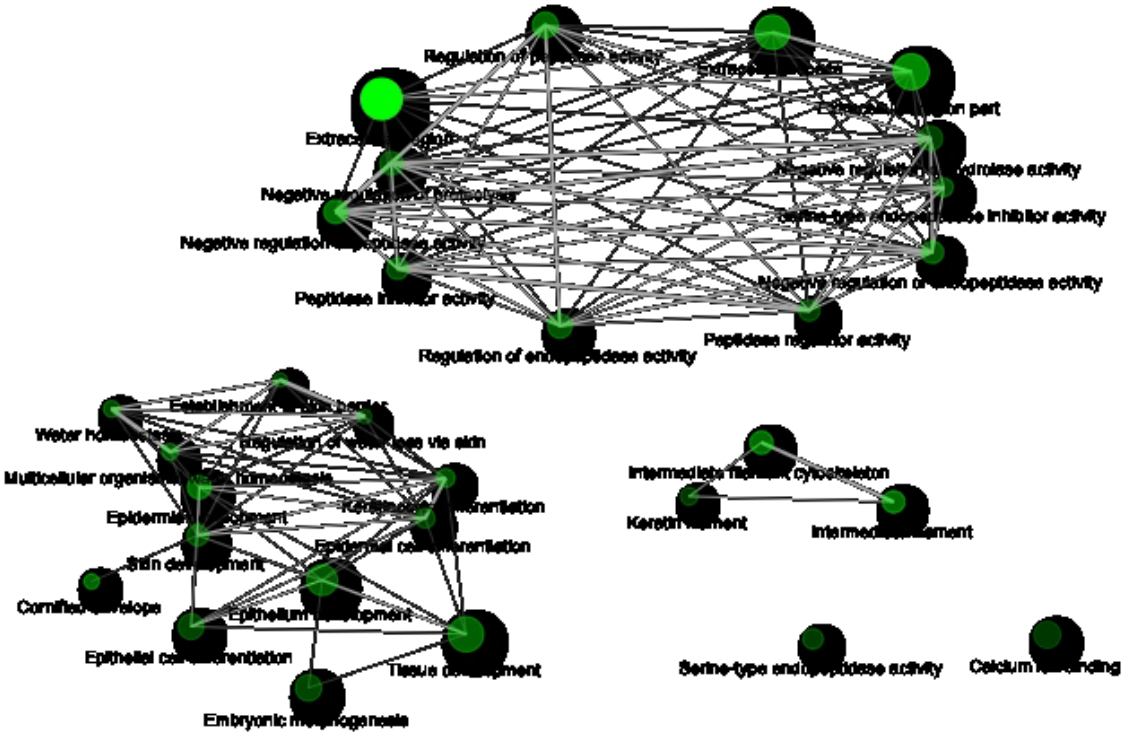


Figure 4.14. A proposed network summarizing GO terms in the WT-T5 group.

Two pathways (nodes) are connected if they share 20% or more genes. Darker nodes are more significantly enriched gene sets. Bigger nodes represent larger gene sets. Thicker edges represent a higher proportion of transcripts were shared among each term.

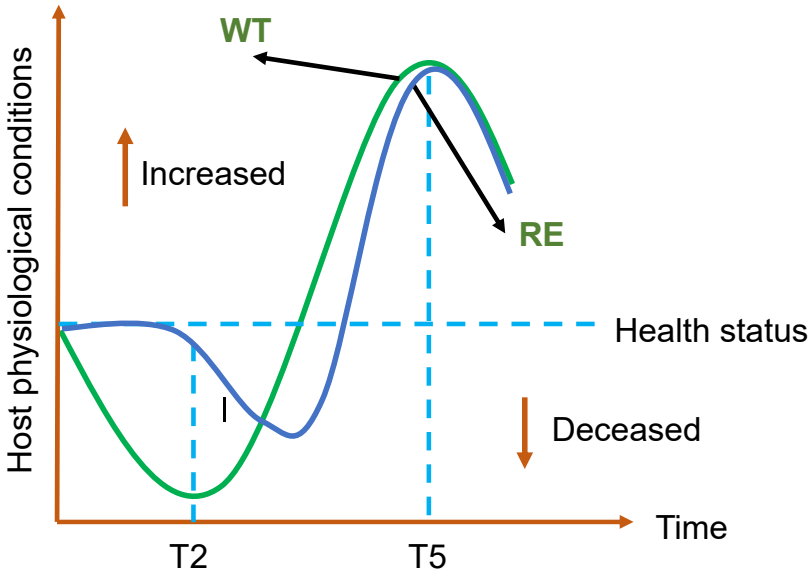


Figure 4.15. Proposed dynamic pattern of host responses to strain-specific STEC O157 colonization.

Chapter 5. Strain-specific Shiga toxins producing *Escherichia coli* O157 colonization affected microbial interactions and assembly of the active rectal mucosa-attached microbiome and its interactions with host immune function in calves[3]

5.1 Introduction

Shiga toxins producing *Escherichia coli* (STEC) is a critical foodborne pathogen for humans and accounts for more than two million acute illnesses annually worldwide (Majowicz et al., 2014). Among STEC, STEC O157:H7 is the major serotype responsible for severe sequala in humans including hemolytic uremic syndrome (HUS) and hemorrhagic colitis (HC) (Kim and Song, 2022). Shiga toxins are the main virulence factors in STEC O157 and can be broadly classified into Shiga toxin 1 (stx1) and Shiga toxin 2 (stx2) (Melton-Celsa, 2014). Shiga toxin 2 is more often associated with human illness as it is present in 40%, 41%, and 43% of isolates responsible for HUS, hospitalization, and HC cases, respectively (Panel et al., 2020). Shiga toxin 2 occurs as several variants including stx2 a, b, c, d, e, f, and g (Melton-Celsa, 2014). Although the acquisition of stx2a prophage has been reported to be integral to STEC O157 pathogenicity, a recent study revealed that a phage type (PT) 21/28 associated with STEC diseases in humans contained both stx2a and stx2c encoding prophages but only expressed stx2c(Chase-Topping et al., 2008). For instance, 61% of HUS cases in children were caused by PT 21/28$^{stx2a\text{-}stx2c+}$ in Scotland from 1997 to 2001 (Chase-Topping et al., 2008).

[3] Chapter 5 is a part of a manuscript submitted to Microbiome, and currently under review.

Cattle are the main reservoir for STEC O157 with rectal-anal junction (RAJ) being the major colonization site (Xu et al., 2014; Wang et al., 2017). Cattle shed STEC O157 through feces into the environment with those that shed more than 10^4 CFU STEC per gram of feces being defined as 'super shedders' (SS). Super shedders are considered to be the primary source of STEC transmission on farms and to the food production chain (Chase-Topping et al., 2008; Xu et al., 2014; Munns et al., 2015; Wang et al., 2016). A previous study also found that PT 21/28$^{stx2a\text{-}stx2c+}$ was likely to be associated with SS, accounting for 50% of the cattle isolates in Scotland (Chase-Topping et al., 2008). These findings suggest that this phage type could play a key role in cattle-human transmission. Commensal bacteria have been reported to inhibit STEC O157 colonization in the ruminant digestive tract through direct (*i.e.* competitive exclusion) and indirect (*i.e.* activation of host immune protection) mechanisms (Buffie and Pamer, 2013). Recent studies have highlighted that RAJ mucosal-attached microbiota were functionally and compositionally influenced by STEC O157 (Wang et al., 2016, 2018). Furthermore, microbial interactions differed among rectal mucosal microbial communities in beef cattle that harbored STEC O157 with varied *stx2* expression (Pan et al., 2021b). However, most previous STEC O157 studies in beef cattle used fecal and/or rectal lumen samples as opposed to epithelial RAJ mucosa.

It is known that STEC O157 colonizes the RAJ tissue of young calves, where it has been reported to cause lesions and gut immune dysfunction (Dean-Nystrom et al., 1997; Sandhu and Gyles, 2002; Menge, 2020). We speculate that varied host-microbiome interactions during early life could be one of the factors that influence the development of SS and the persistence of

STEC in cattle. The gut microbial community is gradually established after birth, a period critical for shaping the composition and structure of the gut microbiome that influences host health (Zhou et al., 2013; Malmuthuge and Guan, 2017; Malmuthuge et al., 2019; O'Hara et al., 2020). Both deterministic factors (*i.e.* interspecies interactions, species traits, host) and stochastic factors (*i.e.* birth, death, colonization of microbes) can simultaneously occur and affect the assembly of microbial communities (Zhou and Ning, 2017). For instance, a recent study of the rumen of adult dairy cows revealed that stochasticity played a role in shaping long-term rumen microbiome development and stability in cattle (Furman et al., 2020). Within gut microbial communities, microbes interact within ecological niches that are crucial for the successful establishment and maintenance of microbial populations (Braga et al., 2016). We further speculate that the interaction within the early life microbiome during the pathogen infection can be affected by both deterministic and stochastic factors.

Additionally, STEC O157 colonization has been shown to alter intestinal immunity in beef cattle. For instance, immune genes including chemokine (C-C motif) ligand 21 (*CCL21*), CD19 molecule (*CD19*), and 4-domains, subfamily A, and member 1 (*MS4A1*) were reported to be downregulated in SS compared to NS (Wang et al., 2017; Pan et al., 2021a). Although differentially expressed genes in rectal tissues were found to be associated with predicted hindgut mucosa microbial functions in SS, *i.e.* the expression of Ficolin 2 (*FCN2*) was negatively correlated with the relative abundance of microbial metabolic pathways, including amino acid-related enzymes (Wang et al., 2016). However, the mechanisms whereby STEC O157 influences the composition, interactions, and assembly of hindgut microbiota, as well as

host immune responses are unknown. Also, current microbial studies focused on DNA-based amplicon sequencing, which fails to represent how active microbes respond to gut environment alterations. Therefore, in this study, we aimed to assess the microbial compositional and structural dynamics, microbial interactions, assembly shifts, and stability of the active mucosa-attached microbiome using tissue samples collected from calves orally challenged with STEC O157 strains that lacked (PT 21/28$^{stx2a\text{-}stx2c+}$) or possessed stx2a (RE 21/28$^{stx2a+stx2c+}$) using RNA based sequencing. Additionally, we assessed the shift of host immune responses and host-microbial interactions to strain-specific O157 colonization.

5.2 Materials and methods

5.2.1 Animal study and sample collection

The animal study, sample collection, and fecal STEC O157 enumeration are described in Chapter 4.

5.2.2 RNA extraction and sequencing

RNA extraction and quantification are described in Chapter 4. Total RNA (0.1 µg) was further subjected to reverse transcription to synthesize cDNA using a cDNA Synthesis Kit (Bio-Rad, Hercules, CA, USA). Single-stranded cDNA was amplified using Oligo(dT)12-18 (Life Technologies, Carlsbad, CA, USA) and SuperScript™ II RT (Life Technologies, Carlsbad, CA, USA) was used to synthesize double-strand cDNA.

To generate the mucosa-attached active bacterial compositional profiles, the bacterial V1-V3 region of the 16S rRNA gene was amplified from generated cDNA using bacterial primers

Ba9F (5'-GAGTTTGATCMTGGCTCAG-3') and Ba515Rmod1 (5'-CCGCGGCKGCTGGCAC-3'). The PCR amplification products were verified using agarose gel (2%) electrophoresis and purified with a Qiagen Gel Extraction Kit (Qiagen, Germany). All amplicon libraries were sequenced using an Illumina MiSeq PE 300 platform (2 × 300 pair-end) at Génome Québec, McGill University (Quebec, Canada).

Extracted total RNA (1ug) was used for library construction using the Truseq Stranded Total RNA Sample Preparation kit (Illumina, San Diego, CA, USA) following the manufacturer's instructions. The quality of constructed libraries was assessed using an Agilent 2200 TapeStation and a Qubit 2.0 Fluorometer. RNA sequencing was performed using a HiSeq 4000 sequencing system (Illumina, San Diego, CA, USA), with paired-end (100 bp) sequencing at Genome Quebec Innovation Centre. One sample from WT-T1, two samples from CT-T2, two samples from CT-T2, one sample from WT-T5, and one sample from RE-T5 were omitted due to the low quality of RNA-sequencing reads.

5.2.3 Mucosal attached microbial community analysis

The raw sequence data were assigned to each sample according to the corresponding barcode and were processed using QIIME2 (Version 2019.10) (Bolyen et al., 2019). Quality control, denoising, removal of chimeric sequences, and generation of amplicon sequencing variants (ASVs) were performed using the QIIME2 plugin DADA2 (Callahan et al., 2016). Taxonomic classification was performed in QIIME2 using a taxonomic classifier with the SILVA database (version 132) as the reference. The Good's coverage index was used to evaluate the adequacy of sequencing depth to generate bacterial profiles in each sample.

Alpha diversity was estimated using Shannon (evenness) and Chao1 (richness) indices. The Kolmogorov-Smirnov (KS) test was used to assess if the shifts of alpha diversities from pre- to post-challenge differed ($P \leq 0.01$ as a significance) or shared a similar trend of changing patterns ($P > 0.01$). This nonparametric test assessed whether empirical distributions of variables between two different groups were significantly different (Karson, 1968). Beta diversity was evaluated based on Bray-Curtis distance to determine the similarities of active microbial profiles across times among the three groups with ANOSIM (analysis of similarities) determining the effects of age and challenge on microbial similarities. Linear regression models were constructed to assess the relationship between alpha diversities and STEC O157 fecal shedding levels at T2 and T5 with Shannon/Chao1 indices as independent variables and log10 fecal shedding levels as dependent variables ($P < 0.05$ as significant). The PROC MIXED model in SAS (ver. 9.13; SAS Institute Inc., Cary, NC, United States) was used to analyze the effects of challenge and age on the relative abundance of taxa at the phylum level.

5.2.4 Microbial interactions within the active mucosa microbial community using network analysis

Microbial networks were constructed using Spearman's coefficient (absolute Spearman's $R > 0.6$, $P < 0.05$) and topological properties, including modularity (value > 0.4, suggesting that the network has a modular structure) (Newman, 2006), average degree (the average number of connections per node) (Wolfe, 1997) and clustering coefficient (also termed transitivity, the degree to which nodes tend to cluster together) (Ravasz et al., 2002) were computed for each network. Then, two topological properties (average degree, clustering coefficient) representing

node distributions were compared pairwise across time in microbial communities associated with CT, WT, and RE groups using the Kolmogorov-Smirnov (KS) test in R (Karson, 1968).

Within-module connectivity (Zi) and among-module connectivity (Pi) were computed to characterize the topological role of nodes with the classification as follows: network hubs (Zi>2.5; Pi>0.62), module hubs (Zi >2.5; Pi<0.62), connectors (Zi<2.5; Pi>0.62) and peripherals (Zi<2.5; Pi<0.62) (Guimerà and Amaral, 2005). Network generalists refer to taxa that are highly connected with others both within and among modules (network hubs), within a module (module hubs), and among different modules within a network (connectors). Network specialists represent peripheral taxa that interact less with other taxa (*i.e.*, nodes were considered connected within a module if at least 60% of the links were within the module) (Guimerà and Amaral, 2005).

Natural connectivity measures the network stability based on the following algorithm:

$$\mathrm{ave}(\lambda) = \ln(\frac{1}{N}\sum_{i=1}^{N} e^{\lambda_i})$$

where $\mathrm{ave}(\lambda)$ is the natural connectivity, N is the number of nodes in the network, and λ_i is the eigenvalue of the adjacency matrix. Up to 80% of total nodes in each group were randomly removed from the adjacency matrix and λ_i and $\mathrm{ave}(\lambda)$ were re-calculated after each removal. The visualization of the natural connectivity was performed using the ggplot2 package in R.

To assess if the microbial taxa abundance affected microbial interactions, the microbial taxa were stratified into three groups based on their relative abundance: high abundance (relative abundance>1%), moderate abundance (0.01%< relative abundance<1%), and low abundance (relative abundance<0.01%) (Zhang et al., 2018; Tian et al., 2023). Then, the

relationships between abundant-specific microbial generalists/specialists and network properties (*i.e.* nodes, edges, modularity, average degree, clustering coefficient, natural connectivity) were determined using Spearman's coefficient (absolute R>0.8 and P<0.05 as a significance) across each group. Significant relationships were further examined using linear regression models (R^2 >0.8, P<0.05 as a significance).

5.2.5 Assessment of microbial community assembly patterns in response to STEC O157 challenge

Determinism highlights strong selections imposed by environments and species interactions, while stochasticity focuses on the random and unpredictable events affecting assembly (Stegen et al., 2012). The Raup-Crick distance (β_{RC}) was used to assess the relative importance of stochastic/deterministic processes in the microbial assemblage. The β_{RC} measures the extent to which the deterministic-driven assembly deviates from the assemblies based on null (stochastic) expectations: a value approaching -1 or 1 (β_{RC}>0.95 or β_{RC}<-0.95) refers to the deterministic factors that drive microbial community assembly (Anderson et al., 2011; Chase et al., 2011). Whereas if β_{RC} does not significantly deviate from 0 (-0.95<β_{RC}<0.95), the community is considered a stochastic-driven assembly.

5.2.6 Microbial specialization patterns in response to STEC O157 colonization and fecal shedding

Microbial specialization is a general biological process that shapes the assembly of microbial communities (Johnson et al., 2012), which can be measured using the niche breadth index:

$$B_j = \frac{1}{\sum_{i=1}^{N} P_{ij}^2}$$

Here, B_j stands for niche breadth, P_{ij} refers to the proportion of any species i in a given sample, j and N is the total number of samples (Levins, n.d.). Niche breadth refers to the diversity of recourses used by an individual (or species) within a certain environment (Carscadden et al., 2020). When available recourses are limited, niche breadth is usually increased for species to gain more resources for survival and such taxa can occupy a broader niche and are defined as generalists. In contrast, species that selectively use specific resources and have a narrower niche breadth are defined as specialists (Carscadden et al., 2020). Based on the 1000-time simulations (quasiswap permutation algorithms, EcolUtils R package), the empirical niche breadth value of certain taxa that exceeded the 95% confidence interval of the null distribution was designated as a generalist, whereas those that were below the 95% confidence interval were defined as a specialist. Taxa that were within the 95% confidence interval were defined as neutralists (Wu et al., 2017). The Chi-square test was used to test the equality of numbers of specialized microbes across CT, WT, and RE groups pre- and post-challenge ($P<0.01$ as significant). Specialized microbes among microbial assemblage that play a role in microbial interactions were identified (that is a taxon involved in both microbial networks and microbial assemblage). The relations between the relative abundance of identified dual role microbes and both β_{RC} and $\log_{10}$ STEC fecal shedding were assessed using linear regression models ($P<0.05$ as a significance).

5.2.7 Identification of host immune-related pathways and host-microbial interactions

RNA-sequencing reads were first subjected to the quality filter and adapter trimming using FastQC and bbDuk. Filtered reads were then mapped against the reference bovine reference

assembly ARS-UCD 1.2 using STAR (Version 2.7.1a). Feature counts of each ensemble ID were then generated using subread (Version 2.0.0) and were then normalized into TPM (transcripts per million). The gene set enrichment analysis (GSEA), which considers all genes instead of only differentially abundant genes was adopted to identify the altered pathways (Mootha et al., 2003; Subramanian et al., 2005). This approach does not need the preselection of genes of interest and is capable of identifying whether a pathway is up-/down- regulated (Subramanian et al., 2005; Reimand et al., 2019). Particularly, the GSEA analysis was used to identify pathways that were altered in response to the STEC O157 challenge in both WT and RE groups using R (altered pathways: absolute normalized enrichment score >1, nominal P value <0.05, FDR q-value <0.1). Among altered pathways, host immune pathways were then selected, and only expressed genes among these selected pathways were considered for further analysis. Host genes from identified host immune-related pathways were correlated with the relative abundance of identified key microbes using the Spearman correlation (Absolute R>0.8 and P<0.01 as a significance) among CT, WT, and RE groups at pre- and post-challenge.

5.2.8 Data availability

All sequence data have been deposited to NCBI Sequence Read Archive (SRA) under accession numbers PRJNA991158 (RNA sequencing) and PRJNA988112 (Amplicon sequencing).

5.3 Results

5.3.1 Strain-specific STEC O157 challenge resulted in a similar active rectal mucosa microbial community structure

One sample from the RE-T5 group was excluded due to low-quality sequence data. An average of 37,449 ±10,711 paired-end raw reads were generated for 71 mucosa samples. After quality control, a total of 11,375 amplicon sequence variants (ASVs) were identified with an average of 160 ± 5 ASVs per sample from an average of 15,774 ± 3,747 filtered pair-end reads. The Good's coverage was > 99.9 % for all samples. A total of 13 phyla were identified across all groups, among which *Actinobacteria*, *Bacteroidota*, *Firmicutes*, and *Proteobacteria* were the most abundant (accumulated relative abundance accounting >90% for each group, Table 5.1). Among identified phyla, the relative abundance of *Bacteroidota*, and *Actinobacteria* were affected by interactions between challenge and age ($P_{Bacteroidota} = 0.02$, $P_{Actinobacteria} = 0.04$, Table 5.1).

Both Shannon and Chao1 indices (alpha diversity) of rectal mucosal microbiota in WT and RE calves exhibited similar change patterns: increasing from pre-challenge (T1) to peak shedding periods (T2), and then decreasing from T2 to T5 (All KS test P>0.01, P>0.01 stands for a similar trend of changing patterns, Figure 5.1A, B). However, the shift in alpha diversity of the CT group differed from WT and RE groups as both indices were the lowest at T2 (Figure 5.1A, B). Furthermore, all three groups differed ($P_{Shannon}<0.01$, $P_{Chao1} =0.09$) in Shannon and Chao1 indices at T2. Similarly, the Bray-Curtis distance (beta diversity) of microbial similarity showed a separation of rectal bacteria between T1 and T5, and the cluster of rectal bacteria at T2 overlapped with T1 and T5 in both WT and RE groups, while this pattern was not apparent in the CT group (Figure 5.1C). The ANOSIM analysis identified several interactions between age and STEC O157 challenge regarding the similarity of microbial communities

($P_{Age*Challenge}$=0.01, P_{Age}=0.21, $P_{Challenge}$=0.36). Although microbial richness and evenness were altered similarly in both WT and RE groups, the relationship between alpha diversities and log_{10} fecal shedding levels varied (Figure 5.2). In particular, the Shannon index was negatively correlated to log_{10} fecal shedding at T2 in the WT group (P=0.041, Figure 5.2), while no correlation was observed with the RE group.

5.3.2 Differential microbial interactions in active mucosal attached microbiota in response to strain-specific O157 challenges

Varied dynamic patterns of microbial networks and their properties were identified among three groups. The pattern of network stability in calves challenged with different O157 strains was similar with the greatest microbial network stability occurring during peak shedding (T2) and the lowest at T5 (Figure 5.3A-C). It is noted that PT 21/28$^{stx2a\text{-}stx2c+}$ colonization in calves led to greater variability in microbial network stability as compared to those inoculated with RE 21/28$^{stx2+stx2c+}$ (Figure 5.3B, C). In contrast, a continuous increase in microbial network stability was identified in the CT group from T1 to T5 (Figure 5.3A).

The number of nodes (microbial taxa) exhibited similar change patterns in all three groups during the experimental period, in that the number of nodes increased from T1 to T2 and decreased from T2 to T5 (Figure 5.3D, Table 5.2). However, the number of edges peaked at T2 for the two challenge groups (WT and RE), while it increased from T2 to T5 in the CT group (Figure 5.3E). The modularity [(the index measures the strength of division of a network into modules with its value >0.4 suggesting the network has a modular structure (Newman, 2006)]) was greater than 0.4 for CT, WT, and RE groups (Figure 5.3F). The average degree (an index

that measures the number of edges connected to a node) was the highest at T2 for WT and RE, while it was the highest at T5 in the CT group (Figure 5.3G). For clustering coefficients, the highest value was observed at T2 for all three groups (Figure 5.3H), while the Kolmogorov-Smirnov tests showed differences ($P<0.01$) in the average degrees and clustering coefficients among the three groups during the experimental period.

From the network modularity perspective, taxa were classified into network hubs, module hubs, connectors that represent generalists, and peripherals that represent specialists in the community. In our study, network hubs and module hubs were not identified in CT, WT, or RE groups. An average of 2% (ranging from 0 to 3.8%) of total nodes were designated as connectors while an average of 98 % ± 0.4 % of total nodes were classified as peripherals for CT, WT, and RE groups (Table 5.3).

5.3.3 Moderate abundant genera featured as network specialists relate to network stability

According to the aforementioned stratification criteria, 78%-90% of total genera (from 98 to 164) belonged to the moderate abundant taxa among CT, WT, and RE groups, while 7%-18% and 0-1% of total genera (from 98 to 164) were designated as high- and low- abundant taxa, respectively (Table 5.4). Out of the 23 classified connectors, 17 (74%) were members of the moderate-abundant taxa among three groups (Figure 5.5A), and none of the network connectors was related to network stability (Figure 5.5). Among the 1,073 classified peripherals, 912 (85%) belonged to the moderate-abundant taxa among the three groups (Figure 5.4B). The number of peripherals ($R^2_{adj}=0.82$, $P<0.01$), particularly moderately ($R^2_{adj}=0.86$, $P<0.01$) abundant

peripherals were positively correlated to network stability, whereas low- abundant peripherals tended to be significant (R^2_{adj}=0.47, P=0.025, Figure 5.4C-E).

5.3.4 STEC O157 challenge affected the dynamics of mucosa microbial assembly patterns

The Raup-Crick distance revealed that microbial community assembly patterns in the CT group were consistently stochastic-driven (Figure 5.6A), while it transitioned from deterministic-driven (T1 and T2) to a stochastic-driven (T5) assembly in the WT group (Figure 5.6A). For the RE group, the assembly of the microbial community was stochastic-driven at T2, while a deterministic process made the major contribution to assemblies at T1 and T5 (Figure 5.6A). In addition, the Raup-Crick distance was negatively correlated with $\log_{10}$ STEC O157 fecal shedding in the WT group, with the β_{RC} being increased for more stochastic-driven mucosal attached microbiota of calves with lower O157 fecal shedding (R^2_{adj}=0.67, P<0.01, Figure 5.6B). However, the Raup-Crick distance measured for mucosal-attached microbiota of the RE group was not related to $\log_{10}$ STEC O157 fecal shedding (Figure 5.6C).

5.3.5 STEC O157 challenge affected microbes participating in both microbial interactions and assembly

Microbial specialization quantifies the role of each microbe within the microbial community by estimating the availability of resources/nutrients to individuals (Carscadden et al., 2020). A range of 6 to 29 (6% -18%) assembly generalists and 5 to 19 (5% -14%) assembly specialists were identified for all three groups from T1 to T5 (All P>0.1, Table 5.5, Figure 5.7). Among identified specialized microbes, two bacterial taxa (*Paeniclostridium* and *Gallibacterium*) were the only ones (assembly specialists) designated as network connectors. Particularly, the relative

abundance of both *Paeniclostridium* and *Gallibacterium* were similar among CT, WT, and RE groups pre-challenge (Figure 5.8A, B). However, the relative abundance of *Paeniclostridium* was increased in WT as compared to CT at T2 ($P<0.01$), while CT and RE have a similar relative abundance to *Paeniclostridium* at this timepoint ($P=0.25$, Figure 5.8A). At T5, the relative abundance of *Paeniclostridium* showed a tendency of being higher in both WT and RE as compared to CT ($P_{WT\ vs.\ CT}=0.081$, $P_{RE\ vs.\ CT}=0.011$, Figure 5.8A). At T2, the relative abundance of *Gallibacterium* showed an increasing trend in CT as compared to the WT, while it did not differ between CT and RE (Figure 5.8B). Furthermore, the relative abundance of *Gallibacterium* showed an increasing trend in both WT and RE compared to the CT at T5 ($P_{WT\ vs.\ CT}=0.05$, $P_{RE\ vs.\ CT}=0.011$, Figure 5.8B). The relative abundance of *Paeniclostridium* was positively correlated with β_{RC} ($R^2_{adj}=0.68$, $P=0.028$) at T2 and negatively correlated with β_{RC} at T5 ($R^2_{adj}=0.66$, $P=0.058$) in the WT group (Figure 5.9A top left and top right). For the relative abundance of *Gallibacterium,* only samples at T2 tended to be positively correlated with β_{RC} ($R^2_{adj}=0.36$, $P=0.091$, Figure 5.9B bottom left).

5.3.6 Strain-dependent host immune responses and host-microbial interactions in calves after STEC O157 challenge

The GSEA analysis revealed varied host responses to the STEC O157 challenge between WT and RE, among them altered pathways related to host immunities were selected (defined as host immune-related pathways). The percentage of expressed genes within each pathway (that is the number of genes expressed within each pathway among all samples/number of genes belonging to each pathway) were assessed including antigen processing and presentation (43 out of 88

expressed genes), chemokine signaling pathway (171 out of 189 expressed genes), the intestinal immune network for IgA production (32 out of 48 expressed genes), natural killer cell-mediated cytotoxicity (95 out of 137 expressed genes), MAPK signaling pathway (259 out of 267 expressed genes), T cell receptor signaling pathway (107 out of 108 expressed genes), B cell receptor pathway (73 out of 75 expressed genes). Three host immune-related pathways including the MAPK signaling pathway, antigen processing and presentation, and the T-cell receptor signaling pathway were upregulated at T2 compared to both T1 and T5 in WT, with the B-cell receptor signaling pathway being the only pathway upregulated at T2 as compared to T1 in WT (Figure 5.10A). For RE, both T-cell and B-cell receptor signaling pathways were upregulated at T5 and T2 as compared to T1. Four pathways, including antigen processing and presentation, the chemokine signaling pathway, intestinal immune network for IgA production, and the natural killer cell mediated cytotoxicity were upregulated at T5 compared to T1 (Figure 5.10B).

There were no significant interactions between the relative abundance of active mucosal-attached microbes and expressions of host immune related genes (defined as host-microbial interactions in our study) for CT from T1 to T5, except for *MAPT* from the MAPK signaling pathway being negatively correlated to *UCG.010* at T2 (Table 5.6). For WT, a range of 10 to 49, 0 to 6, and 9 to 59 positive host-microbial interactions were identified at T1, T2, and T5, respectively (Table 5.6). A range of 6 to 35, 3 to 22, and 4 to 82 negative host-microbial interactions were identified in WT at T1, T2, and T5, respectively (Table 5.6). For RE, only positive host-microbial interactions were identified at T1. Two positive (*NCR3* from cell

mediated cytotoxicity interacted with *Odoribacter*; *MAP3K12* from MAPK signaling pathway interacted with *Rikenellaceae RC9 gut group*) and 1 negative (*MAP2K3* from MAPK signaling pathway interacted with *Oscillospiraceae*) host-microbial interactions were identified at T2 in RE (Table 5.6). A range of 1 to 6 positive and 1 to 29 negative host-microbial interactions were identified in RE at T5 (Table 5.6).

5.3.7 Observed varied host-microbial interactions for beneficial and pathogenic microbes in response to the STEC O157 challenge

We further assessed interactions between mucosal active microbes and expressions of host genes involved in host-immune related pathways (Figure 5.11-5.13). The commensal bacteria, *Prevotella* (Larsen, 2017; Iljazovic et al., 2021)*, Rikenellaceae RC9 gut group* (Sun et al., 2019; Huang et al., 2021)*, Fecalibacterium* (Ferreira-Halder et al., 2017), and *Dorea* (Shahi et al., 2017) were detected in the active mucosal attached microbiota with their relative abundance (% ± SD) being 0.5 ± 0.008, 1.6 ± 0.013, 0.8 ± 0.021, 0.1 ± 0.004 in all samples. Their potential interactions with genes involved in host immune-related pathways revealed that in the WT-T1, the relative abundance of *Prevotella* (0.2 ± 0.001) exhibited the greatest interaction with host immune genes, being negatively correlated with *CALR*, *HSP90AA1*, and *PDIA3* involved in antigen processing and presentation as well as *PLA2G12A* and *CACNA2D2* involved in the MAPK signaling pathway (All $P<0.01$, Figure 5.14A). The relative abundance of *Dorea* (0.2 ± 0.002) was negatively correlated with *BLNK* from B-cell receptor signaling pathway and *ADCY2* involved in Chemokine signaling pathway, while it was positively correlated with *GADD45A* involved in MAPK signaling pathway and *IFNGR1* from natural killer cell mediated

cytotoxicity in WT-T1 (All P<0.01, Figure 5.14A). Before the challenge of WT, the relative abundance of *Escherichia- Shigella* (1.1 ± 0.01) was negatively correlated with the expression of genes involved in the MAPK signaling pathway, chemokine signaling pathway, and antigen processing and presentation pathway (Figure 5.14A). At peak fecal shedding, in WT, only the relative abundance of *Paeniclostridium* (0.7 ± 0.008) was negatively correlated with the expression of genes involved in antigen processing and presentation, intestinal immune network for IgA production, T cell receptor signaling pathway, MAPK signaling pathway, B cell receptor signaling pathway and natural killer cell mediated cytotoxicity (Figure 5.14B). At T5 in WT, only the relative abundance of *Rikenellaceae RC9 gut group* (0.7 ± 0.003) and *Escherichia- Shigella* (7.8 ± 0.17) showed the most interactions with host immune-related gene expressions (Figure 5.14C). For RE, only the relative abundance of *Faecelibacterium* (4.7 ± 0.04) was positively correlated with expression of genes involved in antigen processing and presentation, MAPK signaling pathway, and natural killer cell mediated cytotoxicity at T1, and only relative abundance of *Rikenellaceae RC9 gut group* (1.7 ± 0.009) was positively correlated with *MAP3K12* expression in the MAPK signaling pathway at T2 (Figure 5.15A, B). The relative abundance of *Paeniclostridium* (1.0 ± 0.01) showed varied interactions with expression of genes involved in antigen processing and presentation, intestinal immune network for IgA production, T cell receptor signaling pathway, MAPK signaling pathway, B cell receptor signaling pathway, natural killer cell mediated cytotoxicity and the chemokine signaling pathway in RE-T5 (Figure 5.15C).

5.4 Discussion

Previous studies suggested that *E. coli* O157 supper-shedders had significantly different microbial profiles associated with the rectal mucosa and feces in beef cattle (Pan et al., 2021b). However, studies that have examined the active mucosal microbiome and its relationship with host gene expression are limited (RNA level). Different from previous studies using DNA-based amplicon sequencing (Wang et al., 2017), the current study is the first to reveal the active mucosal attached microbiota associated with the rectal anal junction and its shift in response to STEC challenge. Firstly, we found that the mucosal attached microbial community in the rectal tissue of veal calves was dominated by Firmicutes, Bacteroidetes, Proteobacteria, and Actinobacteria, phyla which are similar to that found in adult cattle using DNA-based amplicon sequencing (Xu et al., 2014; Mao et al., 2015; Wang et al., 2017). Compared to DNA-based amplicon sequencing which suggested *Treponema* was the most abundant classified genus (9.13%) associated with the rectal mucosa of beef cattle (Mao et al., 2015), we found that *Treponema* was only a highly abundant taxon (1.6%) with *UCG.005* (30%), *UCG.010* (15%), and *Christensenellaceae R7 group* (9%) being the most abundant genus for all three treatment groups. These results suggest that microbial compositions are different at DNA and RNA levels and further microbiome studies using RNA-based amplicon sequencing should be encouraged to uncover 'true' microbial compositions.

Secondly, we found that the profiles and diversity of this community can be affected by STEC O157 challenge. Similar to other super shedder studies (Mir et al., 2019; Vasco et al., 2021), we found that higher fecal shedding levels increased alpha diversity. This likely reflects

the ability of STEC O157 to be a successful competitor within gut lumen bacterial communities (Kaper et al., 2004), which could also be the case for the mucosal-attached community. At peak shedding times (T2), STEC O157 may deprive commensal bacteria of nutrients, enabling it to reach peak population levels. As a result, rectal mucosal attached microbial communities can become more diverse to regain the space and resources to maintain their homeostasis to repel pathogens, as reflected by the increase in alpha diversity at peak shedding in both groups challenged with STEC O157. Once STEC O157 colonization is limited, the microbiota can shift back to the relatively simple community as reflected by the reduced alpha diversity at T5. Similarly, this same pattern was observed in the beta diversity of bacterial communities in calves challenged with STEC O157 as compared to calves that were not. These suggest that active mucosal microbial communities could respond to rectal mucosal STEC O157 colonization, highlighting the importance of examining active mucosal microbiota for O157 colonization in beef cattle.

Although stx2a was more toxic than stx2c which causes a higher degree of host responses (Fraser et al., 2004; Melton-Celsa, 2014), the expression of stx2a genes in STEC did not affect RAJ active mucosal microbial profiles. This finding complied with our previous research, supports the concept that microbial diversities and similarities are not driven by the expression of stx2 in beef cattle (Pan et al., 2021b). The similar changing patterns of microbial profiles may be due to the nature of Shiga toxins. The Shiga toxins attach to epithelium through the glycosphingolipid receptor (Gb3 receptor), which was not expressed on the surface of epithelial cells in beef cattle (Menge et al., 2001; Raa et al., 2009; Melton-Celsa, 2014). Instead, STEC

O157 colonizes the gut epithelium through attachments and adhesions by proteins such as the type III secretion system (T3SS) and intimin in beef cattle (Wu et al., 2010; Sheng et al., 2011; Ji and Dong, 2015; Gaytán et al., 2016; Deng et al., 2017; Xue et al., 2017). Future studies should focus on STEC possessing attachments and adhesions genes and their effects on the structures of the mucosal attached microbiome.

Additionally, we found that colonization of STEC O157 induced dynamic shifts of microbial interactions, and as results crosstalk between mucosal microbes may impact the nature of the metabolites generated. For instance, *Prevotellaceae UCG.003*, a member of the Bacteroidota, was more abundant in the PT 21/28 $^{stx2a\text{-}stx2c+}$ group (0.003 ± 0.003) compared to RE 21/28 $^{stx2a+stx2c+}$ group (0.001 ± 0.001) at T2. Members of the Bacteroidota are known to produce butyrate which can inhibit the growth and colonization of *E.coli* and is critical among microbial interactions (Pryde et al., 2002; Faust et al., 2012; Wang et al., 2019; Silva et al., 2020; Clark et al., 2021; Iljazovic et al., 2021). Hence the higher abundance of *Prevotellaceae* in WT at peak shedding level may infer the higher level of butyrate produced in the gut against STEC O157 colonization and potential alterations of microbial interactions. Whereas butyrate production and STEC O157 colonization may not be the case for RE. The production of stx2a in STEC could affect host gut homeostasis, therefore, augmenting its survival and inhibiting functions of gut commensals, leading to a low abundance of *Prevotellaceae.* Network analysis often highlighted the importance of generalists in driving and stabilizing microbial interactions. For instance, a previous study used ocean bacterial data set to construct microbial networks and revealed that highly connected keystone taxa (that is generalists) were responsible for microbial

compositional changes and their presence promotes microbial stabilities (Herren and McMahon, 2018). However, our study found that peripherals (network specialists), the often-neglected ecotype in microbial interactions, contributed more to the maintenance of network stability. This result suggests that the role of different ecotypes within microbial interactions is dynamic, which may be due to the combined effect of the host and the gut environment. Microbial network stability represents the ability of the microbiota to respond to environmental change (Liu et al., 2022), with the shift to different ecotypes reflecting a response in network stability as a result of pathogen challenge altering the ambient gut environment. Particularly, moderately abundant peripherals made the greatest contribution to network stability with 74% being network generalists (17 out of 23) and 85% (912 out of 1071) belonging to moderate abundant taxa. While other studies focus on rare and high abundant taxa and their effects on microbial communities as was the case for rare taxa serving as keystone taxa for the restoration of multifunctionality of the soil microbiome (Xu et al., 2021), or highly abundant taxa dominating the succession of microbial communities during microbiome establishment under a well-controlled environment (Zhu et al., 2023). However, little is known about the ecological role of moderately abundant taxa within microbial communities. From our research, it is suggested that moderate abundant taxa have a multifunctional role within microbial communities from an ecological perspective, as they serve as keystone taxa within clusters in microbial networks or as peripherals contributing to network stability.

The microbial community assembly is vital since this process affects microbial profiles and their ability to respond to external environmental changes (Liu et al., 2019). However, there is

a lack of such research for the bovine mucosal microbial community. We found that STEC O157 colonization together with stx2 subtype differences affect patterns and the dynamics of rectal mucosal microbial community assembly. Particularly, the production of stx2a but not stx2c shifted assembly patterns from deterministic-driven to stochastic-driven at T2. During the community assembly process, microbes are affected by both stochastic and deterministic factors and can interact with each other leading to the specialization of certain microbial taxa within the niche (Pandit et al., 2009; Braga et al., 2016; Browne et al., 2021). As a result, the production of stx2a in RE may minimize the effect of microbial interactions and responses to gut environmental changes. The fact that the microbial community assembly was driven by the stochastic process in unchallenged calves during the experimental period, suggests that as previously reported, age was a deterministic factor that plays an important role in the assembly of (Furman et al., 2020) the hindgut mucosa microbiome. We found that assembly patterns differed in challenge groups compared to the control group at T1 when all calves were weaned and fed in the same environment. This illustrates that individualized responses play a critical role in gut microbiome assemblage. These results can also facilitate the development of strategies to promote communities that limit the colonization and persistence of STEC O157 at the RAJ.

In our study, the STEC O157 challenge was identified as the deterministic factor affecting microbial assembly, with two assembly specialists being identified that were theoretically related to STEC O157 colonization and rectal mucosal microbial interactions. Two bacterial taxa *Paeclostridium* and *Gallibacterium* were network connectors dominating microbial

interactions and simultaneously being assembly specialists for both WT and RE groups. Assembly specialists harbor narrow ecological niches and are more affected by deterministic processes due to their preferences and sensitivities to external environmental conditions (Logares et al., 2013; Liao et al., 2016; Xun et al., 2019; Xu et al., 2022). Our results confirmed that the relative abundance of pathogenic *Paeclostridium* and opportunistic pathogenic *Gallibacterium* increased after STEC O157 colonization. And such variations can be STEC strain dependent as reflected by their relative abundance increased at 7 days post challenge in the WT group but increased at 26 days post challenge in RE calves. Correspondingly, the expression of *SEMA6A*, the gene encoding for the receptor of exotoxin TcsL in *Paeniclostridium* (Tian et al., 2020), was higher in WT compared to unchallenged calves 7 days post-challenge, while this increase was not observed for RE at T2. Our study is the first to reveal the expression of *SEMA6A* as the receptor of *Paeniclostridium* was affected by the production of stx2a and STEC O157 colonization. Further evidence is that the *Paeclostridium* exhibited varied interactions with community assembly patterns in WT post-challenge, with no effects in the RE pre- and post-challenge. As a bacterial pathogen, *Paeclostridium* can produce hemorrhagic toxins causing acute infectious disease in humans and animals (Vidor et al., 2019; Li et al., 2022). The increased relative abundance of this bacterial taxon after STEC O157 colonization in WT may suggest a potential mutualism between *Paeclostridium* and STEC, and such mutualism could augment the survival and proliferation of these two bacteria. However, the expression of stx2a did not affect interactions between the relative abundance of *Paeclostridium* and microbial assembly in RE at T2, suggesting that stx2a expression may be a

negative signal that inhibits interactions for other microbes, causing no interactions observed for the RE challenge at T2. *Gallibacterium* is the other network connector that is designated to assembly specialists, which is an opportunistic pathogen in animals (Driessche et al., 2020). Whereas no significant relationships with community assembly were found for *Gallibacterium*, suggesting this microbial taxon could play a trivial role in affecting RAJ mucosal microbial assembly. Taken together, our results highlighted that although opportunistic pathogens are in low abundance in the rectal mucosa, pathogen colonization can still affect its abundance and its roles in rectal mucosal microbial interactions and assembly. We also highlighted the approach used for the selection of microbes is novel that we consider the role of microbes in both interactions and assembly, instead of only one process, to fully identify how microbes serve a function within the community.

Previous studies reported cellular and humoral immune responses in calves challenged with stx2-positive STEC O157 (Hoffman et al., 2006). For instance, the serum IgG specific to STEC O157 intimin was significantly increased following oral challenge with a STEC O157 stx2 positive strain in adult cattle (Bretschneider et al., 2007a, 2007b). However, such a study failed to assess how different stx2 subtypes in STEC alter host responses. Another study revealed that the level of systemic H7-specific IgA antibody was increased in calves challenged with PT 21/28$^{stx2a\text{-}stx2c+}$ compared to RE 21/28$^{stx2\text{-}stx2c+}$ (Fitzgerald et al., 2019). Our study revealed a higher magnitude of host immune responses in calves after STEC O157 challenge from the transcriptomic perspective, but host immune responses were affected by stx2 subtype differences. Particularly, wild-type O157 challenge induced intensive but inhibited host

immune responses at peak fecal shedding, with responses then enhanced when shedding returned to normal. While the production of stx2a did not affect host responses at T2 and similar host responses were induced at T5. Our analysis showed how the host responds to strain-specific STEC O157 colonization from the transcriptomic perspective, which extends our knowledge of host-O157 interactions. However, further research using samples collected from other time points (T3, T4) is needed for deciphering the nature of host responses as a result of challenge with strain-specific STEC O157.

In addition to the pathogen-driven immune responses, gut microbiota also plays a key role in regulating host immune function, which is capable of producing metabolites that influence host homeostasis (Nicholson et al., 2012; Jansma and Aidy, 2021). Among the gut microbiota, commensal organisms such as *Prevotella* (Larsen, 2017; Iljazovic et al., 2021)*, Rikenellaceae RC9 gut group* (Sun et al., 2019; Huang et al., 2021)*, Fecalibacterium* (Ferreira-Halder et al., 2017), and *Dorea* (Shahi et al., 2017) were reported to be beneficial to the host. For instance, *Dorea* has been reported to be one of the most abundant genera in non-shedders (Wang et al., 2017)*,* which is a beneficial butyrate-producing bacteria in the gut that enhances tight junctions of the epithelium (Louis and Flint, 2017). Indeed, we found that interactions between beneficial microbes and host immunities were dominant before the challenge in both WT and RE groups. For example, *Prevotella* a mucosa-attached commensal that is positively associated with high fiber consumption can distinctively modulate host immune responses and gut barrier functions in human epithelial cells (Ilhan et al., 2020). In accordance with our findings that host immune-related pathways showed intense positive interactions with *Prevotella* pre-challenge,

suggesting that beneficial microbes are positively related to host immunities when the host remains in its homeostasis. However, STEC O157 colonization shifted such beneficial microbes- host immune genes interactions for both WT and RE groups. The aforementioned *Paeniclostridium*, the identified microbes involved in both microbial interactions and assembly showed significant negative interactions with host immune genes at T2 in PT 21/28 *stx2a-stx2c+* challenged calves, suggesting that *Paeniclostridium* is a critical microbe involved in both host immune responses and STEC O157 colonization. Hence, we proposed a pathogen-gut commensals-host model in that the pathogen colonization initiated the shift of interactions and assembly of microbial communities, which further affect host responses and differentiate host-microbial interactions. In our study, the colonization of STEC O157 first induced active rectal mucosal microbial interactions and assembly variations (Figure 5.16). Through such processes, keystone opportunistic pathogens were identified to be associated with host immunity alterations and differed host-microbial interactions (Figure 5.16). We also noticed that the expression of stx2a in STEC can be a factor affecting the pathogen-gut commensals-host model in our study. It can delay the occurrence of host- *Paeniclostridium* interactions and limit interactions between gut commensals and host immune genes, highlighting the role of stx2a in shifting host-microbial interactions instead of affecting rectal mucosal microbial profiles. The proposed model can be useful to further our understanding of STEC O157 colonization mechanisms and can apply to other similar studies.

Lastly, before challenge when calves were confirmed STEC O157 negative and shared a similar health condition, the quantity of host-microbial interactions was not similar among three

groups, indicating there are specific differences among hosts and possibly individual responses that could affect such host-microbial interactions. Besides, we found that calf growth (age changes) plays a trivial role in affecting interactions between host immune-related pathways and the relative abundance of microbial genera from T1 to T5 since a limited number of host-microbial interactions was identified in unchallenged calves across ages.

5.5 Conclusions

Our research comprehensively revealed the RAJ mucosa attached active microbial profiles and its interactions and assembly patterns shifts in response to strain-specific STEC O157 colonization in calves. Different from studies using fecal samples and DNA-based amplicon sequencing, we collected RAJ mucosa samples and used cDNA-based amplicon sequencing to reflect active microbial communities. Our results revealed that STEC O157 colonization affected RAJ active microbial profiling, however, the changing patterns of microbial structures were not strain-dependent. Instead, both microbial interactions and assembly patterns were strain-dependent. Furthermore, our study suggested STEC O157 colonization can be the deterministic factor affecting RAJ microbial assembly patterns. This study is the first to identify how opportunistic pathogenic taxa (*Paeniclostridium* and *Gallibacterium*) could be the keystone members to affect microbial assembly and microbial interactions during STEC colonization, revealing potential interactions between pathogen and commensal organisms. Further RNA-seq identified how host immunities and host-microbial interactions varied during colonization with responses being strain specific. Our findings identified the dynamic

interactions between active mucosal-attached microbes and host immune genes, suggesting that beneficial microbes dominate interactions with host immunities under homeostasis, while pathogenic microbes bear such responsibility with host immunities once homeostasis is disrupted. Therefore, we speculate that STEC O157 indirectly relates to the host through RAJ mucosal microbial interactions and assembly, particularly by regulating the relative abundance of *Paeniclostridium* and its interactions with host immunities. Our findings and speculations require further *in vivo* and *in vitro* models to validate the interaction mechanism between *Paeniclostridium* and STEC O157 as well as the role of *Paeniclostridium* on the hindgut mucosa microbiome and host. Regardless, our findings provide fundamental knowledge of STEC O157 colonization mechanisms and how microbial communities and the host participated in response to the production of different Shiga toxin subtypes, the main virulence factors involved in pathogenesis.

5.6 References

Anderson, M. J., Crist, T. O., Chase, J. M., Vellend, M., Inouye, B. D., Freestone, A. L., et al. (2011). Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. *Ecol Lett* 14, 19–28. doi: 10.1111/j.1461-0248.2010.01552.x.

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., et al. (2019). Author Correction: Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* 37, 1091–1091. doi: 10.1038/s41587-019-0252-6.

Braga, R. M., Dourado, M. N., and Araújo, W. L. (2016). Microbial interactions: ecology in a molecular perspective. *Braz J Microbiol* 47, 86–98. doi: 10.1016/j.bjm.2016.10.005.

Bretschneider, G., Berberov, E. M., and Moxley, R. A. (2007a). Isotype-specific antibody responses against Escherichia coli O157:H7 locus of enterocyte effacement proteins in adult beef cattle following experimental infection. *Vet Immunol Immunop* 118, 229–238. doi: 10.1016/j.vetimm.2007.06.005.

Bretschneider, G., Berberov, E. M., and Moxley, R. A. (2007b). Reduced intestinal colonization of adult beef cattle by Escherichia coli O157:H7 tir deletion and nalidixic-acid-resistant mutants lacking flagellar expression. *Vet Microbiol* 125, 381–386. doi: 10.1016/j.vetmic.2007.06.009.

Browne, H. P., Almeida, A., Kumar, N., Vervier, K., Adoum, A. T., Viciani, E., et al. (2021). Host adaptation in gut Firmicutes is associated with sporulation loss and altered transmission cycle. *Genome Biol* 22, 204. doi: 10.1186/s13059-021-02428-6.

Buffie, C. G., and Pamer, E. G. (2013). Microbiota-mediated colonization resistance against intestinal pathogens. *Nat Rev Immunol* 13, 790–801. doi: 10.1038/nri3535.

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* 13, 581–583. doi: 10.1038/nmeth.3869.

Carscadden, K. A., Emery, N. C., Arnillas, C. A., Cadotte, M. W., Afkhami, M. E., Gravel, D., et al. (2020). Niche Breadth: Causes and Consequences for Ecology, Evolution, and Conservation. *Q Rev Biology* 95, 179–214. doi: 10.1086/710388.

Chase, J. M., Kraft, N. J. B., Smith, K. G., Vellend, M., and Inouye, B. D. (2011). Using null models to disentangle variation in community dissimilarity from variation in α-diversity. *Ecosphere* 2, 1–11. doi: 10.1890/es10-00117.1.

Chase-Topping, M., Gally, D., Low, C., Matthews, L., and Woolhouse, M. (2008). Super-shedding and the link between human infection and livestock carriage of Escherichia coli O157. *Nat Rev Microbiol* 6, 904–912. doi: 10.1038/nrmicro2029.

Clark, R. L., Connors, B. M., Stevenson, D. M., Hromada, S. E., Hamilton, J. J., Amador-Noguez, D., et al. (2021). Design of synthetic human gut microbiome assembly and butyrate production. *Nat Commun* 12, 3254. doi: 10.1038/s41467-021-22938-y.

Dean-Nystrom, E. A., Bosworth, B. T., Cray, W. C., and Moon, H. W. (1997). Pathogenicity of Escherichia coli O157:H7 in the intestines of neonatal calves. *Infect Immun* 65, 1842–1848. doi: 10.1128/iai.65.5.1842-1848.1997.

Deng, W., Marshall, N. C., Rowland, J. L., McCoy, J. M., Worrall, L. J., Santos, A. S., et al. (2017). Assembly, structure, function and regulation of type III secretion systems. *Nat Rev Microbiol* 15, 323–337. doi: 10.1038/nrmicro.2017.20.

Driessche, L. V., Vanneste, K., Bogaerts, B., Keersmaecker, S. C. J. D., Roosens, N. H., Haesebrouck, F., et al. (2020). Isolation of Drug-Resistant Gallibacterium anatis from Calves with Unresponsive Bronchopneumonia - Volume 26, Number 4—April 2020 - Emerging Infectious Diseases journal - CDC. *Emerg Infect Dis* 26, 721–730. doi: 10.3201/eid2604.190962.

Faust, K., Sathirapongsasuti, J. F., Izard, J., Segata, N., Gevers, D., Raes, J., et al. (2012). Microbial Co-occurrence Relationships in the Human Microbiome. *Plos Comput Biol* 8, e1002606. doi: 10.1371/journal.pcbi.1002606.

Ferreira-Halder, C. V., Faria, A. V. de S., and Andrade, S. S. (2017). Action and function of Faecalibacterium prausnitzii in health and disease. *Best Pract Res Clin Gastroenterology* 31, 643–648. doi: 10.1016/j.bpg.2017.09.011.

Fitzgerald, S. F., Beckett, A. E., Palarea-Albaladejo, J., McAteer, S., Shaaban, S., Morgan, J., et al. (2019). Shiga toxin sub-type 2a increases the efficiency of Escherichia coli O157

transmission between animals and restricts epithelial regeneration in bovine enteroids. *Plos Pathog* 15, e1008003. doi: 10.1371/journal.ppat.1008003.

Fraser, M. E., Fujinaga, M., Cherney, M. M., Melton-Celsa, A. R., Twiddy, E. M., O'Brien, A. D., et al. (2004). Structure of Shiga Toxin Type 2 (Stx2) from Escherichia coli O157:H7*. *J Biol Chem* 279, 27511–27517. doi: 10.1074/jbc.m401939200.

Furman, O., Shenhav, L., Sasson, G., Kokou, F., Honig, H., Jacoby, S., et al. (2020). Stochasticity constrained by deterministic effects of diet and age drive rumen microbiome assembly dynamics. *Nat Commun* 11, 1904. doi: 10.1038/s41467-020-15652-8.

Gaytán, M. O., Martínez-Santos, V. I., Soto, E., and González-Pedrajo, B. (2016). Type Three Secretion System in Attaching and Effacing Pathogens. *Front Cell Infect Mi* 6, 129. doi: 10.3389/fcimb.2016.00129.

Guimerà, R., and Amaral, L. A. N. (2005). Functional cartography of complex metabolic networks. *Nature* 433, 895–900. doi: 10.1038/nature03288.

Herren, C. M., and McMahon, K. D. (2018). Keystone taxa predict compositional change in microbial communities. *Environ Microbiol* 20, 2207–2217. doi: 10.1111/1462-2920.14257.

Hoffman, M. A., Menge, C., Casey, T. A., Laegreid, W., Bosworth, B. T., and Dean-Nystrom, E. A. (2006). Bovine Immune Response to Shiga-Toxigenic Escherichia coli O157:H7. *Clin Vaccine Immunol* 13, 1322–1327. doi: 10.1128/cvi.00205-06.

Huang, C., Ge, F., Yao, X., Guo, X., Bao, P., Ma, X., et al. (2021). Microbiome and Metabolomics Reveal the Effects of Different Feeding Systems on the Growth and Ruminal Development of Yaks. *Front Microbiol* 12, 682989. doi: 10.3389/fmicb.2021.682989.

Ilhan, Z. E., Łaniewski, P., Tonachio, A., and Herbst-Kralovetz, M. M. (2020). Members of Prevotella Genus Distinctively Modulate Innate Immune and Barrier Functions in a Human Three-Dimensional Endometrial Epithelial Cell Model. *J Infect Dis* 222, 2082–2092. doi: 10.1093/infdis/jiaa324.

Iljazovic, A., Roy, U., Gálvez, E. J. C., Lesker, T. R., Zhao, B., Gronow, A., et al. (2021). Perturbation of the gut microbiome by Prevotella spp. enhances host susceptibility to mucosal inflammation. *Mucosal Immunol* 14, 113–124. doi: 10.1038/s41385-020-0296-4.

Jansma, J., and Aidy, S. E. (2021). Understanding the host-microbe interactions using metabolic modeling. *Microbiome* 9, 16. doi: 10.1186/s40168-020-00955-1.

Ji, H., and Dong, H. (2015). Key steps in type III secretion system (T3SS) towards translocon assembly with potential sensor at plant plasma membrane. *Mol Plant Pathol* 16, 762–773. doi: 10.1111/mpp.12223.

Johnson, D. R., Goldschmidt, F., Lilja, E. E., and Ackermann, M. (2012). Metabolic specialization and the assembly of microbial communities. *Isme J* 6, 1985–1991. doi: 10.1038/ismej.2012.46.

Kaper, J. B., Nataro, J. P., and Mobley, H. L. T. (2004). Pathogenic Escherichia coli. *Nat Rev Microbiol* 2, 123–140. doi: 10.1038/nrmicro818.

Karson, M. (1968). Handbook of Methods of Applied Statistics. Volume I: Techniques of Computation Descriptive Methods, and Statistical Inference. Volume II: Planning of Surveys and Experiments. I. M. Chakravarti, R. G. Laha, and J. Roy, New York, John Wiley; 1967, $9.00. *J Am Stat Assoc* 63, 1047–1049. doi: 10.1080/01621459.1968.11009335.

Kim, H.-J., and Song, W.-J. (2022). Inactivation of Escherichia coli O157: H7 in foods by emerging technologies: a review. *Lett Appl Microbiol* 76. doi: 10.1093/lambio/ovac007.

Larsen, J. M. (2017). The immune response to Prevotella bacteria in chronic inflammatory disease. *Immunology* 151, 363–374. doi: 10.1111/imm.12760.

Levins, R. (n.d.). *Evolution in Changing Environments*. Princeton University Press doi: 10.1515/9780691209418.

Li, X., He, L., Luo, J., Zheng, Y., Zhou, Y., Li, D., et al. (2022). Paeniclostridium sordellii hemorrhagic toxin targets TMPRSS2 to induce colonic epithelial lesions. *Nat Commun* 13, 4331. doi: 10.1038/s41467-022-31994-x.

Liao, J., Cao, X., Zhao, L., Wang, J., Gao, Z., Wang, M. C., et al. (2016). The importance of neutral and niche processes for bacterial community assembly differs between habitat generalists and specialists. *Fems Microbiol Ecol* 92, fiw174. doi: 10.1093/femsec/fiw174.

Liu, J., Meng, Z., Liu, X., and Zhang, X.-H. (2019). Microbial assembly, interaction, functioning, activity and diversification: a review derived from community compositional data. *Mar Life Sci Technology* 1, 112–128. doi: 10.1007/s42995-019-00004-3.

Liu, S., Yu, H., Yu, Y., Huang, J., Zhou, Z., Zeng, J., et al. (2022). Ecological stability of microbial communities in Lake Donghu regulated by keystone taxa. *Ecol. Indic.* 136, 108695. doi: 10.1016/j.ecolind.2022.108695.

Logares, R., Lindström, E. S., Langenheder, S., Logue, J. B., Paterson, H., Laybourn-Parry, J., et al. (2013). Biogeography of bacterial communities exposed to progressive long-term environmental change. *Isme J* 7, 937–948. doi: 10.1038/ismej.2012.168.

Louis, P., and Flint, H. J. (2017). Formation of propionate and butyrate by the human colonic microbiota. *Environ. Microbiol.* 19, 29–41. doi: 10.1111/1462-2920.13589.

Majowicz, S. E., Scallan, E., Jones-Bitton, A., Sargeant, J. M., Stapleton, J., Angulo, F. J., et al. (2014). Global Incidence of Human Shiga Toxin–Producing Escherichia coli Infections and Deaths: A Systematic Review and Knowledge Synthesis. *Foodborne Pathog. Dis.* 11, 447–455. doi: 10.1089/fpd.2013.1704.

Malmuthuge, N., and Guan, L. L. (2017). Understanding host-microbial interactions in rumen: searching the best opportunity for microbiota manipulation. *J Anim Sci Biotechno* 8, 8. doi: 10.1186/s40104-016-0135-3.

Malmuthuge, N., Liang, G., Griebel, P. J., and Guan, L. L. (2019). Taxonomic and Functional Compositions of the Small Intestinal Microbiome in Neonatal Calves Provide a Framework for Understanding Early Life Gut Health. *Appl Environ Microb* 85. doi: 10.1128/aem.02534-18.

Mao, S., Zhang, M., Liu, J., and Zhu, W. (2015). Characterising the bacterial microbiota across the gastrointestinal tracts of dairy cattle: membership and potential function. *Sci Rep-uk* 5, 16116. doi: 10.1038/srep16116.

Melton-Celsa, A. R. (2014). Shiga Toxin (Stx) Classification, Structure, and Function. *Microbiol Spectr* 2. doi: 10.1128/microbiolspec.ehec-0024-2013.

Menge, C. (2020). The Role of Escherichia coli Shiga Toxins in STEC Colonization of Cattle. *Toxins* 12, 607. doi: 10.3390/toxins12090607.

Menge, C., Stamm, I., Wuhrer, M., Geyer, R., Wieler, L. H., and Baljer, G. (2001). Globotriaosylceramide (Gb3/CD77) is synthesized and surface expressed by bovine lymphocytes upon activation in vitro. *Vet Immunol Immunop* 83, 19–36. doi: 10.1016/s0165-2427(01)00365-8.

Mir, R. A., Schaut, R. G., Allen, H. K., Looft, T., Loving, C. L., Kudva, I. T., et al. (2019). Cattle intestinal microbiota shifts following Escherichia coli O157:H7 vaccination and colonization. *Plos One* 14, e0226099. doi: 10.1371/journal.pone.0226099.

Mootha, V. K., Lindgren, C. M., Eriksson, K.-F., Subramanian, A., Sihag, S., Lehar, J., et al. (2003). PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet* 34, 267–273. doi: 10.1038/ng1180.

Munns, K. D., Selinger, L. B., Stanford, K., Guan, L., Callaway, T. R., and McAllister, T. A. (2015). Perspectives on Super-Shedding of Escherichia coli O157:H7 by Cattle. *Foodborne Pathog Dis* 12, 89–103. doi: 10.1089/fpd.2014.1829.

Newman, M. E. J. (2006). Modularity and community structure in networks. *Proc National Acad Sci* 103, 8577–8582. doi: 10.1073/pnas.0601602103.

Nicholson, J. K., Holmes, E., Kinross, J., Burcelin, R., Gibson, G., Jia, W., et al. (2012). Host-Gut Microbiota Metabolic Interactions. *Science* 336, 1262–1267. doi: 10.1126/science.1223813.

O'Hara, E., Neves, A. L. A., Song, Y., and Guan, L. L. (2020). The Role of the Gut Microbiome in Cattle Production and Health: Driver or Passenger? *Annu Rev Anim Biosci* 8, 199–220. doi: 10.1146/annurev-animal-021419-083952.

Pan, Z., Chen, Y., McAllister, T. A., Gänzle, M., Plastow, G., and Guan, L. L. (2021a). Abundance and Expression of Shiga Toxin Genes in Escherichia coli at the Recto-Anal Junction Relates to Host Immune Genes. *Front Cell Infect Mi* 11, 633573. doi: 10.3389/fcimb.2021.633573.

Pan, Z., Chen, Y., Zhou, M., McAllister, T. A., and Guan, L. L. (2021b). Microbial interaction-driven community differences as revealed by network analysis. *Comput Struct Biotechnology J* 19, 6000–6008. doi: 10.1016/j.csbj.2021.10.035.

Pandit, S. N., Kolasa, J., and Cottenie, K. (2009). Contrasts between habitat generalists and specialists: an empirical extension to the basic metacommunity framework. *Ecology* 90, 2253–2262. doi: 10.1890/08-0851.1.

Panel, E. B., Koutsoumanis, K., Allende, A., Alvarez-Ordóñez, A., Bover-Cid, S., Chemaly, M., et al. (2020). Pathogenicity assessment of Shiga toxin-producing Escherichia coli (STEC) and the public health risk posed by contamination of food with STEC. *Efsa J* 18. doi: 10.2903/j.efsa.2020.5967.

Pryde, S. E., Duncan, S. H., Hold, G. L., Stewart, C. S., and Flint, H. J. (2002). The microbiology of butyrate formation in the human colon. *Fems Microbiol Lett* 217, 133–139. doi: 10.1111/j.1574-6968.2002.tb11467.x.

Raa, H., Grimmer, S., Schwudke, D., Bergan, J., Wälchli, S., Skotland, T., et al. (2009). Glycosphingolipid Requirements for Endosome-to-Golgi Transport of Shiga Toxin. *Traffic* 10, 868–882. doi: 10.1111/j.1600-0854.2009.00919.x.

Ravasz, E., Somera, A. L., Mongru, D. A., Oltvai, Z. N., and Barabási, A.-L. (2002). Hierarchical Organization of Modularity in Metabolic Networks. *Science* 297, 1551–1555. doi: 10.1126/science.1073374.

Reimand, J., Isserlin, R., Voisin, V., Kucera, M., Tannus-Lopes, C., Rostamianfar, A., et al. (2019). Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. *Nat Protoc* 14, 482–517. doi: 10.1038/s41596-018-0103-9.

Sandhu, K. S., and Gyles, C. L. (2002). Pathogenic Shiga toxin-producing Escherichia coli in the intestine of calves. *Can J Vet Res Revue Can De Recherche Veterinaire* 66, 65–72.

Shahi, S. K., Freedman, S. N., and Mangalam, A. K. (2017). Gut microbiome in multiple sclerosis: The players involved and the roles they play. *Gut Microbes* 8, 607–615. doi: 10.1080/19490976.2017.1349041.

Sheng, H., Wang, J., Lim, J. Y., Davitt, C., Minnich, S. A., and Hovde, C. J. (2011). Internalization of Escherichia Coli O157:H7 by Bovine Rectal Epithelial Cells. *Front Microbiol* 2, 32. doi: 10.3389/fmicb.2011.00032.

Silva, Y. P., Bernardi, A., and Frozza, R. L. (2020). The Role of Short-Chain Fatty Acids From Gut Microbiota in Gut-Brain Communication. *Front Endocrinol* 11, 25. doi: 10.3389/fendo.2020.00025.

Stegen, J. C., Lin, X., Konopka, A. E., and Fredrickson, J. K. (2012). Stochastic and deterministic assembly processes in subsurface microbial communities. *ISME J.* 6, 1653–1664. doi: 10.1038/ismej.2012.22.

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-

wide expression profiles. *Proc National Acad Sci* 102, 15545–15550. doi: 10.1073/pnas.0506580102.

Sun, L., Jia, H., Li, J., Yu, M., Yang, Y., Tian, D., et al. (2019). Cecal Gut Microbiota and Metabolites Might Contribute to the Severity of Acute Myocardial Ischemia by Impacting the Intestinal Permeability, Oxidative Stress, and Energy Metabolism. *Front Microbiol* 10, 1745. doi: 10.3389/fmicb.2019.01745.

Tian, L., Zhang, Y., Zhang, L., Zhang, L., Gao, X., and Feng, B. (2023). Biogeographic Pattern and Network of Rhizosphere Fungal and Bacterial Communities in Panicum miliaceum Fields: Roles of Abundant and Rare Taxa. *Microorg* 11, 134. doi: 10.3390/microorganisms11010134.

Tian, S., Liu, Y., Wu, H., Liu, H., Zeng, J., Choi, M. Y., et al. (2020). Genome-Wide CRISPR Screen Identifies Semaphorin 6A and 6B as Receptors for Paeniclostridium sordellii Toxin TcsL. *Cell Host Microbe* 27, 782-792.e7. doi: 10.1016/j.chom.2020.03.007.

Vasco, K., Nohomovich, B., Singh, P., Venegas-Vargas, C., Mosci, R. E., Rust, S., et al. (2021). Characterizing the Cattle Gut Microbiome in Farms with a High and Low Prevalence of Shiga Toxin Producing Escherichia coli. *Microorg* 9, 1737. doi: 10.3390/microorganisms9081737.

Vidor, C. J., Bulach, D., Awad, M., and Lyras, D. (2019). Paeniclostridium sordellii and Clostridioides difficile encode similar and clinically relevant tetracycline resistance loci in diverse genomic locations. *Bmc Microbiol* 19, 53. doi: 10.1186/s12866-019-1427-5.

Wang, L., Chauliac, D., Moritz, B. E., Zhang, G., Ingram, L. O., and Shanmugam, K. T. (2019). Metabolic engineering of Escherichia coli for the production of butyric acid at high titer and productivity. *Biotechnol. Biofuels* 12, 62. doi: 10.1186/s13068-019-1408-9.

Wang, O., Liang, G., McAllister, T. A., Plastow, G., Stanford, K., Selinger, B., et al. (2016). Comparative Transcriptomic Analysis of Rectal Tissue from Beef Steers Revealed Reduced Host Immunity in Escherichia coli O157:H7 Super-Shedders. *Plos One* 11, e0151284. doi: 10.1371/journal.pone.0151284.

Wang, O., McAllister, T. A., Plastow, G., Stanford, K., Selinger, B., and Guan, L. L. (2017). Host mechanisms involved in cattle Escherichia coli O157 shedding: a fundamental understanding for reducing foodborne pathogen in food animal production. *Sci Rep-uk* 7, 7630. doi: 10.1038/s41598-017-06737-4.

Wang, O., McAllister, T. A., Plastow, G., Stanford, K., Selinger, B., and Guan, L. L. (2018). Interactions of the Hindgut Mucosa-Associated Microbiome with Its Host Regulate Shedding of Escherichia coli O157:H7 by Cattle. *Appl Environ Microb* 84, e01738-17. doi: 10.1128/aem.01738-17.

Wolfe, A. W. (1997). Social Network Analysis: Methods and Applications. *Am Ethnol* 24, 219–220. doi: 10.1525/ae.1997.24.1.219.

Wu, W., Logares, R., Huang, B., and Hsieh, C. (2017). Abundant and rare picoeukaryotic sub-communities present contrasting patterns in the epipelagic waters of marginal seas in the northwestern Pacific Ocean. *Environ Microbiol* 19, 287–300. doi: 10.1111/1462-2920.13606.

WU, Y., HINENOYA, A., TAGUCHI, T., NAGITA, A., SHIMA, K., TSUKAMOTO, T., et al. (2010). Distribution of Virulence Genes Related to Adhesins and Toxins in Shiga Toxin-Producing Escherichia coli Strains Isolated from Healthy Cattle and Diarrheal Patients in Japan. *J Vet Med Sci* 72, 589–597. doi: 10.1292/jvms.09-0557.

Xu, M., Huang, Q., Xiong, Z., Liao, H., Lv, Z., Chen, W., et al. (2021). Distinct Responses of Rare and Abundant Microbial Taxa to In Situ Chemical Stabilization of Cadmium-Contaminated Soil. *Msystems* 6, e01040-21. doi: 10.1128/msystems.01040-21.

Xu, Q., Vandenkoornhuyse, P., Li, L., Guo, J., Zhu, C., Guo, S., et al. (2022). Microbial generalists and specialists differently contribute to the community diversity in farmland soils. *J Adv Res* 40, 17–27. doi: 10.1016/j.jare.2021.12.003.

Xu, Y., Dugat-Bony, E., Zaheer, R., Selinger, L., Barbieri, R., Munns, K., et al. (2014). Escherichia coli O157:H7 Super-Shedder and Non-Shedder Feedlot Steers Harbour Distinct Fecal Bacterial Communities. *Plos One* 9, e98115. doi: 10.1371/journal.pone.0098115.

Xue, Y., Du, M., Sheng, H., Hovde, C. J., and Zhu, M.-J. (2017). Escherichia coli O157:H7 suppresses host autophagy and promotes epithelial adhesion via Tir-mediated and cAMP-

independent activation of protein kinase A. *Cell Death Discov* 3, 17055. doi: 10.1038/cddiscovery.2017.55.

Xun, W., Li, W., Xiong, W., Ren, Y., Liu, Y., Miao, Y., et al. (2019). Diversity-triggered deterministic bacterial assembly constrains community functions. *Nat Commun* 10, 3833. doi: 10.1038/s41467-019-11787-5.

Zhang, Y., Wu, G., Jiang, H., Yang, J., She, W., Khan, I., et al. (2018). Abundant and Rare Microbial Biospheres Respond Differently to Environmental and Spatial Factors in Tibetan Hot Springs. *Front Microbiol* 9, 2096. doi: 10.3389/fmicb.2018.02096.

Zhou, J., Liu, W., Deng, Y., Jiang, Y.-H., Xue, K., He, Z., et al. (2013). Stochastic Assembly Leads to Alternative Communities with Distinct Functions in a Bioreactor Microbial Community. *Mbio* 4, e00584-12. doi: 10.1128/mbio.00584-12.

Zhou, J., and Ning, D. (2017). Stochastic Community Assembly: Does It Matter in Microbial Ecology? *Microbiol Mol Biol R* 81, e00002-17. doi: 10.1128/mmbr.00002-17.

Zhu, M., Qi, X., Yuan, Y., Zhou, H., Rong, X., Dang, Z., et al. (2023). Deciphering the distinct successional patterns and potential roles of abundant and rare microbial taxa of urban riverine plastisphere. *J Hazard Mater* 450, 131080. doi: 10.1016/j.jhazmat.2023.131080.

5.7 Tables and figures

Table 5.1. The average relative abundance of bacterial phylum across each group

Group	CT			WT			RE			P		
Time	T1	T2	T5	T1	T2	T5	T1	T2	T5	Group	Time	Group*Time
Actinobacteriota	0.0081±0.0192	0.0136±0.0348	0.0065±0.0196	0.0053±0.0060	0.0149±0.0175	0.0002±0.0005	0.0004±0.0009	0.0114±0.0178	0.0005±0.0008	0.15	0.01*	0.02*
Bacteroidota	0.0502±0.0337	0.0502±0.0328	0.0823±0.0402	0.0370±0.0256	0.0573±0.0281	0.0306±0.0279	0.0735±0.0573	0.0699±0.0300	0.0698±0.0452	0.03*	0.51	0.04*
Firmicutes	0.8265±0.1078	0.6851±0.3084	0.7383±0.2057	0.8163±0.0766	0.6935±0.0871	0.7982±0.1095	0.7339±0.2323	0.7664±0.1457	0.7958±0.1675	0.56	0.31	0.67
Proteobacteria	0.0780±0.1013	0.1913±0.3262	0.1399±0.1686	0.0398±0.0294	0.1756±0.0967	0.0398±0.0312	0.1542±0.2389	0.1164±0.0960	0.1049±0.1567	0.71	0.35	0.45

The PROC MIXED models was used to determine factors affecting the differences of bacterial phylum across each group ($P<0.05$ as significant).

Table 5.2. The network topological properties of microbial interactions among rectal mucosa microbial communities among three groups.

Group	Week	Nodes	Edges	Modularity	Average degree	Clustering coefficient
	T1	132	775	0.46	11.74	0.56
CT	T2	163	906	0.55	11.12	0.72
	T5	144	1173	0.49	16.29	0.66
	T1	105	300	0.73	5.71	0.80
WT	T2	117	494	0.75	8.44	0.83
	T5	97	261	0.72	5.38	0.63
	T1	116	550	0.52	9.48	0.67
RE	T2	122	642	0.54	10.52	0.73
	T5	95	353	0.69	7.43	0.70

Table 5.3. The attributions of microbial genera based on their network roles.

	Group	Week			χ^2 test	
		T1	**T2**	**T5**	χ^2	***p***
Connectors	CT	0	4	3	N/A	0.45
	WT	4	3	3		
	RE	2	3	1		
Peripherals	CT	132	160	142	2.1	0.71
	WT	101	114	94		
	RE	114	120	94		

Table 5.4. The stratification of abundant-specific genera across each group.

Abundant	Group	Week			χ^2 test	
		T1	T2	T5	χ^2	p
High	CT	0.08 (n=11)	0.070 (n=12)	0.081 (n=12)	1.1	0.9
	WT	0.10 (n=11)	0.092 (n=11)	0.18 (n=18)		
	RE	0.094 (n=11)	0.097 (n=12)	0.13 (n=13)		
Moderate	CT	0.78 (n=104)	0.82 (n=134)	0.82 (n=122)	4.1	0.4
	WT	0.89 (n=93)	0.90 (n=106)	0.82 (n=80)		
	RE	0.89 (n=104)	0.90 (n=111)	0.87 (n=85)		
Low	CT	0.14 (n=18)	0.11 (n=18)	0.01 (n=14)	N/A	0.8
	WT	0.010 (n=1)	0.017 (n=2)	0 (n=0)		
	RE	0.016 (n=2)	0.0080 (n=1)	0 (n=0)		

Fisher's exact test was used as the count of one cell was below 5.

Table 5.5. The quantity of specialized microbes among microbial community assembly in CT, WT and RE groups from T1 to T5.

Attributes	Group	Week			χ^2 test	
		T1	T2	T5	χ^2	p
Generalists	CT	19	29	22	2.2	0.7
	WT	9	12	13		
	RE	10	10	6		
Non-significant	CT	95	126	112	3.7	0.4
	WT	91	101	76		
	RE	100	108	85		
Specialists	CT	19	9	14	2.6	0.6
	WT	5	6	9		
	RE	7	6	7		

Chi-square test was used to test the equality of numbers of specialized microbes.

Table 5.6. The quantity of host-microbial interactions in CT, WT, and RE groups from T1 to T5.

Group	Antigen		B cell		Chemo		IgA		Killer		MAPK		T cell	
	P	N	P	N	P	N	P	N	P	N	P	N	P	N
CT-T1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CT-T2	0	0	0	0	0	0	0	0	0	0	0	1	0	0
CT-T5	0	0	0	0	0	0	0	0	0	0	0	0	0	0
WT-T1	11	10	11	10	43	24	10	6	22	8	49	35	22	15
WT-T2	3	3	3	14	6	18	0	6	4	11	6	17	2	22
WT-T5	9	9	29	14	44	45	10	4	28	22	59	82	27	25
RE-T1	1	0	1	0	2	0	1	0	2	0	3	0	1	0
RE-T2	0	0	0	0	0	0	0	0	1	0	1	1	0	0
RE-T5	1	5	6	11	13	22	4	1	3	17	4	29	3	16

The value in each cell represents the quantify of interactions between genes involved in host immune pathways and microbes. Absolute $R>0.8$ and $P<0.01$ as the cut off for significant interactions. P and N stand for the positive and negative interactions, respectively. Antigen, B cell, Chemo, IgA, Killer, MAPK, T cell refer to antigen processing and presentation, B cell receptor signaling pathway, chemokine signaling pathway, intestinal immune network for IgA production, natural killer cell mediated cytotoxicity, MAPK signaling pathway, T cell receptor signaling pathway, respectively.

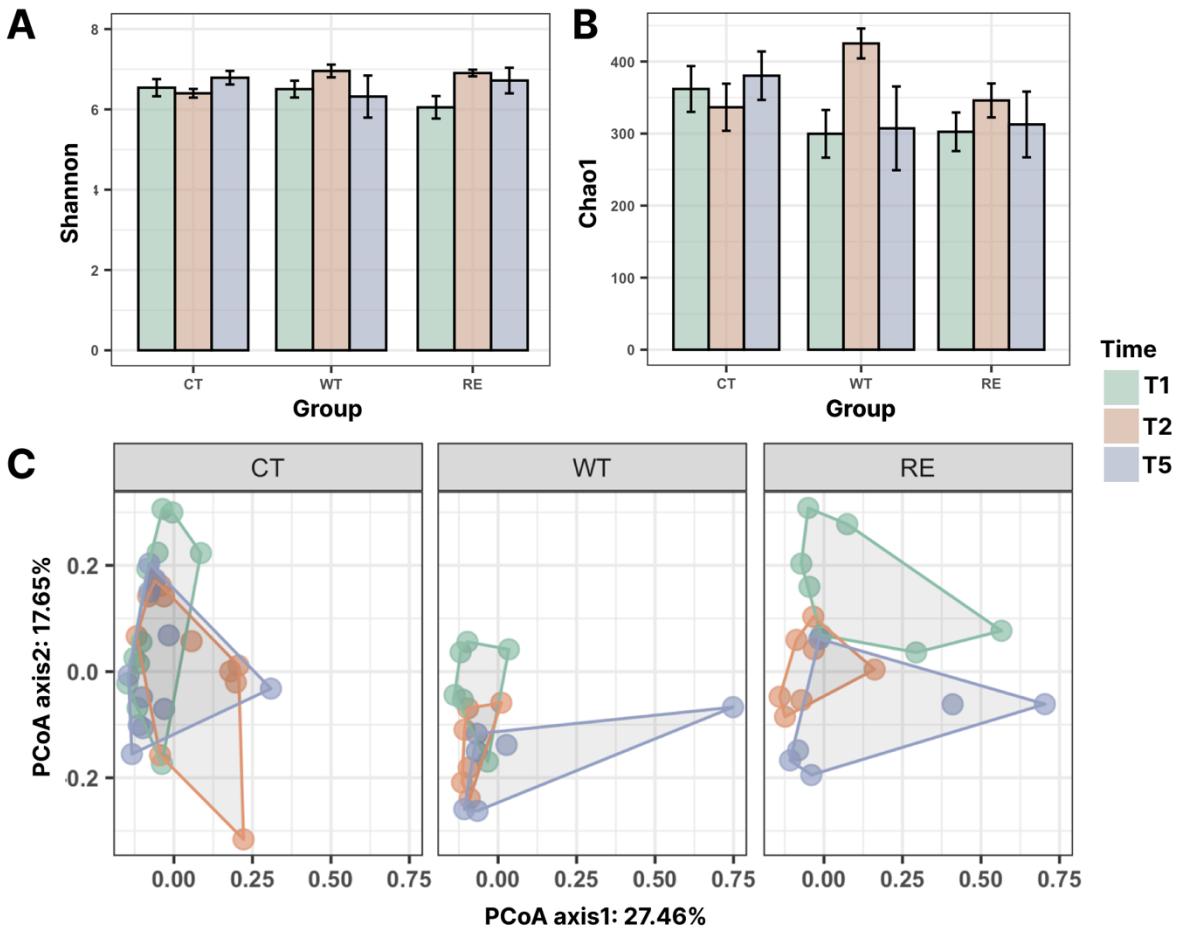


Figure 5.1. Comparison of diversity metrics among three groups from pre- (T1) to post- (T5) challenge.

Shannon (A) and Chao1 (B) indices were used to estimate the evenness and richness across three groups with three different colors of bars representing samples collected from T1, T2, and T5. The horizontal bars within boxes represent medians. The Kruskal-Wallis test was used to determine whether indices between any two groups were significant. ($P \leq 0.05$).

C. Principal coordinate analysis (PCoA) was used for the visualization of the Bray-Curtis distance. The PERMANOVA was used to test for the similarity of clustering patterns among different ages within each group. Differences were considered significant at $P \leq 0.05$.

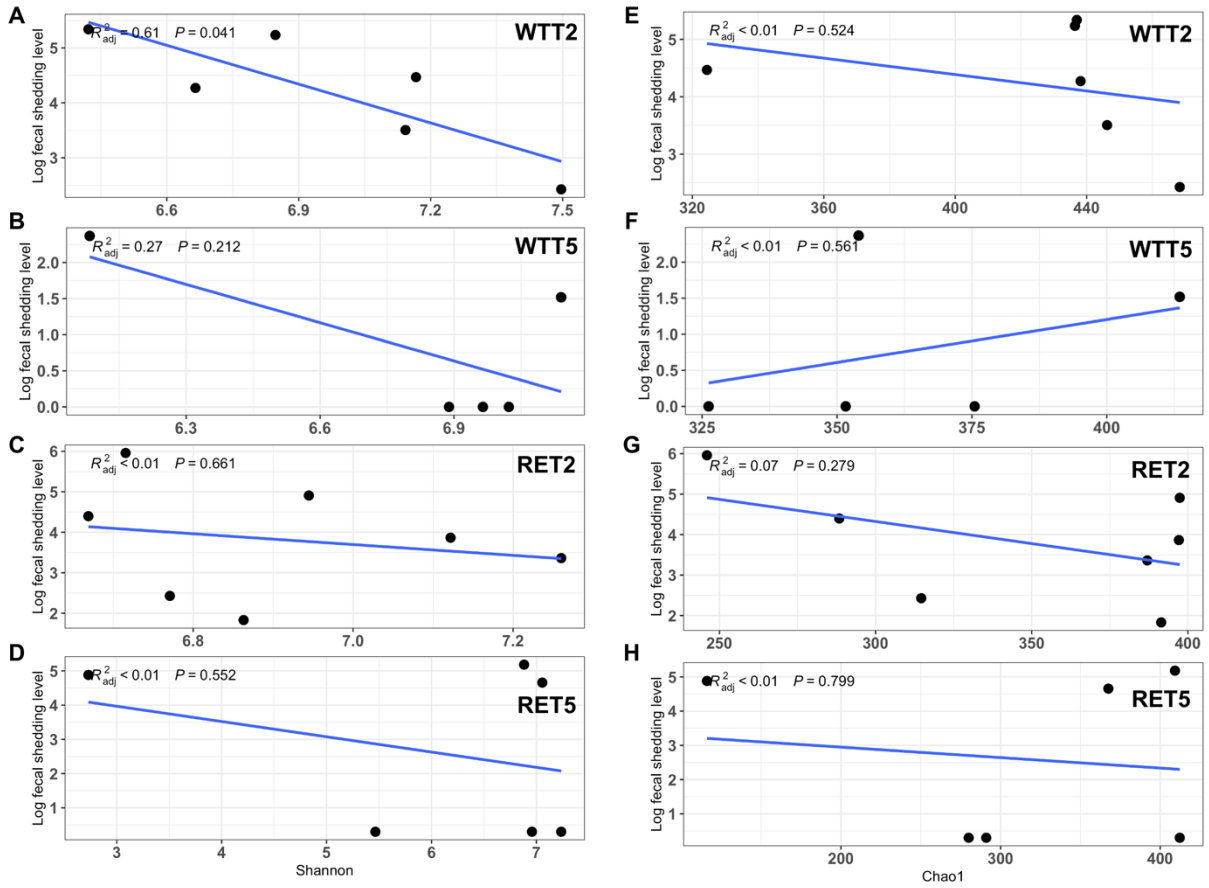


Figure 5.2. Relations between microbial diversities and log_{10} STEC O157 fecal shedding in calves in WT and RE groups from T1 to T5, respectively.

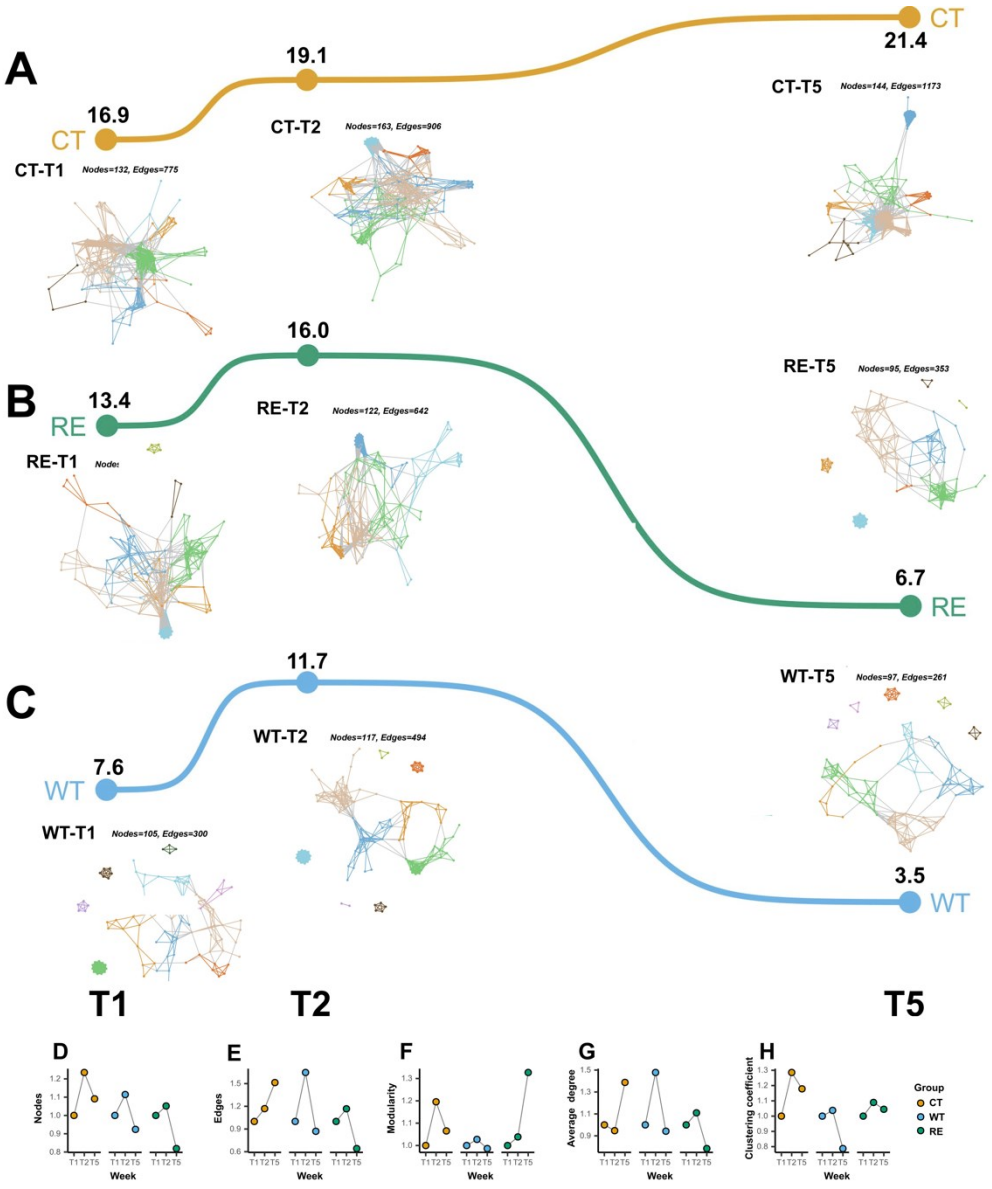


Figure 5.3. Dynamic microbial interactions were revealed by network analysis in the CT, RE, and WT groups at different ages.

A, B, and C represents networks constructed in CT, RE, and WT groups, respectively. The line represents network stability changes with values near the line and microbial interactions were plotted below each line. The changing patterns of network properties including nodes (D), edges (E), modularity (F), average degree (G), and clustering coefficient (H) were visualized from within each group.

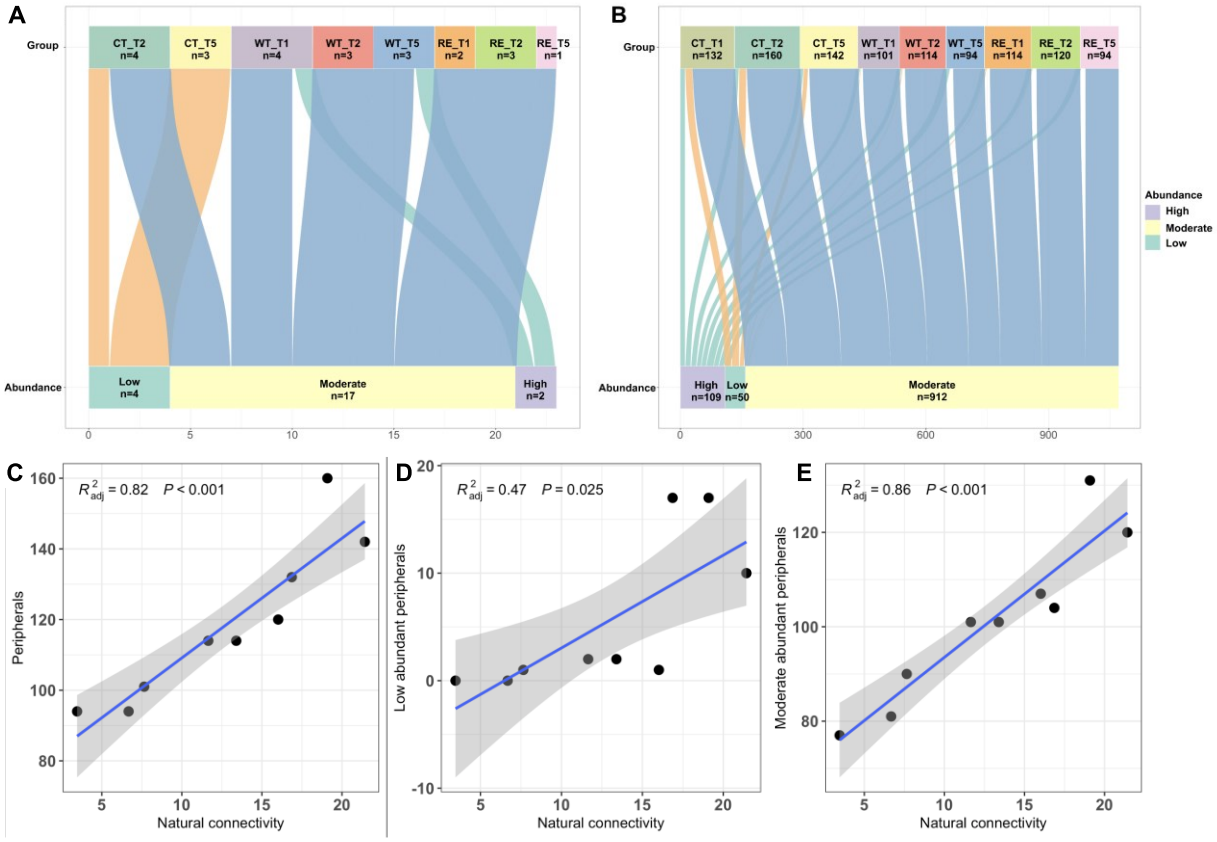


Figure 5.4. Attributes of network connectors and peripherals to the abundant-specific genera and significant relations between network stabilities.

Attributes of network connectors (A) and peripherals (B) to the abundant-specific genera and significant relations between network stabilities and network peripherals (C), low abundant (D), and moderate abundant (E) peripherals using linear regression models.

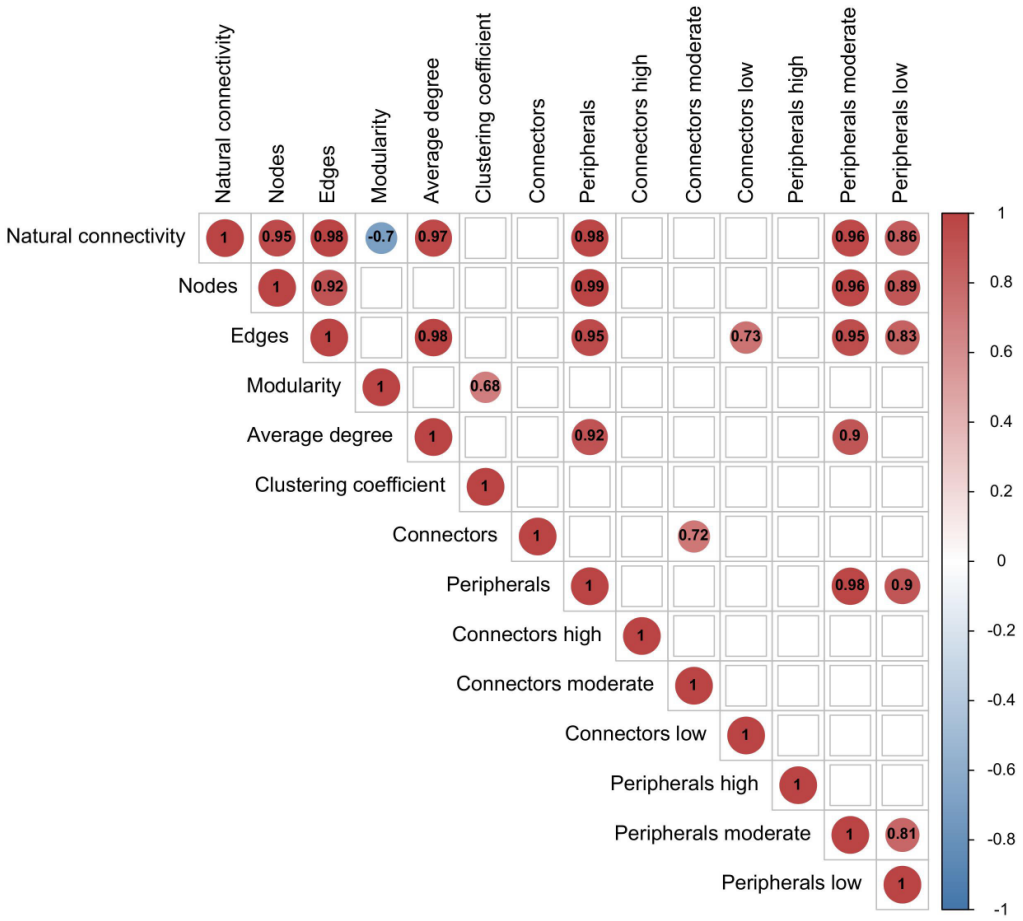


Figure 5.5. Spearman rank-based correlations between network stabilities and network properties and abundant-specific connectors and peripherals.

The absolute R> 0.6 and P<0.05 were considered as significant. Only significant correlations were shown in the plot.

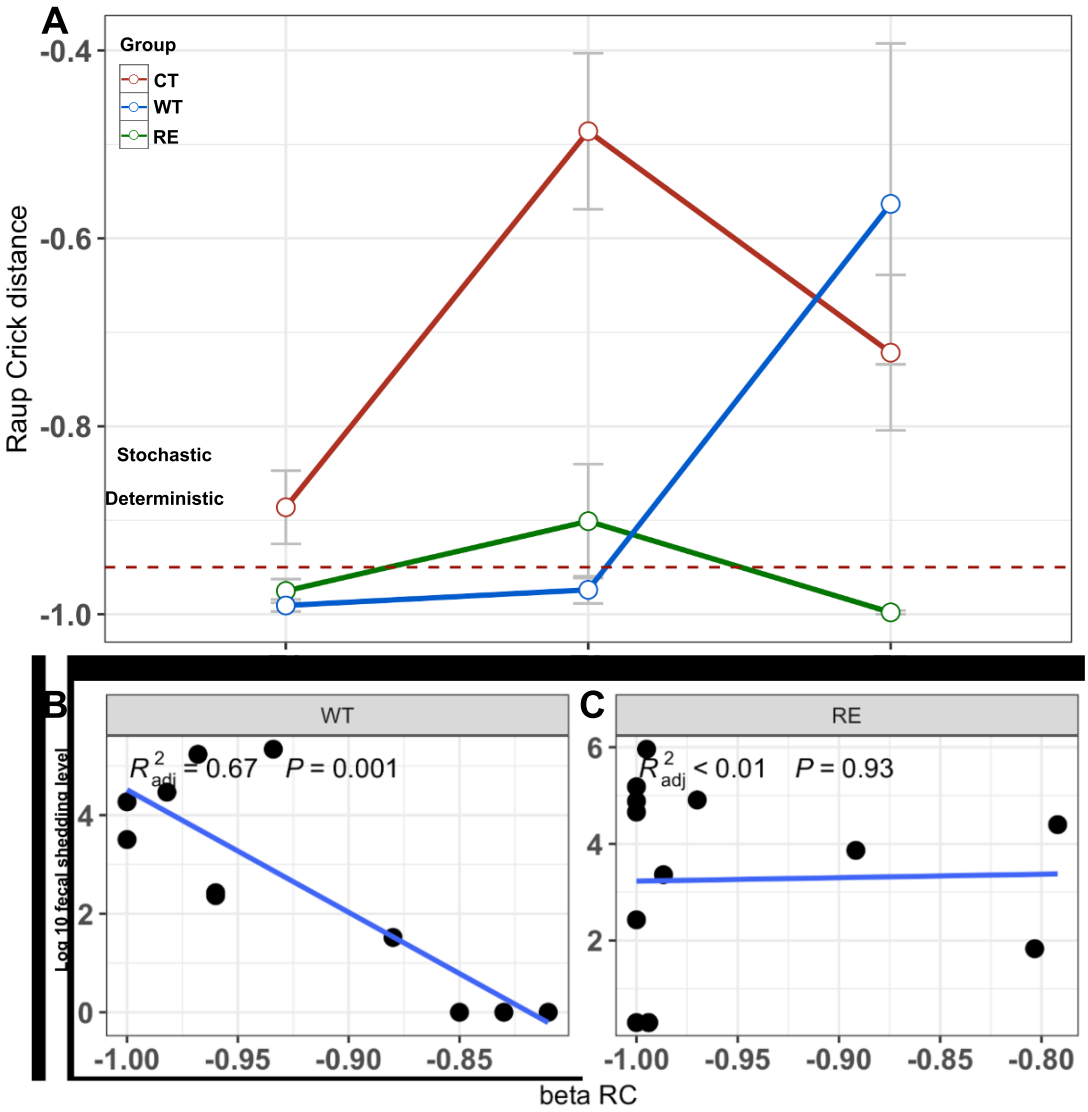


Figure 5.6. Microbial assembly patterns determined by Raup-Crick distance and its relations with log $_{10}$ STEC O157 fecal shedding level.

The dynamics of microbial community assembly patterns in three groups from T1 to T5 (A). The dotted horizontal line represents the boundary line that separates assembly patterns being more stochastic- (above) or deterministic- (below) driven. The linear regression model shows the relations between Raup-Crick distance and $\log_{10}$ STEC O157 fecal shedding in the WT (B) and RE (C) groups.

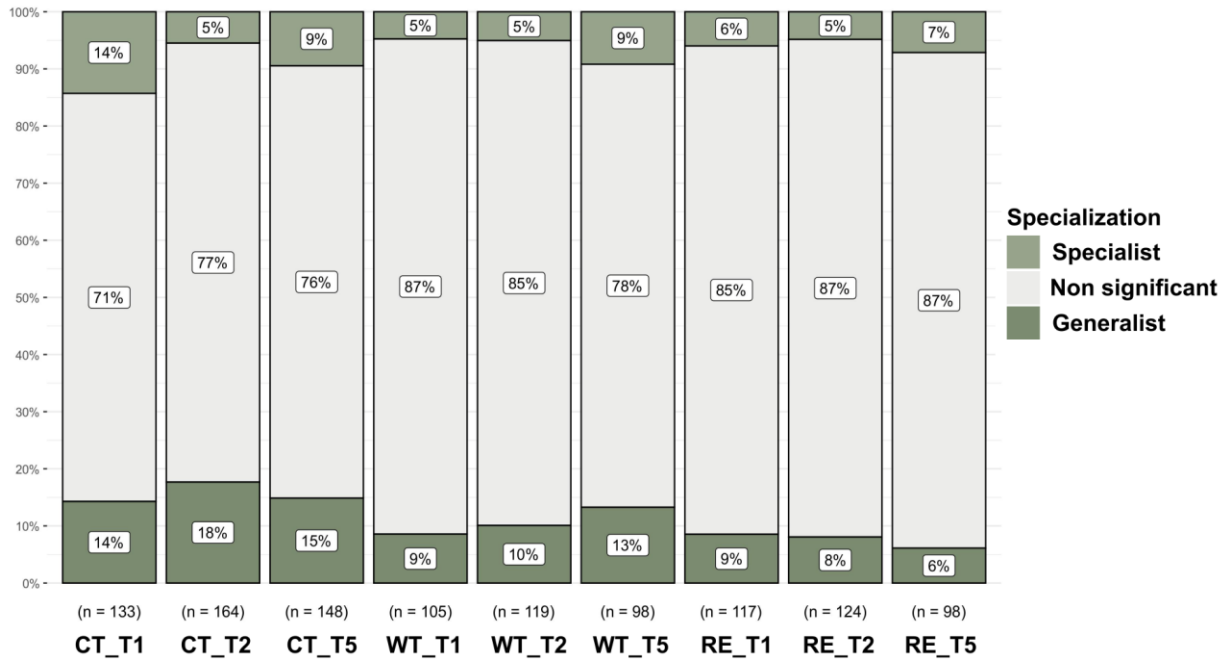


Figure 5.7. The specialization patterns of mucosa-attached microbes during microbiome assembly from T1 to T5 for CT, WT, and RE. The proportion of microbes belonging to specialists (on the top of each column) and generalists (on the bottom of each column) were assessed and microbes without specializations were considered as non-significant in the plot.

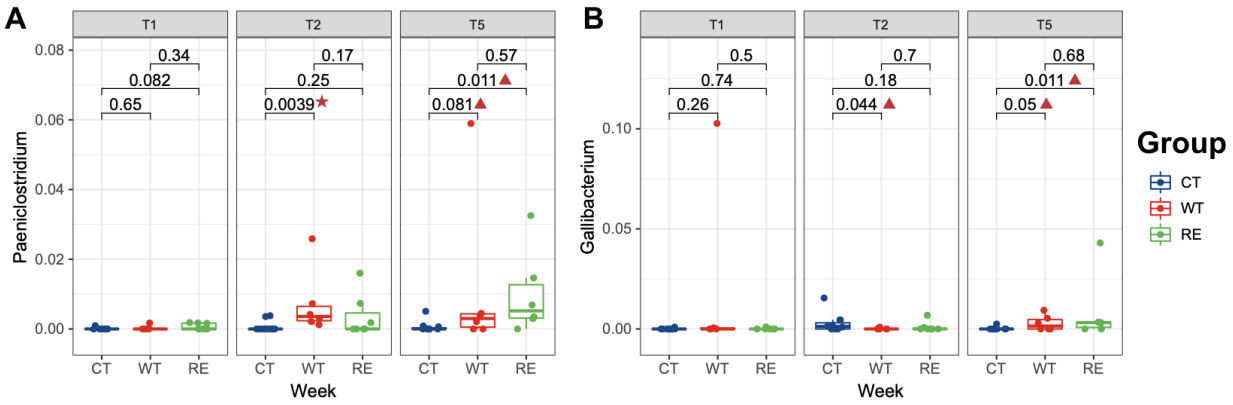


Figure 5.8. The comparison of the relative abundance of *Paeniclostridium* and *Gallibacterium* across each group from T1 to T5.

(A) and (B) refer to the comparison for the relative abundance *Paeniclostridium* and *Gallibacterium*. P< 0.01 as significance (marked as a triangle) and P<0.05 as a trend being significant (marked as a star).

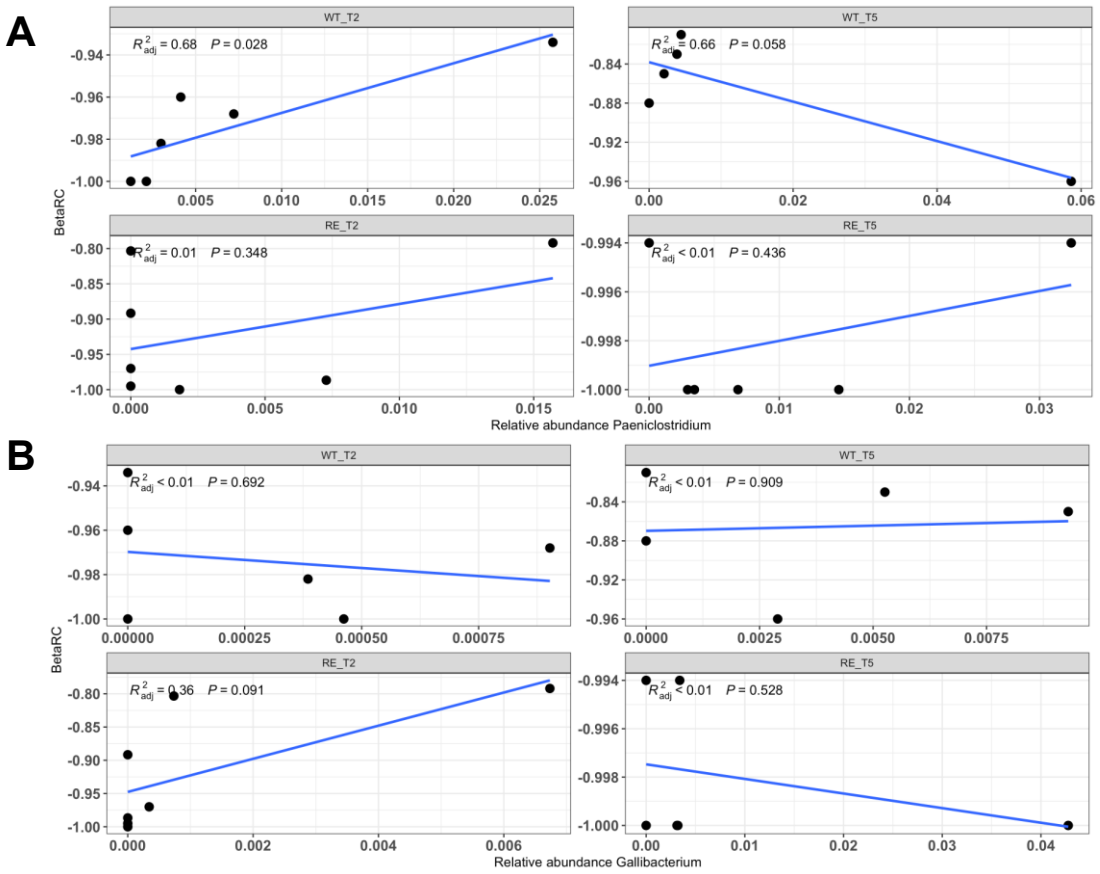


Figure 5.9. The linear regression models showing the relative abundance of *Paeniclostridium* and *Gallibacterium* and their relationships with Raup-Crick distance in WT and RE groups at T2 and T5.

The linear regression models showing the relative abundance of *Paeniclostridium* (A) and *Gallibacterium* (B) and their relationships with Raup-Crick distance in WT (first and third row) and RE (second and fourth row) groups at T2 (left column) and T5 (right column).

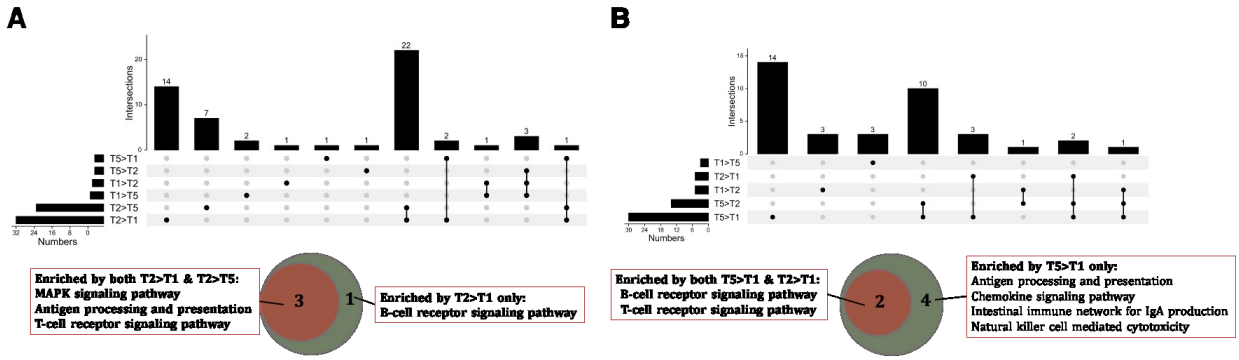


Figure 5.10. The upset plot shows intersections of host pathways response to STEC O157 colonization in WT and RE groups.

The Venn plots showed the interaction for WT (A) and RE (B) groups, and below refers to intersections of host immune-related pathways. For example, T2>T1 refers to the host immune pathways that are upregulated in T2 compared to T1. Values on top of each column in the upset plot refer to the number of enriched pathways.

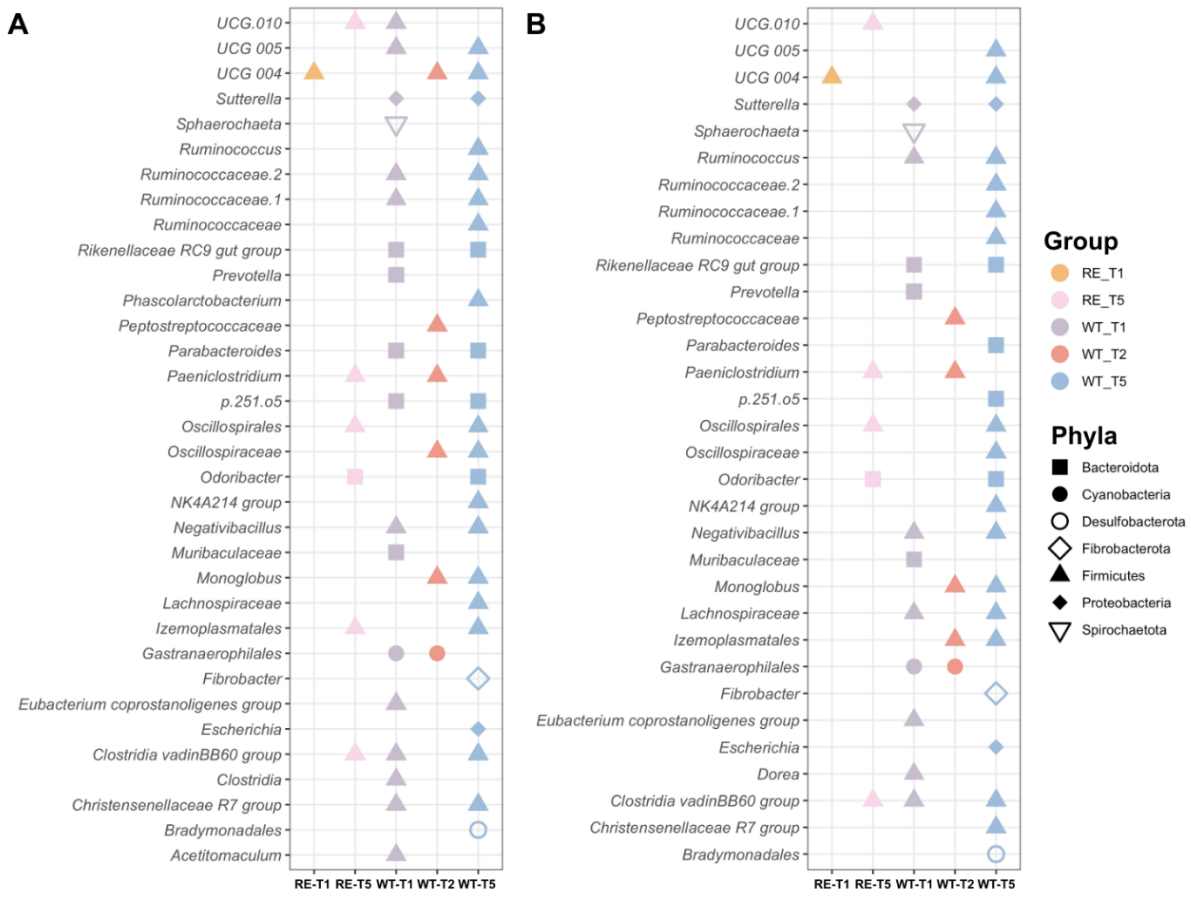


Figure 5.11. Significant interactions between host-immune related pathways and the relative abundance of rectal mucosal microbes.

Host genes involved in T cell receptor signaling pathway (A) and B-cell receptor signaling pathway (B).

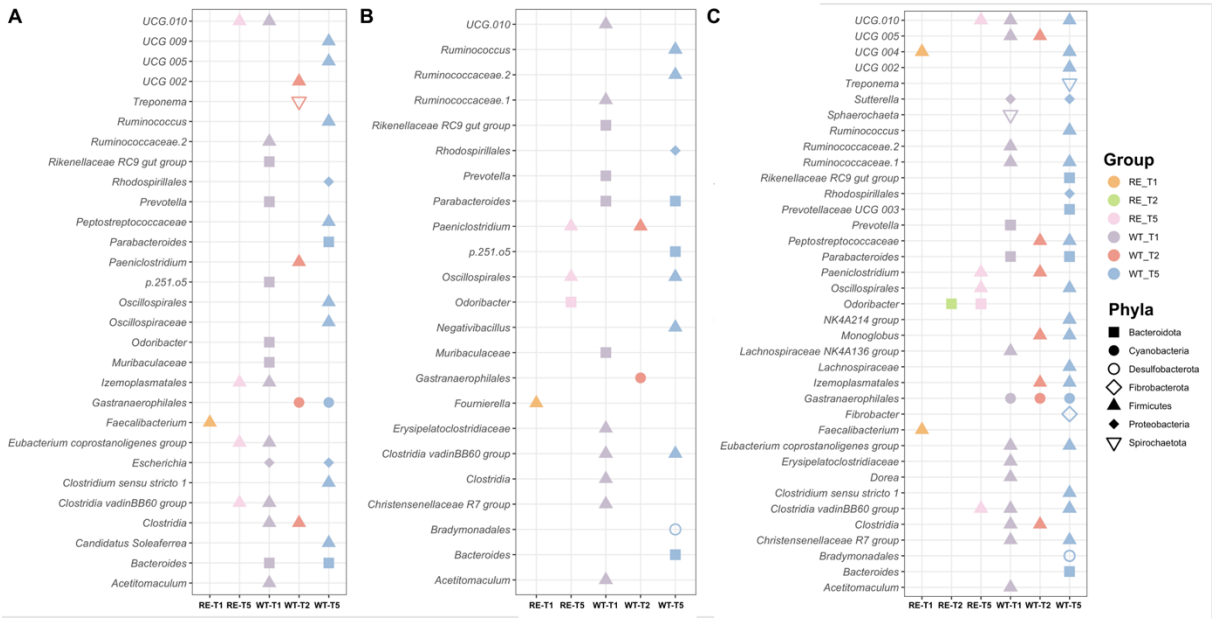


Figure 5.12. Significant interactions between host-immune related pathways and the relative abundance of rectal mucosal microbes.

Host genes involved in antigen processing and presentation (A) and intestinal immune network for IgA production (B), and natural killer cell mediated cytotoxicity (C).

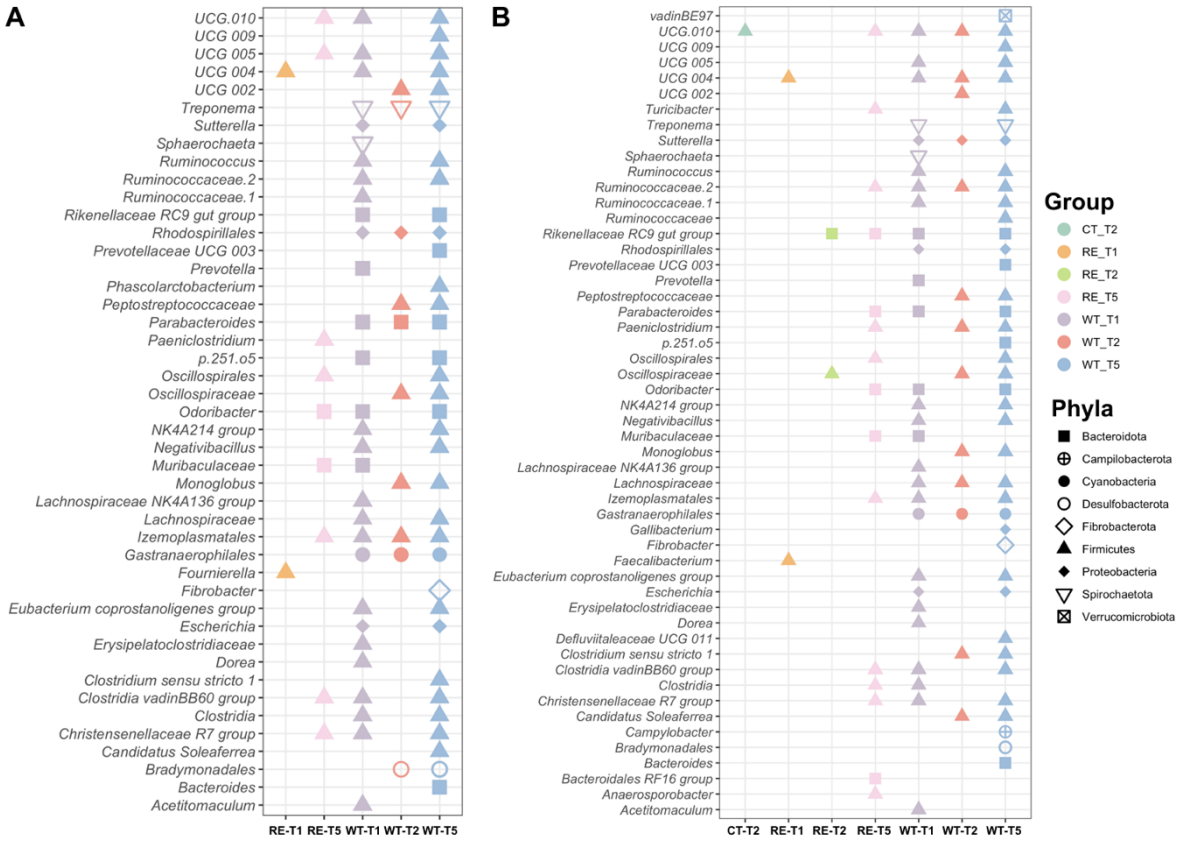


Figure 5.13. Significant interactions between host-immune related pathways and the relative abundance of rectal mucosal microbes.

Host genes involved in chemokine signaling pathway (A) and MAPK signaling pathway (B).

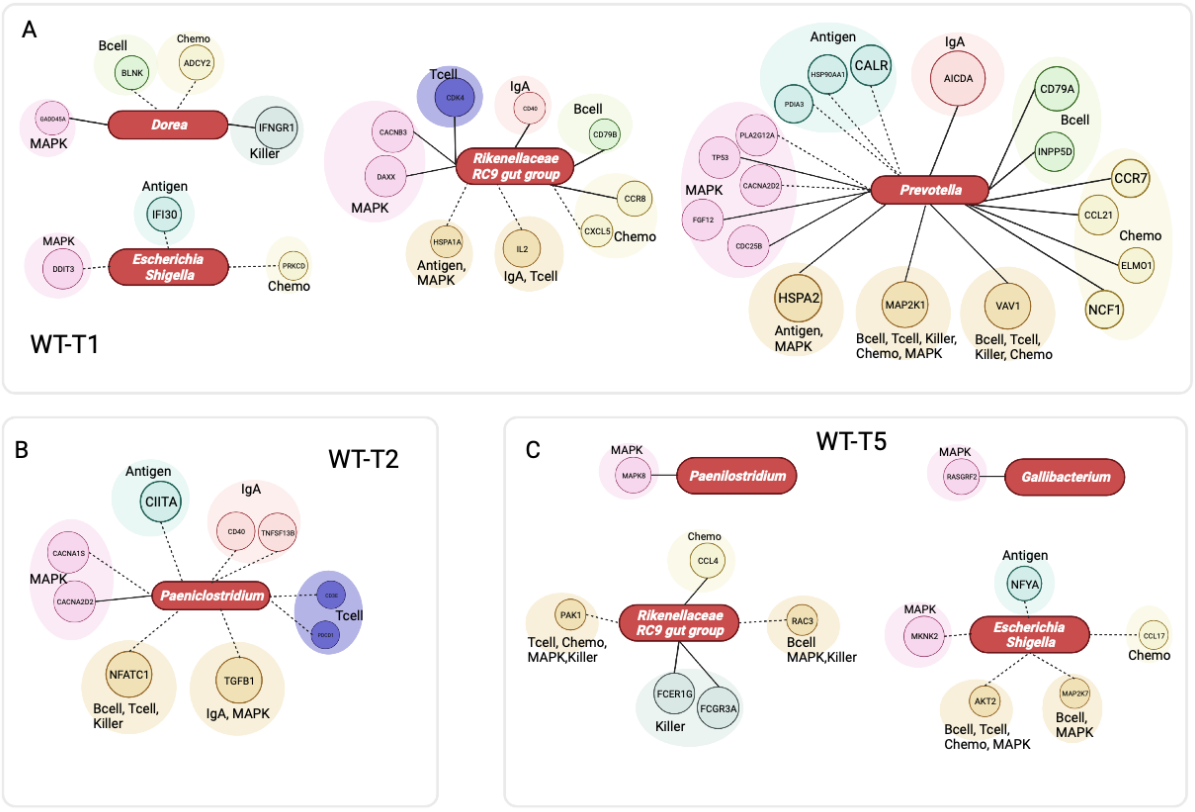


Figure 5.14. Interactions between selected microbes and host immune genes in the WT group across T1, T2, and T5.

The plots A, B, C refer to T1, T2, and T5, respectively. The solid line and dotted line refer to positive and negative interactions, respectively. The rod shape represents mucosal microbes and the circle refers to host immune-related genes with divergent colors representing different host immune-related pathways. For certain genes that were involved in more than one pathway, all pathways were marked. Antigen, Bcell, Chemo, MAPK, Killer, IgA, Tcell refer to antigen processing and presentation, B-cell receptor signaling pathway, chemokine signaling pathway, MAPK signaling pathway, natural killer cell mediated cytotoxicity, intestinal immune network for IgA production, T-cell receptor signaling pathway, respectively.

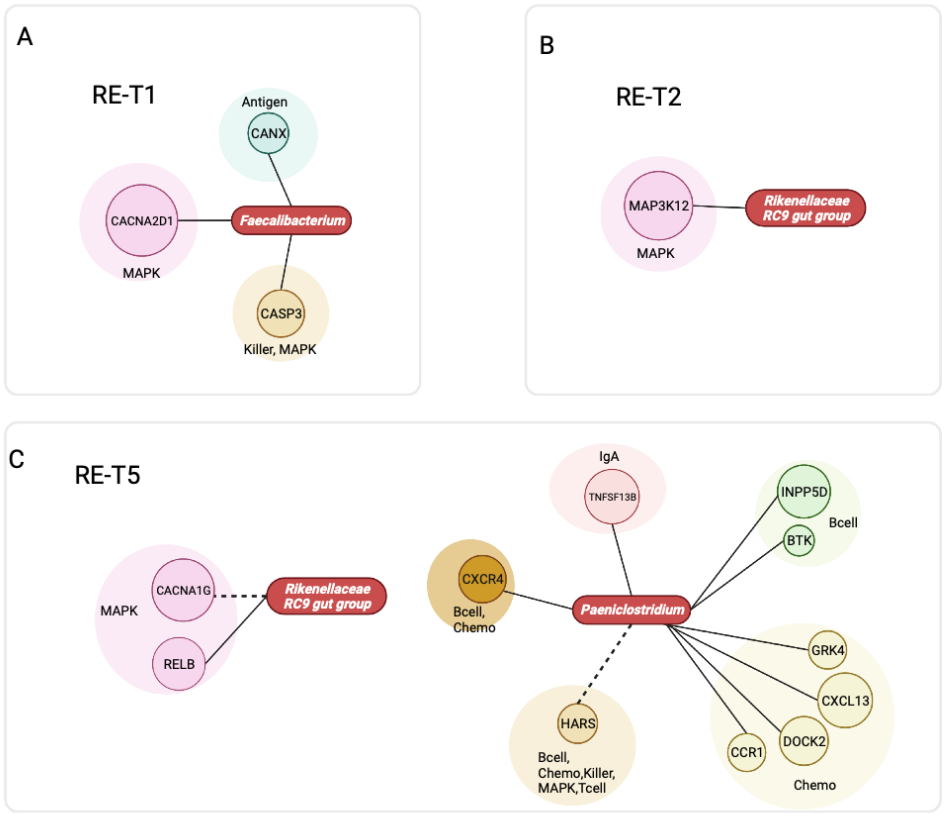


Figure 5.15. Interactions between selected microbes and host immune genes in the RE group across T1, T2, and T5.

The plots A, B, C refer to T1, T2, and T5, respectively. The solid line and dotted line refer to positive and negative interactions, respectively. The rod shape represents mucosal microbes and the circle refers to host immune-related genes with divergent colors representing different host immune-related pathways. For certain genes that were involved in more than one pathway, all pathways were marked. Antigen, Bcell, Chemo, MAPK, Killer, IgA, Tcell refer to antigen processing and presentation, B-cell receptor signaling pathway, chemokine signaling pathway, MAPK signaling pathway, natural killer cell mediated cytotoxicity, intestinal immune network for IgA production, T-cell receptor signaling pathway, respectively.

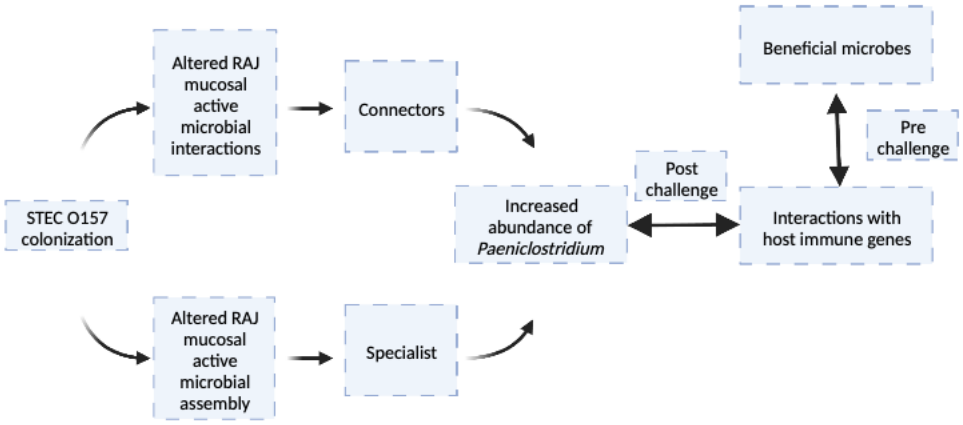


Figure 5.16. The proposed pathogen-gut commensals-host model for host-microbial interactions upon STEC O157 colonization in veal calves.

The diagram is created using Biorender (www. biorender.com).

Chapter 6. General discussion

Shiga toxin producing *E.coli* (STEC) colonization in beef cattle is a complex process, of which super-shedders (shedding more than 10^4 CFU/g *E.coli* O157) are associated with cattle-human transmission (Xu et al., 2014; Cote et al., 2015; Munns et al., 2015; McCabe et al., 2019). Identification of the prevalence of Shiga toxins in STEC and how their expression affects hindgut fecal and mucosal-attached microbiota at rectal-anal junction (RAJ), and host responses could provide fundamental knowledge on STEC colonization in beef cattle and to the development of on-farm intervention strategies. In this thesis, I performed research using integrated omics and molecular approaches to reveal how STEC colonization and expression of *stx* gene affect host-microbial interactions at the RAJ. The epidemiological surveys suggested that abundance (DNA) and expressions (RNA) of *stx1* and *stx2* were not evenly identified from fecal and RAJ mucosa samples (Chapter 2). Particularly, the *stx2* gene expression in bacteria could affect fecal microbial interactions as compared to those that lacked expression (Chapter 3). Furthermore, we revealed that STEC O157 colonization and *stx2a* expression influences host responses using the veal calf challenge model with strain-specific STEC O157 (stx2a+ vs. stx2a-) (Chapters 4 and 5). In addition, the STEC O157 challenge also affected gut homeostasis through the regulation of RAJ mucosal attached microbial community, its interactions and assembly (Chapter 5). Additionally, *stx2a* expression could be a factor that contributes to the varied interactions between rectal mucosa microbes and expression of host immune genes.

6.1 Varied abundance and expressions of *stx1* and *stx2* in STEC from feces and RAJ mucosa

In Chapter 2, the abundance (DNA) of *stx1* and *stx2* were compared between fecal and RAJ mucosal samples. The findings indicate that the abundance of *stx1* and *stx2* differed between fecal and RAJ mucosal samples and *stx* in STEC from mucosal samples were more abundant, suggesting that STEC were colonizing the RAJ. This finding is in accordance with previous report that copy numbers of *E.coli* O157 were inconsistent between the RAJ epithelium and feces (Durso et al., 2010). In addition, *stx* abundance was found to be affected by beef cattle breeds (Angus, Charolais, Kinsella Composite) and sampling year. Different breeds of beef cattle have remarkably different genetic backgrounds, and such variation may alter the host gut environment, which could be the reason for varied copy numbers of *stx*. Taken together, for most current studies, researchers tend to use fecal samples to quantify STEC copy numbers and *stx*, but such an approach can fail to reveal the 'true' abundance of STEC and Shiga toxin genes and neglects other factors that may affect these observations. Therefore, in Chapter 2, we suggested using both fecal and mucosal swabs or biopsies for accurate estimation of stx gene abundance in beef cattle.

In Chapter 2, we demonstrated that the expression of *stx2* was identified only in mucosal samples, while the *stx1* expression was not identified in STEC from either fecal or mucosal samples. As the stx2 toxin is 400 times more toxic than stx1 and its expression is commonly associated with human diseases (Schmidt et al., 2000; Fraser et al., 2004). When STEC

colonizes the RAJ epithelium, the adherence factor intimin, encoded by *eae* enables STEC colonization (Farfan and Torres, 2012). Our study identified that expressions of *eae* and *stx2* genes were simultaneously detected in STEC from mucosal samples, suggesting that the expression of *stx2* may act as a cofactor with *eae* to contribute to higher mucosal colonization of STEC in beef cattle.

6.2 The *stx2* expression relates to fecal and RAJ mucosal microbiome

As mentioned above, we realized that the expression of *stx2* is vital for STEC colonization, and previous studies highlighted that fecal microbial communities were altered during STEC colonization compared to healthy beef steers (Xu et al., 2014). Therefore, we assessed if the expression of *stx2* affects fecal microbial community profiles (Chapter 3). Our findings indicate that the expression of *stx2* did not affect the structure or profile of fecal microbial communities rather, fecal microbial interactions were altered, and microbial community stability was decreased in response to *stx2* expression. Several microbes were only present in the *stx2* expression group and were found to be the keystone taxa connecting other microbes within microbial interactions. This is the first research to reveal the ecological role of *stx2* on the fecal microbiome. Our work highlighted that expression of *stx2* could change fecal microbial community interactions and decrease the stability of fecal microbial communities.

In Chapter 3 we did not identify whether production of stx2 subtypes altered the profile of microbial communities at the rectal mucosa. Therefore, we assessed the shift of RAJ mucosal attached microbiome responses to the production of stx2 subtypes (stx2a, the most commonly

associated with human disease and found in SS (Fitzgerald et al., 2019)) using an *in vivo* challenge model (Chapter 5). Our results first indicate that the challenge of STEC O157 significantly shifted mucosal microbial diversities and similarities, whereas the production of stx2a did not alter microbial structures. Instead, microbial interactions and network properties (*i.e.* nodes, edges, average degree, clustering coefficients) were significantly changed upon stx2a+ STEC colonization as compared to unchallenged calves. Secondly, the production of stx2a in STEC was considered as a deterministic factor that shaped the rectal mucosa microbial assembly, which in turn affected microbial responses to STEC O157 colonization. Taken together, we explored the role of stx2, particularly stx2a in STEC on hindgut lumen and mucosal microbiota from multiple ecological perspectives, contributing to the fundamental knowledge of STEC O157-microbiota interactions.

6.3 The expression of *stx2* in STEC relates to host responses and host-microbial interactions

Previous studies identified alterations in host gene expression in SS compared to NS, indicating that host immune functions may inhibit STEC shedding in SS (Wang et al., 2016, 2017). However, how host genes relate to expressions of *stx* in STEC has not been studied. Hence, in Chapter 2, we selected four host immune genes previously reported to be downregulated in SS including chemokine (C-C motif) ligand 21 (*CCL21*), lymphotoxin beta (*LTB*), CD19 molecule (*CD19*), and 4-domains, subfamily A, member 1 (*MS4A1*) and investigated their relationship with the expression of *stx2*. We found that the expression of *stx2* could be predicted by the

expression of four host immune genes, with *MS4A1* being the most predictive. These results highlighted that stx2+ STEC colonization in beef cattle could affect host immune gene expressions. However, considering the fact that *stx2* possesses several subtypes with stx2a being the most prevalent toxin associated with SS and human infections (Lisboa et al., 2019). And previous SS studies neglect the effect of subtype stx2 on host responses and how host responses vary in addition to host immune responses. In Chapter 4, we assessed host transcriptomic response variations in terms of stx2a+/stx2a- STEC O157 (different in terms of the ability of the stx2a production) colonization using a veal calf challenge model. This study monitored fecal shedding of STEC O157 from pre- to 26 days post-challenge and found that challenged calves (regardless of stx2a+/ stx2a- STEC) shed most STEC O157 7 days post-challenge (Fitzgerald et al., 2019). The previous study also found that host rectal and systemic STEC-specific antibody responses changed significantly 7 days post challenge (Fitzgerald et al., 2019). This suggests that the dynamics of O157 fecal shedding was influenced by host responses. Therefore, we examined that changing patterns of host transcriptomic expression from pre- (T1) to 7 days (T2) and 26 days (T5) post-challenge.

We revealed that the production of stx2a in STEC affected the kinetics of host responses post-challenge. In particular, stx2a- STEC O157 colonization resulted in the down regulation of host functions involved in two major clusters at T2: host extracellular region and gut barrier integrity (Chapter 4), with upregulation of host immune-related pathways (Chapter 5). Whereas stx2a+ STEC O157 challenge did not induce any host responses as compared to unchallenged calves at T2 (Chapter 4) and only induced increased host immune functions at T2 (Chapter 5).

At T5, both stx2a- and stx2a+ STEC O157 induced similar host responses with enhanced functions from host extracellular region and gut barrier integrity clusters (Chapter 4) and upregulated host immune-related pathways (Chapter 5). The extracellular region refers to the outermost structure of the cell where *E.coli* O157 can colonize and interact with host cells, of which we found the IL-17 signaling pathway (T-cell related host response) was involved with its functions related to protection against bacterial infection at mucosal sites (Mills, 2023). Although previous studies rarely explored the T-cell mediated host adaptive immunities in STEC colonization, our results suggest that pathogenic STEC O157 colonization in calves affecst host T-cell mediated responses. In addition, the integrity of tissue barrier and epithelial regeneration has been found to be impacted by *E.coli* O157 colonization in calves and mice models (Roxas et al., 2010; Fitzgerald et al., 2019). *E.coli* O157 can induce localized effacement of microvilli causing attach/effacement lesions (A/E lesion) as a result of colonization, which is detrimental to the epithelium and causes a negative effect on tissue barrier integrity (Roxas et al., 2010). However, host-pathogen interactions were dynamic, of which hosts also defend itself to combat stxa2-/stx2a+ STEC O157 colonization on the rectal mucosa site as revealed by the increased host immune related pathways including B-cell and T-cell signaling pathway at T2 and T5. These two pathways were elevated in both stx2a- and stx2a+ STEC challenged calves, suggesting that changing host systemic antibody and cellular responses were not specific to the production of stx2a.

Our previous study revealed that hindgut mucosa-attached microbiota related to host immune responses in SS, constituting the basis for studying host-microbial interactions in terms

of stx2a-/stx2a+ STEC O157 colonization in veal calves. We found that beneficial microbes (*i.e. Prevotella*) dominate interactions with host immune gene expressions pre-challenge in both stx2a- and stx2a+ STEC O157 challenged calves, suggesting that they play a critical role in regulating host immune functions when host gut environment remains stable. While the opportunistic pathogen, *Paeniclostridium* bears such responsibility in stx2a- O157 challenged calves at T2, highlighting that the pathogen colonization could be a factor driven different host-microbial interactions. Therefore, in Chapter 5, we proposed a pathogen-gut commensal-host model, of which the pathogen colonization can initiate the alterations of gut microbial community profiles and such variations can further affect host responses and differentiate host-microbial interactions. Specifically, we found that the production of stx2a can be a factor affecting the proposed model, which can delay the occurrence of host- *Paeniclostridium* interactions and inhibit interactions between gut commensals and host immune genes, highlighting that stx2a production could shift host-microbial interactions instead of directly affecting host responses and gut microbial community structures.

6.4 Translational knowledge of advanced machine learning and statistical approach to STEC studies

Our studies adopted advanced approaches in machine learning and statistics to uncover how STEC colonization and expressions of *stx* interact with host and fecal-/mucosal microbiomes. Firstly, we used Isomap to differentiate host immune gene patterns in response to the expression of *stx2* (Chapter 2). Isomap is a dimensional reduction approach. As compared to principal

component analysis (PCA), it is less restrictive since it does not require any specific distribution (*i.e.* normal distribution) of data (Wang, 2012). It is also suitable for animal studies, since interactions among genetics, environment, and microbes are by nature nonlinear (Nicholson et al., 2004). Secondly, through the usage of random forest and Boruta methods, we determined the predictive accuracy of host immune gene expression on *stx2* expression (Chapter 2). The random forest and Boruta methods are machine-learning based methods that enable the robust classification of predictive genes/microbial markers from the dataset. Our study showed that the feasibility of adopting this approach to determine host immune gene markers for the prediction of *stx2* expression, which can be a potential method for other animal research programs to identify biomarkers for prediction of host phenotypes. Lastly, we constructed microbial interaction networks using Random matrix theory (RMT)-based method to reflect how microbial interaction varied in response to *stx2* gene expression in STEC (Chapter 3). This approach is different from correlation-based networks and can represent the true interaction within microbes and neglect spurious interactions due to the low abundant taxa or noise since it sets up a threshold to differentiate true interactions. Most of advanced statistical and machine learning methods have been developed in theory and our studies showed the feasibility of adopting such approaches to uncover biological questions. These novel approaches can solve problems that previous methods cannot effectively resolve, *i.e.* selections of potential microbial/genetic markers from high throughput dataset. This suggests that future research using advanced machine learning and statistical methods should be prompted to fully uncover biological questions.

6.5 Future directions

This is the first study showing how Shiga toxin 2 and its subtype in STEC affect host responses, fecal/rectal mucosal microbiome variations, and host-microbial interactions using high throughput sequencing. Although our studies shed light on the mechanisms of STEC colonization affecting host-microbial interactions, there are still some limitations.

First, we performed the epidemiological survey on DNA and RNA levels, no *stx* expression at the protein level was assessed. Therefore, information on how Shiga toxin proteins could affect host-microbial interactions is missing. Also, STEC is not the only bacteria that carries the *stx* gene, other bacteria such as *Enterobacter* have also been reported to produce Shiga toxins (Paton and Paton, 1996). Our study did not quantify STEC in Chapter 2 and 3, therefore, it is possible that the detection of *stx* gene expression may be due to other stx-carrying bacteria. Future studies are needed to quantify *stx* proteins and STEC to ensure *stx* gene expressions are functionally valid in STEC. Secondly, we only examined the effect of stx on host-microbial interactions. Previous studies reported that virulence factors such as T3SS could alter host immune responses and alter microbiomes (Bliska et al., 2013; Gaytán et al., 2016; Caballero-Flores et al., 2021). However, our study did not characterize virulence factors in addition to stx and how other virulence factors in STEC could potentially affect host-microbial interactions. Future studies on the effect of colonization of different STEC O157 serotypes on host-microbial interactions are needed. Thirdly, our challenge trial only observed STEC O157 fecal shedding from pre- to 26 days post-challenge. A previous study monitored *E. coli* O157

shedding patterns of beef steers for 77 days and found three shedding patterns including non-persistent (last ~7 days), moderately persistent (last ~30 days), and persistent shedders (last > 30 days)(Baines et al., 2008). However, samples collected from other time points post challenge were not assessed. Therefore, the challenged calves in my study may belong to different types of SS, and the differences in transcriptome profiles in the challenged calves may be due to the combined effect of different STEC O157 strains (stx2a+ vs. stx2a-) and different shedding types. Therefore, it is not possible to completely understand how hosts respond to strain-specific STEC O157 colonization. Also, the proposed effect of stx2a production on host responses cannot be confirmed, future studies using samples from other time points are needed to illustrate the complete dynamic changing patterns of host and microbial responses. Also, for future SS research, it is recommended to observe STEC O157 fecal shedding over a longer term, thus minimizing the effects of shedding types on host responses. Fourthly, although DEGs were identified between SS and NS, the genetic variations related to DEGs and potential linkage to Single Nucleotide Polymorphisms (SNPs) were not assessed. A previous study highlighted that SNPs play a role in the differential expression of genes between SS and NS (Wang et al., 2017., However, how the production of stx2a could interact with SNPs in DEGs requires further study. Lastly, our studies focused on host-microbial interactions using omics approaches and most conclusions are based on the bioinformatic analyses. Both *in vivo* (*i.e.* intestinal loop model (Vlisidou et al., 2004)) and *in vitro* models (*i.e.* bovine intestinal epithelial cell model (Miyazawa et al., 2010)) need to be employed to verify or refute our assumptions. Also, proteomic and metabolomic analysis should be performed to further define potential

metabolites and proteins related to host-microbial interactions upon strain-specific STEC O157 colonization. A previous study identified microRNAs (miRNAs) play a regulatory role in regulating host responses in SS compared to NS (Wang et al., 2021). Further studies using miRNAomes are needed to examine if miRNAs play such role in regulating host-microbial interactions upon strain-specific STEC O157 colonization.

6.6 Implications

The goal of my research is to reveal mechanisms of STEC O157 colonization and *stx2* in affecting host-microbial interactions, to identify potential microbial/genetic markers that could be useful for on farm identification of SS and to identify potential direct fed microbials that may reduce mucosal STEC colonization in beef cattle. Learning the mechanisms of STEC O157 colonization and expression of *stx2* could facilitate development of effective preharvest interventions, such as probiotics and direct fed microbials. Also, diagnostic tools using potential candidate genes could allow early and quick identification of SS and stx2 positive O157 strain, allowing real time treatments to minimize E. *coli* O157 colonization and fecal shedding prior to slaughter. By investigating *E.coli* O157-host-microial interaction, our ultimate goal is to develop methods that can effectively reduce fecal *E. coli* O157 and limit its cattle-human transmission, leading to reduced risk of beef recalls and human illness.

6.7 References

Akbar, S., Gu, L., Sun, Y., Zhang, L., Lyu, K., Huang, Y., et al. (2022). Understanding host-microbiome-environment interactions: Insights from Daphnia as a model organism. *Sci. Total Environ.* 808, 152093. doi: 10.1016/j.scitotenv.2021.152093.

Baines, D., Lee, B., and McAllister, T. (2008). Heterogeneity in enterohemorrhagic Escherichia coli O157:H7 fecal shedding in cattle is related to Escherichia coli O157:H7 colonization of the small and large intestine. *Can J Microbiol* 54, 984–995. doi: 10.1139/w08-090.

Bliska, J. B., Wang, X., Viboud, G. I., and Brodsky, I. E. (2013). Modulation of innate immune responses by Yersinia type III secretion system translocators and effectors. *Cell Microbiol* 15, 1622–1631. doi: 10.1111/cmi.12164.

Caballero-Flores, G., Pickard, J. M., and Núñez, G. (2021). Regulation of Citrobacter rodentium colonization: virulence, immune response and microbiota interactions. *Curr. Opin. Microbiol.* 63, 142–149. doi: 10.1016/j.mib.2021.07.003.

Cote, R., Katani, R., Moreau, M. R., Kudva, I. T., Arthur, T. M., DebRoy, C., et al. (2015). Comparative Analysis of Super-Shedder Strains of Escherichia coli O157:H7 Reveals Distinctive Genomic Features and a Strongly Aggregative Adherent Phenotype on Bovine Rectoanal Junction Squamous Epithelial Cells. *Plos One* 10, e0116743. doi: 10.1371/journal.pone.0116743.

Durso, L. M., Harhay, G. P., Smith, T. P. L., Bono, J. L., DeSantis, T. Z., Harhay, D. M., et al. (2010). Animal-to-Animal Variation in Fecal Microbial Diversity among Beef Cattle ▿ . *Appl Environ Microb* 76, 4858–4862. doi: 10.1128/aem.00207-10.

Fitzgerald, S. F., Beckett, A. E., Palarea-Albaladejo, J., McAteer, S., Shaaban, S., Morgan, J., et al. (2019). Shiga toxin sub-type 2a increases the efficiency of Escherichia coli O157 transmission between animals and restricts epithelial regeneration in bovine enteroids. *Plos Pathog* 15, e1008003. doi: 10.1371/journal.ppat.1008003.

Fraser, M. E., Fujinaga, M., Cherney, M. M., Melton-Celsa, A. R., Twiddy, E. M., O'Brien, A. D., et al. (2004). Structure of Shiga Toxin Type 2 (Stx2) from Escherichia coli O157:H7*. *J Biol Chem* 279, 27511–27517. doi: 10.1074/jbc.m401939200.

Gaytán, M. O., Martínez-Santos, V. I., Soto, E., and González-Pedrajo, B. (2016). Type Three Secretion System in Attaching and Effacing Pathogens. *Front Cell Infect Mi* 6, 129. doi: 10.3389/fcimb.2016.00129.

Lisboa, Szelewicki, Lin, Latonas, Li, Zhi, et al. (2019). Epidemiology of Shiga Toxin-Producing Escherichia coli O157 in the Province of Alberta, Canada, 2009–2016. *Toxins* 11, 613. doi: 10.3390/toxins11100613.

McCabe, E., Burgess, C. M., Lawal, D., Whyte, P., and Duffy, G. (2019). An investigation of shedding and super-shedding of Shiga toxigenic Escherichia coli O157 and E. coli O26 in cattle

presented for slaughter in the Republic of Ireland. *Zoonoses Public Hlth* 66, 83–91. doi: 10.1111/zph.12531.

Mills, K. H. G. (2023). IL-17 and IL-17-producing cells in protection versus pathology. *Nat. Rev. Immunol.* 23, 38–54. doi: 10.1038/s41577-022-00746-9.

Miyazawa, K., Hondo, T., Kanaya, T., Tanaka, S., Takakura, I., Itani, W., et al. (2010). Characterization of newly established bovine intestinal epithelial cell line. *Histochem. Cell Biol.* 133, 125–134. doi: 10.1007/s00418-009-0648-3.

Munns, K. D., Selinger, L. B., Stanford, K., Guan, L., Callaway, T. R., and McAllister, T. A. (2015). Perspectives on Super-Shedding of Escherichia coli O157:H7 by Cattle. *Foodborne Pathog Dis* 12, 89–103. doi: 10.1089/fpd.2014.1829.

Nicholson, J. K., Holmes, E., Lindon, J. C., and Wilson, I. D. (2004). The challenges of modeling mammalian biocomplexity. *Nat. Biotechnol.* 22, 1268–1274. doi: 10.1038/nbt1015.

Paton, A. W., and Paton, J. C. (1996). Enterobacter cloacae producing a Shiga-like toxin II-related cytotoxin associated with a case of hemolytic-uremic syndrome. *J. Clin. Microbiol.* 34, 463–465. doi: 10.1128/jcm.34.2.463-465.1996.

Roxas, J. L., Koutsouris, A., Bellmeyer, A., Tesfay, S., Royan, S., Falzari, K., et al. (2010). Enterohemorrhagic E. coli alters murine intestinal epithelial tight junction protein expression and barrier function in a Shiga toxin independent manner. *Lab. Investig.* 90, 1152–1168. doi: 10.1038/labinvest.2010.91.

Schmidt, H., Scheef, J., Morabito, S., Caprioli, A., Wieler, L. H., and Karch, H. (2000). A New Shiga Toxin 2 Variant (Stx2f) fromEscherichia coli Isolated from Pigeons. *Appl Environ Microb* 66, 1205–1208. doi: 10.1128/aem.66.3.1205-1208.2000.

Sheng, H., Lim, J. Y., Knecht, H. J., Li, J., and Hovde, C. J. (2006). Role of Escherichia coli O157:H7 Virulence Factors in Colonization at the Bovine Terminal Rectal Mucosa. *Infect Immun* 74, 4685–4693. doi: 10.1128/iai.00406-06.

Sheng, H., Wang, J., Lim, J. Y., Davitt, C., Minnich, S. A., and Hovde, C. J. (2011). Internalization of Escherichia Coli O157:H7 by Bovine Rectal Epithelial Cells. *Front Microbiol* 2, 32. doi: 10.3389/fmicb.2011.00032.

Vlisidou, I., Lyte, M., Diemen, P. M. van, Hawes, P., Monaghan, P., Wallis, T. S., et al. (2004). The Neuroendocrine Stress Hormone Norepinephrine Augments Escherichia coli O157:H7-Induced Enteritis and Adherence in a Bovine Ligated Ileal Loop Model of Infection. *Infect. Immun.* 72, 5446–5451. doi: 10.1128/iai.72.9.5446-5451.2004.

Wang, J. (2012). Geometric Structure of High-Dimensional Data and Dimensionality Reduction. 151–180. doi: 10.1007/978-3-642-27497-8_8.

Wang, O., Liang, G., McAllister, T. A., Plastow, G., Stanford, K., Selinger, B., et al. (2016). Comparative Transcriptomic Analysis of Rectal Tissue from Beef Steers Revealed Reduced Host Immunity in Escherichia coli O157:H7 Super-Shedders. *Plos One* 11, e0151284. doi: 10.1371/journal.pone.0151284.

Wang, O., McAllister, T. A., Plastow, G., Stanford, K., Selinger, B., and Guan, L. L. (2017). Host mechanisms involved in cattle Escherichia coli O157 shedding: a fundamental understanding for reducing foodborne pathogen in food animal production. *Sci Rep-uk* 7, 7630. doi: 10.1038/s41598-017-06737-4.

Wang, O., Zhou, M., Chen, Y., McAllister, T. A., Plastow, G., Stanford, K., et al. (2021). MicroRNAomes of Cattle Intestinal Tissues Revealed Possible miRNA Regulated Mechanisms Involved in Escherichia coli O157 Fecal Shedding. *Front Cell Infect Mi* 11, 634505. doi: 10.3389/fcimb.2021.634505.

Xu, Y., Dugat-Bony, E., Zaheer, R., Selinger, L., Barbieri, R., Munns, K., et al. (2014). Escherichia coli O157:H7 Super-Shedder and Non-Shedder Feedlot Steers Harbour Distinct Fecal Bacterial Communities. *Plos One* 9, e98115. doi: 10.1371/journal.pone.0098115.

Zaheer, R., Dugat-Bony, E., Holman, D., Cousteix, E., Xu, Y., Munns, K., et al. (2017). Changes in bacterial community composition of Escherichia coli O157:H7 super-shedder cattle occur in the lower intestine. *Plos One* 12, e0170050. doi: 10.1371/journal.pone.0170050.

Bibliography

Aitchison, J. (1981). A new approach to null correlations of proportions. *J Int Ass Math Geol* 13, 175–189. doi: 10.1007/bf01031393.

Aitchison, J., and Egozcue, J. J. (2005). Compositional Data Analysis: Where Are We and Where Should We Be Heading? *Math Geol* 37, 829–850. doi: 10.1007/s11004-005-7383-7.

Akbar, S., Gu, L., Sun, Y., Zhang, L., Lyu, K., Huang, Y., et al. (2022). Understanding host-microbiome-environment interactions: Insights from Daphnia as a model organism. *Sci. Total Environ.* 808, 152093. doi: 10.1016/j.scitotenv.2021.152093.

Albert, R., and Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Rev Mod Phys* 74, 47–97. doi: 10.1103/revmodphys.74.47.

Alharbi, M. G., Al-Hindi, R. R., Esmael, A., Alotibi, I. A., Azhari, S. A., Alseghayer, M. S., et al. (2022). The "Big Six": Hidden Emerging Foodborne Bacterial Pathogens. *Tropical Medicine Infect Dis* 7, 356. doi: 10.3390/tropicalmed7110356.

Alonso, C. A., Mora, A., Díaz, D., Blanco, M., González-Barrio, D., Ruiz-Fons, F., et al. (2017). Occurrence and characterization of stx and/or eae-positive Escherichia coli isolated from wildlife, including a typical EPEC strain from a wild boar. *Vet. Microbiol.* 207, 69–73. doi: 10.1016/j.vetmic.2017.05.028.

Anderson, M. J., Crist, T. O., Chase, J. M., Vellend, M., Inouye, B. D., Freestone, A. L., et al. (2011). Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. *Ecol Lett* 14, 19–28. doi: 10.1111/j.1461-0248.2010.01552.x.

Andreoli, S. P., Trachtman, H., Acheson, D. W. K., Siegler, R. L., and Obrig, T. G. (2002). Hemolytic uremic syndrome: epidemiology, pathophysiology, and therapy. *Pediatr Nephrol* 17, 293–298. doi: 10.1007/s00467-001-0783-0.

Araki, R., Fukumura, R., Sasaki, N., Kasama, Y., Suzuki, N., Takahashi, H., et al. (2006). More Than 40,000 Transcripts, Including Novel and Noncoding Transcripts, in Mouse Embryonic Stem Cells. *STEM CELLS* 24, 2522–2528. doi: 10.1634/stemcells.2006-0005.

Arumugam, M., Raes, J., Pelletier, E., Paslier, D. L., Yamada, T., Mende, D. R., et al. (2011). Enterotypes of the human gut microbiome. *Nature* 473, 174–180. doi: 10.1038/nature09944.

Astel, A., Tsakovski, S., Barbieri, P., and Simeonov, V. (2007). Comparison of self-organizing maps classification approach with cluster and principal components analysis for large environmental data sets. *Water Res* 41, 4566–4578. doi: 10.1016/j.watres.2007.06.030.

Baines, D., Lee, B., and McAllister, T. (2008). Heterogeneity in enterohemorrhagic Escherichia coli O157:H7 fecal shedding in cattle is related to Escherichia coli O157:H7 colonization of the small and large intestine. *Can J Microbiol* 54, 984–995. doi: 10.1139/w08-090.

Barberán, A., Bates, S. T., Casamayor, E. O., and Fierer, N. (2012). Using network analysis to explore co-occurrence patterns in soil microbial communities. *Isme J* 6, 343–351. doi: 10.1038/ismej.2011.119.

Bartolomaeus, H., McParland, V., and Wilck, N. (2020). Darm-Herz-Achse. *Herz* 45, 134–141. doi: 10.1007/s00059-020-04897-0.

Berge, A. C., Hancock, D. D., Sischo, W. M., and Besser, T. E. (2010). Geographic, farm, and animal factors associated with multiple antimicrobial resistance in fecal Escherichia coli isolates from cattle in the western United States. *J Am Vet Med Assoc* 236, 1338–1344. doi: 10.2460/javma.236.12.1338.

Bertin, Y., Girardeau, J. P., Chaucheyras-Durand, F., Lyan, B., Pujos-Guillot, E., Harel, J., et al. (2011). Enterohaemorrhagic Escherichia coli gains a competitive advantage by using ethanolamine as a nitrogen source in the bovine intestinal content. *Environ Microbiol* 13, 365–377. doi: 10.1111/j.1462-2920.2010.02334.x.

Bizama, C., Benavente, F., Salvatierra, E., Gutiérrez-Moraga, A., Espinoza, J. A., Fernández, E. A., et al. (2014). The low-abundance transcriptome reveals novel biomarkers, specific intracellular pathways and targetable genes associated with advanced gastric cancer. *Int. J. Cancer* 134, 755–764. doi: 10.1002/ijc.28405.

Bliska, J. B., Wang, X., Viboud, G. I., and Brodsky, I. E. (2013). Modulation of innate immune responses by Yersinia type III secretion system translocators and effectors. *Cell Microbiol* 15, 1622–1631. doi: 10.1111/cmi.12164.

Blount, Z. D. (2015). The unexhausted potential of E. coli. *Elife* 4, e05826. doi: 10.7554/elife.05826.

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., et al. (2019). Author Correction: Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* 37, 1091–1091. doi: 10.1038/s41587-019-0252-6.

Bono, J. L., Keen, J. E., Miller, L. C., Fox, J. M., Chitko-McKown, C. G., Heaton, M. P., et al. (2004). Evaluation of a Real-Time PCR Kit for Detecting Escherichia coli O157 in Bovine Fecal Samples. *Appl Environ Microb* 70, 1855–1857. doi: 10.1128/aem.70.3.1855-1857.2004.

Bosilevac, J. M., and Koohmaraie, M. (2012). Predicting the Presence of Non-O157 Shiga Toxin-Producing Escherichia coli in Ground Beef by Using Molecular Tests for Shiga Toxins, Intimin, and O Serogroups. *Appl. Environ. Microbiol.* 78, 7152–7155. doi: 10.1128/aem.01508-12.

Bouskra, D., Brézillon, C., Bérard, M., Werts, C., Varona, R., Boneca, I. G., et al. (2008). Lymphoid tissue genesis induced by commensals through NOD1 regulates intestinal homeostasis. *Nature* 456, 507–510. doi: 10.1038/nature07450.

Bowman, E. P., Campbell, J. J., Soler, D., Dong, Z., Manlongat, N., Picarella, D., et al. (2000). Developmental Switches in Chemokine Response Profiles during B Cell Differentiation and Maturation. *J. Exp. Med.* 191, 1303–1318. doi: 10.1084/jem.191.8.1303.

Braga, R. M., Dourado, M. N., and Araújo, W. L. (2016). Microbial interactions: ecology in a molecular perspective. *Braz J Microbiol* 47, 86–98. doi: 10.1016/j.bjm.2016.10.005.

Brandelli, J. R., Griener, T. P., Laing, A., Mulvey, G., and Armstrong, G. D. (2015). The Effects of Shiga Toxin 1, 2 and Their Subunits on Cytokine and Chemokine Expression by Human Macrophage-Like THP-1 Cells. *Toxins* 7, 4054–4066. doi: 10.3390/toxins7104054.

Bretschneider, G., Berberov, E. M., and Moxley, R. A. (2007a). Isotype-specific antibody responses against Escherichia coli O157:H7 locus of enterocyte effacement proteins in adult beef cattle following experimental infection. *Vet Immunol Immunop* 118, 229–238. doi: 10.1016/j.vetimm.2007.06.005.

Bretschneider, G., Berberov, E. M., and Moxley, R. A. (2007b). Reduced intestinal colonization of adult beef cattle by Escherichia coli O157:H7 tir deletion and nalidixic-acid-resistant mutants lacking flagellar expression. *Vet Microbiol* 125, 381–386. doi: 10.1016/j.vetmic.2007.06.009.

Browne, A. S., Midwinter, A. C., Withers, H., Cookson, A. L., Biggs, P. J., Marshall, J. C., et al. (2021a). Transmission Dynamics of Shiga Toxin-Producing Escherichia coli in New Zealand Cattle from Farm to Slaughter. *Appl Environ Microb* 87, e02907-20. doi: 10.1128/aem.02907-20.

Browne, H. P., Almeida, A., Kumar, N., Vervier, K., Adoum, A. T., Viciani, E., et al. (2021b). Host adaptation in gut Firmicutes is associated with sporulation loss and altered transmission cycle. *Genome Biol* 22, 204. doi: 10.1186/s13059-021-02428-6.

Browning, J. L., Dougas, I., Ngam-ek, A., Bourdon, P. R., Ehrenfels, B. N., Miatkowski, K., et al. (1995). Characterization of surface lymphotoxin forms. Use of specific monoclonal antibodies and soluble receptors. *J. Immunol. (Baltim., Md : 1950)* 154, 33–46.

Browning, K. N., and Travagli, R. A. (2021). Comprehensive Physiology. *Compr. Physiol.* 4, 1339–1368. doi: 10.1002/cphy.c130055.

Buffie, C. G., and Pamer, E. G. (2013). Microbiota-mediated colonization resistance against intestinal pathogens. *Nat Rev Immunol* 13, 790–801. doi: 10.1038/nri3535.

Bulcke, M. V. den, Lievens, A., Barbau-Piednoir, E., MbongoloMbella, G., Roosens, N., Sneyers, M., et al. (2010). A theoretical introduction to “Combinatory SYBR®Green qPCR Screening”, a matrix-based approach for the detection of materials derived from genetically modified plants. *Anal Bioanal Chem* 396, 2113–2123. doi: 10.1007/s00216-009-3286-7.

Bürk, C., Dietrich, R., Açar, G., Moravek, M., Bülte, M., and Märtlbauer, E. (2003). Identification and Characterization of a New Variant of Shiga Toxin 1 in Escherichia coli ONT:H19 of Bovine Origin. *J Clin Microbiol* 41, 2106–2112. doi: 10.1128/jcm.41.5.2106-2112.2003.

Bustin, S. A. (2005). Real-time, fluorescence-based quantitative PCR: a snapshot of current procedures and preferences. *Expert Rev Mol Diagn* 5, 493–498. doi: 10.1586/14737159.5.4.493.

Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., et al. (2009). The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. *Clin Chem* 55, 611–622. doi: 10.1373/clinchem.2008.112797.

Caballero-Flores, G., Pickard, J. M., and Núñez, G. (2021). Regulation of Citrobacter rodentium colonization: virulence, immune response and microbiota interactions. *Curr. Opin. Microbiol.* 63, 142–149. doi: 10.1016/j.mib.2021.07.003.

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* 13, 581–583. doi: 10.1038/nmeth.3869.

Carscadden, K. A., Emery, N. C., Arnillas, C. A., Cadotte, M. W., Afkhami, M. E., Gravel, D., et al. (2020). Niche Breadth: Causes and Consequences for Ecology, Evolution, and Conservation. *Q Rev Biology* 95, 179–214. doi: 10.1086/710388.

Casey, M. E., Meade, K. G., Nalpas, N. C., Taraktsoglou, M., Browne, J. A., Killick, K. E., et al. (2015). Analysis of the Bovine Monocyte-Derived Macrophage Response to Mycobacterium avium Subspecies Paratuberculosis Infection Using RNA-seq. *Front Immunol* 6, 23. doi: 10.3389/fimmu.2015.00023.

Castro, V. S., Carvalho, R. C. T., Conte-Junior, C. A., and Figuiredo, E. E. S. (2017). Shiga-toxin Producing Escherichia coli: Pathogenicity, Supershedding, Diagnostic Methods, Occurrence, and Foodborne Outbreaks. *Compr Rev Food Sci F* 16, 1269–1280. doi: 10.1111/1541-4337.12302.

Chaffron, S., Rehrauer, H., Pernthaler, J., and Mering, C. von (2010). A global network of coexisting microbes from environmental and whole-genome sequence data. *Genome Res* 20, 947–959. doi: 10.1101/gr.104521.109.

Chapman, P., Siddons, C., Wright, D., et al. (1993). Cattle as a possible source of verocytotoxin-producing Escherichia coli O157 infections in man | Epidemiology & Infection | Cambridge Core. *Cattle as a possible source of verocytotoxin-producing Escherichia coli O157 infections in man.* Available at: https://www.cambridge.org/core/journals/epidemiology-and-infection/article/cattle-as-a-possible-source-of-verocytotoxinproducing-escherichia-coli-o157-infections-in-man/E350CA666395D1B61BEA2A68E9419BE1 [Accessed March 20, 2021].

Chase, J. M., Kraft, N. J. B., Smith, K. G., Vellend, M., and Inouye, B. D. (2011). Using null models to disentangle variation in community dissimilarity from variation in α-diversity. *Ecosphere* 2, 1–11. doi: 10.1890/es10-00117.1.

Chase-Topping, M., Gally, D., Low, C., Matthews, L., and Woolhouse, M. (2008). Super-shedding and the link between human infection and livestock carriage of Escherichia coli O157. *Nat Rev Microbiol* 6, 904–912. doi: 10.1038/nrmicro2029.

Chatelier, E. L., Nielsen, T., Qin, J., Prifti, E., Hildebrand, F., Falony, G., et al. (2013). Richness of human gut microbiome correlates with metabolic markers. *Nature* 500, 541–546. doi: 10.1038/nature12506.

Chave, J. (2004). Neutral theory and community ecology. *Ecol Lett* 7, 241–253. doi: 10.1111/j.1461-0248.2003.00566.x.

Chess, A. (2013). Random and Non-Random Monoallelic Expression. *Neuropsychopharmacology* 38, 55–61. doi: 10.1038/npp.2012.85.

Cho, S., Fossler, C. P., Diez-Gonzalez, F., Wells, S. J., Hedberg, C. W., Kaneene, J. B., et al. (2007). Cattle-level risk factors associated with fecal shedding of Shiga toxin-encoding bacteria on dairy farms, Minnesota, USA. *Can. J. Vet. Res. Rev. Can. Rech. veterinaire* 73, 151–6.

Chopyk, J., Moore, R. M., DiSpirito, Z., Stromberg, Z. R., Lewis, G. L., Renter, D. G., et al. (2016). Presence of pathogenic Escherichia coli is correlated with bacterial community diversity and composition on pre-harvest cattle hides. *Microbiome* 4, 9. doi: 10.1186/s40168-016-0155-4.

Cimolai, N., Basalyga, S., Mah, D. G., Morrison, B. J., and Carter, J. E. (1994). A continuing assessment of risk factors for the development of Escherichia coli O157:H7-associated hemolytic uremic syndrome. *Clin Nephrol* 42, 85–9.

Clark, R. L., Connors, B. M., Stevenson, D. M., Hromada, S. E., Hamilton, J. J., Amador-Noguez, D., et al. (2021). Design of synthetic human gut microbiome assembly and butyrate production. *Nat Commun* 12, 3254. doi: 10.1038/s41467-021-22938-y.

Cobbaut, K., Berkvens, D., Houf, K., Deken, R. D., and Zutter, L. D. (2009). Escherichia coli O157 Prevalence in Different Cattle Farm Types and Identification of Potential Risk Factors. *J Food Protect* 72, 1848–1853. doi: 10.4315/0362-028x-72.9.1848.

Cobbold, R., and Desmarchelier, P. (2000). A longitudinal study of Shiga-toxigenic Escherichia coli (STEC) prevalence in three Australian dairy herds. *Vet Microbiol* 71, 125–137. doi: 10.1016/s0378-1135(99)00173-x.

Corbishley, A., Ahmad, N. I., Hughes, K., Hutchings, M. R., McAteer, S. P., Connelley, T. K., et al. (2014). Strain-Dependent Cellular Immune Responses in Cattle following Escherichia coli O157:H7 Colonization. *Infect Immun* 82, 5117–5131. doi: 10.1128/iai.02462-14.

Cote, R., Katani, R., Moreau, M. R., Kudva, I. T., Arthur, T. M., DebRoy, C., et al. (2015). Comparative Analysis of Super-Shedder Strains of Escherichia coli O157:H7 Reveals Distinctive Genomic Features and a Strongly Aggregative Adherent Phenotype on Bovine Rectoanal Junction Squamous Epithelial Cells. *Plos One* 10, e0116743. doi: 10.1371/journal.pone.0116743.

Couronné, R., Probst, P., and Boulesteix, A.-L. (2018). Random forest versus logistic regression: a large-scale benchmark experiment. *Bmc Bioinformatics* 19, 270. doi: 10.1186/s12859-018-2264-5.

Croxen, M. A., Law, R. J., Scholz, R., Keeney, K. M., Wlodarska, M., and Finlay, B. B. (2013). Recent Advances in Understanding Enteric Pathogenic Escherichia coli. *Clin Microbiol Rev* 26, 822–880. doi: 10.1128/cmr.00022-13.

Csermely, P. (2009). Weak Links, The Universal Key to the Stability of Networks and Complex Systems. *Frontiers Collect*, 53–100. doi: 10.1007/978-3-540-31157-7_3.

Cua, D. J., and Tato, C. M. (2010). Innate IL-17-producing cells: the sentinels of the immune system. *Nat Rev Immunol* 10, 479–489. doi: 10.1038/nri2800.

Das, G., Chattopadhyay, M., and Gupta, S. (2015). A Comparison of Self-organising Maps and Principal Components Analysis. *Int J Market Res* 58, 815–834. doi: 10.2501/ijmr-2016-039.

David, C. C., and Jacobs, D. J. (2013). Protein Dynamics, Methods and Protocols. *Methods Mol Biology* 1084, 193–226. doi: 10.1007/978-1-62703-658-0_11.

Davies, D. L., and Bouldin, D. W. (1979). A cluster separation measure. *IEEE Trans. pattern Anal. Mach. Intell.* 1, 224–7.

Dawson, K., Rodriguez, R. L., and Malyj, W. (2005). Sample phenotype clusters in high-density oligonucleotide microarray data sets are revealed using Isomap, a nonlinear algorithm. *BMC Bioinform.* 6, 195. doi: 10.1186/1471-2105-6-195.

Dean-Nystrom, E. A., Bosworth, B. T., Cray, W. C., and Moon, H. W. (1997). Pathogenicity of Escherichia coli O157:H7 in the intestines of neonatal calves. *Infect Immun* 65, 1842–1848. doi: 10.1128/iai.65.5.1842-1848.1997.

DeGrandis, S., Law, H., Brunton, J., Gyles, C., and Lingwood, C. A. (1989). Globotetraosylceramide Is Recognized by the Pig Edema Disease Toxin. *J Biol Chem* 264, 12520–12525. doi: 10.1016/s0021-9258(18)63888-8.

Delgado, S., Morán, F., Mora, A., Merelo, J. J., and Briones, C. (2015). A novel representation of genomic sequences for taxonomic clustering and visualization by means of self-organizing maps. *Bioinformatics* 31, 736–744. doi: 10.1093/bioinformatics/btu708.

Deng, W., Marshall, N. C., Rowland, J. L., McCoy, J. M., Worrall, L. J., Santos, A. S., et al. (2017). Assembly, structure, function and regulation of type III secretion systems. *Nat Rev Microbiol* 15, 323–337. doi: 10.1038/nrmicro.2017.20.

Deng, Y., Jiang, Y.-H., Yang, Y., He, Z., Luo, F., and Zhou, J. (2012). Molecular ecological network analyses. *Bmc Bioinformatics* 13, 113. doi: 10.1186/1471-2105-13-113.

Derakhshani, H., Plaizier, J. C., Buck, J. D., Barkema, H. W., and Khafipour, E. (2020). Composition and co-occurrence patterns of the microbiota of different niches of the bovine

mammary gland: potential associations with mastitis susceptibility, udder inflammation, and teat-end hyperkeratosis. *Animal Microbiome* 2, 11. doi: 10.1186/s42523-020-00028-6.

Deurenberg, R. H., Bathoorn, E., Chlebowicz, M. A., Couto, N., Ferdous, M., García-Cobos, S., et al. (2017). Application of next generation sequencing in clinical microbiology and infection prevention. *J Biotechnol* 243, 16–24. doi: 10.1016/j.jbiotec.2016.12.022.

Donkersgoed, J. V., Berg, J., Potter, A., Hancock, D., Besser, T., Rice, D., et al. (2001). Environmental sources and transmission of Escherichia coli O157 in feedlot cattle. *Can Vet J La Revue Vétérinaire Can* 42, 714–20.

Dowd, S. E., Sun, Y., Secor, P. R., Rhoads, D. D., Wolcott, B. M., James, G. A., et al. (2008). Survey of bacterial diversity in chronic wounds using Pyrosequencing, DGGE, and full ribosome shotgun sequencing. *Bmc Microbiol* 8, 43. doi: 10.1186/1471-2180-8-43.

Driessche, L. V., Vanneste, K., Bogaerts, B., Keersmaecker, S. C. J. D., Roosens, N. H., Haesebrouck, F., et al. (2020). Isolation of Drug-Resistant Gallibacterium anatis from Calves with Unresponsive Bronchopneumonia - Volume 26, Number 4—April 2020 - Emerging Infectious Diseases journal - CDC. *Emerg Infect Dis* 26, 721–730. doi: 10.3201/eid2604.190962.

Dundas, S., Todd, W. T. A., Stewart, A. I., Murdoch, P. S., Chaudhuri, A. K. R., and Hutchinson, S. J. (2001). The Central Scotland Escherichia coli O157:H7 Outbreak: Risk Factors for the

Hemolytic Uremic Syndrome and Death among Hospitalized Patients. *Clin Infect Dis* 33, 923–931. doi: 10.1086/322598.

Durso, L. M., Harhay, G. P., Smith, T. P. L., Bono, J. L., DeSantis, T. Z., Harhay, D. M., et al. (2010). Animal-to-Animal Variation in Fecal Microbial Diversity among Beef Cattle ▿ . *Appl Environ Microb* 76, 4858–4862. doi: 10.1128/aem.00207-10.

Easton, D. M., Totsika, M., Allsopp, L. P., Phan, M.-D., Idris, A., Wurpel, D. J., et al. (2011). Characterization of EhaJ, a New Autotransporter Protein from Enterohemorrhagic and Enteropathogenic Escherichia coli. *Front. Microbiol.* 2, 120. doi: 10.3389/fmicb.2011.00120.

Eckburg, P. B., Bik, E. M., Bernstein, C. N., Purdom, E., Dethlefsen, L., Sargent, M., et al. (2005). Diversity of the Human Intestinal Microbial Flora. *Science* 308, 1635–1638. doi: 10.1126/science.1110591.

EKIRI, A. B., LANDBLOM, D., DOETKOTT, D., OLET, S., SHELVER, W. L., and KHAITSA, M. L. (2016). Isolation and Characterization of Shiga Toxin–Producing Escherichia coli Serogroups O26, O45, O103, O111, O113, O121, O145, and O157 Shed from Range and Feedlot Cattle from Postweaning to Slaughter. *J Food Protect* 77, 1052–1061. doi: 10.4315/0362-028x.jfp-13-373.

Etcheverría, A. I., and Padola, N. L. (2013). Shiga toxin-producing Escherichia coli. *Virulence* 4, 366–372. doi: 10.4161/viru.24642.

Fagarasan, S., Muramatsu, M., Suzuki, K., Nagaoka, H., Hiai, H., and Honjo, T. (2002). Critical Roles of Activation-Induced Cytidine Deaminase in the Homeostasis of Gut Flora. *Science* 298, 1424–1427. doi: 10.1126/science.1077336.

Fairbrother, J. M., and Nadeau, E. (2006). Escherichia coli: on-farm contamination of animals. *Revue Sci Et Technique Int Office Epizootics* 25, 555–69.

Fan, P., Nelson, C. D., Driver, J. D., Elzo, M. A., Peñagaricano, F., and Jeong, K. C. (2021). Host genetics exerts lifelong effects upon hindgut microbiota and its association with bovine growth and immunity. *Isme J* 15, 2306–2321. doi: 10.1038/s41396-021-00925-x.

Farfan, M. J., and Torres, A. G. (2012). Molecular Mechanisms That Mediate Colonization of Shiga Toxin-Producing Escherichia coli Strains. *Infect. Immun.* 80, 903–913. doi: 10.1128/iai.05907-11.

Faust, K., and Raes, J. (2012). Microbial interactions: from networks to models. *Nat Rev Microbiol* 10, 538–550. doi: 10.1038/nrmicro2832.

Faust, K., Sathirapongsasuti, J. F., Izard, J., Segata, N., Gevers, D., Raes, J., et al. (2012). Microbial Co-occurrence Relationships in the Human Microbiome. *Plos Comput Biol* 8, e1002606. doi: 10.1371/journal.pcbi.1002606.

Ferreira-Halder, C. V., Faria, A. V. de S., and Andrade, S. S. (2017). Action and function of Faecalibacterium prausnitzii in health and disease. *Best Pract Res Clin Gastroenterology* 31, 643–648. doi: 10.1016/j.bpg.2017.09.011.

Fitzgerald, S. F., Beckett, A. E., Palarea-Albaladejo, J., McAteer, S., Shaaban, S., Morgan, J., et al. (2019). Shiga toxin sub-type 2a increases the efficiency of Escherichia coli O157 transmission between animals and restricts epithelial regeneration in bovine enteroids. *Plos Pathog* 15, e1008003. doi: 10.1371/journal.ppat.1008003.

Frankel, G., Phillips, A. D., Rosenshine, I., Dougan, G., Kaper, J. B., and Knutton, S. (1998). Enteropathogenic and enterohaemorrhagic Escherichia coli : more subversive elements. *Mol Microbiol* 30, 911–921. doi: 10.1046/j.1365-2958.1998.01144.x.

Fraser, M. E., Fujinaga, M., Cherney, M. M., Melton-Celsa, A. R., Twiddy, E. M., O'Brien, A. D., et al. (2004). Structure of Shiga Toxin Type 2 (Stx2) from Escherichia coli O157:H7*. *J Biol Chem* 279, 27511–27517. doi: 10.1074/jbc.m401939200.

Fuller, C. A., Pellino, C. A., Flagler, M. J., Strasser, J. E., and Weiss, A. A. (2011). Shiga Toxin Subtypes Display Dramatic Differences in Potency. *Infect Immun* 79, 1329–1337. doi: 10.1128/iai.01182-10.

Furman, O., Shenhav, L., Sasson, G., Kokou, F., Honig, H., Jacoby, S., et al. (2020). Stochasticity constrained by deterministic effects of diet and age drive rumen microbiome assembly dynamics. *Nat Commun* 11, 1904. doi: 10.1038/s41467-020-15652-8.

Galanis, E., Longmore, K., Hasselback, P., Swann, D., Ellis, A., and Panaro, L. (2003). Investigation of an E. coli O157:H7 outbreak in Brooks, Alberta, June-July 2002: the role of

occult cases in the spread of infection within a daycare setting. *Can Commun Dis Rep Relevé Des Maladies Transm Au Can* 29, 21–8.

Gaytán, M. O., Martínez-Santos, V. I., Soto, E., and González-Pedrajo, B. (2016). Type Three Secretion System in Attaching and Effacing Pathogens. *Front Cell Infect Mi* 6, 129. doi: 10.3389/fcimb.2016.00129.

Glass, E. J., and Jensen, K. (2007). Resistance and susceptibility to a protozoan parasite of cattle—Gene expression differences in macrophages from different breeds of cattle. *Vet. Immunol. Immunopathol.* 120, 20–30. doi: 10.1016/j.vetimm.2007.07.013.

Glielmo, A., Husic, B. E., Rodriguez, A., Clementi, C., Noé, F., and Laio, A. (2021). Unsupervised Learning Methods for Molecular Simulation Data. *Chem Rev* 121, 9722–9758. doi: 10.1021/acs.chemrev.0c01195.

Gomes, T. A. T., Elias, W. P., Scaletsky, I. C. A., Guth, B. E. C., Rodrigues, J. F., Piazza, R. M. F., et al. (2016). Diarrheagenic Escherichia coli. *Braz J Microbiol* 47, 3–30. doi: 10.1016/j.bjm.2016.10.015.

Griffin, P. M., and Tauxe, R. V. (1991). 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. doi: 10.1093/oxfordjournals.epirev.a036079.

Group, O. W. (2018). Monitoring the incidence and causes of diseases potentially transmitted by food in Australia: Annual report of the OzFoodNet network, 2012. *Commun Dis Intell 2018* 42.

Guimerà, R., and Amaral, L. A. N. (2005). Functional cartography of complex metabolic networks. *Nature* 433, 895–900. doi: 10.1038/nature03288.

Guo, W., Zhou, M., Ma, T., Bi, S., Wang, W., Zhang, Y., et al. (2020). Survey of rumen microbiota of domestic grazing yak during different growth stages revealed novel maturation patterns of four key microbial groups and their dynamic interactions. *Animal Microbiome* 2, 23. doi: 10.1186/s42523-020-00042-8.

Gupta, S., Mortensen, M. S., Schjørring, S., Trivedi, U., Vestergaard, G., Stokholm, J., et al. (2019). Amplicon sequencing provides more accurate microbiome information in healthy children compared to culturing. *Commun Biology* 2, 291. doi: 10.1038/s42003-019-0540-1.

Hale, T. L., and Formal, S. B. (1980). Cytotoxicity of Shigella dysenteriae 1 for cultured mammalian cells. *Am J Clin Nutrition* 33, 2485–2490. doi: 10.1093/ajcn/33.11.2485.

Hall, G., Kurosawa, S., and Stearns-Kurosawa, D. J. (2017). Shiga Toxin Therapeutics: Beyond Neutralization. *Toxins* 9, 291. doi: 10.3390/toxins9090291.

Halliday, J. E., Chase-Topping, M. E., Pearce, M. C., McKendrick, I. J., Allison, L., Fenlon, D., et al. (2006). Herd-level risk factors associated with the presence of Phage type 21/28 E. coli O157 on Scottish cattle farms. *Bmc Microbiol* 6, 99. doi: 10.1186/1471-2180-6-99.

Han, H., Guo, X., and Yu, H. (2016). Variable Selection Using Mean Decrease Accuracy and Mean Decrease Gini Based on Random Forest. *2016 7th IEEE Int. Conf. Softw. Eng. Serv. Sci. (ICSESS)*, 219–224. doi: 10.1109/icsess.2016.7883053.

Han, K., Lim, H. C., and Hong, J. (1992). Acetic acid formation in escherichia coli fermentation. *Biotechnol Bioeng* 39, 663–671. doi: 10.1002/bit.260390611.

Han, Z., Ma, J., Yang, C.-H., and Ibekwe, A. M. (2021). Soil salinity, pH, and indigenous bacterial community interactively influence the survival of E. coli O157:H7 revealed by multivariate statistics. *Environ Sci Pollut R* 28, 5575–5586. doi: 10.1007/s11356-020-10942-6.

Hanson, C. A., Fuhrman, J. A., Horner-Devine, M. C., and Martiny, J. B. H. (2012). Beyond biogeographic patterns: processes shaping the microbial landscape. *Nat Rev Microbiol* 10, 497–506. doi: 10.1038/nrmicro2795.

Harrell, F. E., Lee, K. L., Matchar, D. B., and Reichert, T. A. (1985). Regression models for prognostic prediction: advantages, problems, and suggested solutions. *Cancer Treat Rep* 69, 1071–77.

Hartl, D. L., and Dykhuizen, D. E. (1984). The Population Genetics of Escherichia Coli. *Annu Rev Genet* 18, 31–68. doi: 10.1146/annurev.ge.18.120184.000335.

He, Z., Fischer, A., Song, Y., Steele, M., and Guan, L. L. (2018). Genome wide transcriptome analysis provides bases on colonic mucosal immune system development affected by colostrum feeding strategies in neonatal calves. *BMC Genom.* 19, 635. doi: 10.1186/s12864-018-5017-y.

Heiman, K. E., Mody, R. K., Johnson, S. D., Griffin, P. M., and Gould, L. H. (2015). Escherichia coli O157 Outbreaks in the United States, 2003–2012. *Emerg Infect Dis* 21, 1293–1301. doi: 10.3201/eid2108.141364.

Herbert, L. J., Vali, L., Hoyle, D. V., Innocent, G., McKendrick, I. J., Pearce, M. C., et al. (2014). E. coli O157 on Scottish cattle farms: Evidence of local spread and persistence using repeat cross-sectional data. *Bmc Vet Res* 10, 95–95. doi: 10.1186/1746-6148-10-95.

Herren, C. M., and McMahon, K. D. (2018). Keystone taxa predict compositional change in microbial communities. *Environ Microbiol* 20, 2207–2217. doi: 10.1111/1462-2920.14257.

Hicklin, A. (2004). Network Stability: Opportunity or Obstacles? *Public Organization Rev* 4, 121–133. doi: 10.1023/b:porj.0000031625.78226.bc.

Hirano, H., and Takemoto, K. (2019). Difficulty in inferring microbial community structure based on co-occurrence network approaches. *Bmc Bioinformatics* 20, 329. doi: 10.1186/s12859-019-2915-1.

Hoffman, M. A., Menge, C., Casey, T. A., Laegreid, W., Bosworth, B. T., and Dean-Nystrom, E. A. (2006). Bovine Immune Response to Shiga-Toxigenic Escherichia coli O157:H7. *Clin Vaccine Immunol* 13, 1322–1327. doi: 10.1128/cvi.00205-06.

Høiby, N., Bjarnsholt, T., Givskov, M., Molin, S., and Ciofu, O. (2010). Antibiotic resistance of bacterial biofilms. *Int J Antimicrob Ag* 35, 322–332. doi: 10.1016/j.ijantimicag.2009.12.011.

Horner-Devine, M. C., Silver, J. M., Leibold, M. A., Bohannan, B. J. M., Colwell, R. K., Fuhrman, J. A., et al. (2007). A COMPARISON OF TAXON CO-OCCURRENCE PATTERNS FOR MACRO- AND MICROORGANISMS. *Ecology* 88, 1345–1353. doi: 10.1890/06-0286.

Hsu, B. B., Way, J. C., and Silver, P. A. (2019). Stable neutralization of virulent bacteria using temperate phage in the mammalian gut. *Biorxiv*, 794222. doi: 10.1101/794222.

Huang, C., Ge, F., Yao, X., Guo, X., Bao, P., Ma, X., et al. (2021). Microbiome and Metabolomics Reveal the Effects of Different Feeding Systems on the Growth and Ruminal Development of Yaks. *Front Microbiol* 12, 682989. doi: 10.3389/fmicb.2021.682989.

Hubbell, S. P., and Borda-de-Água, L. (2004). THE UNIFIED NEUTRAL THEORY OF BIODIVERSITY AND BIOGEOGRAPHY: REPLY. *Ecology* 85, 3175–3178. doi: 10.1890/04-0808.

Ilhan, Z. E., Łaniewski, P., Tonachio, A., and Herbst-Kralovetz, M. M. (2020). Members of Prevotella Genus Distinctively Modulate Innate Immune and Barrier Functions in a Human Three-Dimensional Endometrial Epithelial Cell Model. *J Infect Dis* 222, 2082–2092. doi: 10.1093/infdis/jiaa324.

Iljazovic, A., Roy, U., Gálvez, E. J. C., Lesker, T. R., Zhao, B., Gronow, A., et al. (2021). Perturbation of the gut microbiome by Prevotella spp. enhances host susceptibility to mucosal inflammation. *Mucosal Immunol* 14, 113–124. doi: 10.1038/s41385-020-0296-4.

Imamovic, L., and Muniesa, M. (2011). Quantification and Evaluation of Infectivity of Shiga Toxin-Encoding Bacteriophages in Beef and Salad. *Appl. Environ. Microbiol.* 77, 3536–3540. doi: 10.1128/aem.02703-10.

Ising, C., Venegas, C., Zhang, S., Scheiblich, H., Schmidt, S. V., Vieira-Saecker, A., et al. (2019). NLRP3 inflammasome activation drives tau pathology. *Nature* 575, 669–673. doi: 10.1038/s41586-019-1769-z.

Jackson, M. P., Neill, R. J., O'Brien, A. D., Holmes, R. K., and Newland, J. W. (1987). Nucleotide sequence analysis and comparison of the structural genes for Shiga-like toxin I and Shiga-like toxin II encoded by bacteriophages from Escherichia coli 933. *Fems Microbiol Lett* 44, 109–114. doi: 10.1111/j.1574-6968.1987.tb02252.x.

Jansma, J., and Aidy, S. E. (2021). Understanding the host-microbe interactions using metabolic modeling. *Microbiome* 9, 16. doi: 10.1186/s40168-020-00955-1.

Ji, H., and Dong, H. (2015). Key steps in type III secretion system (T3SS) towards translocon assembly with potential sensor at plant plasma membrane. *Mol Plant Pathol* 16, 762–773. doi: 10.1111/mpp.12223.

Jiao, S., Liu, Z., Lin, Y., Yang, J., Chen, W., and Wei, G. (2016). Bacterial communities in oil contaminated soils: Biogeography and co-occurrence patterns. *Soil Biology Biochem* 98, 64–73. doi: 10.1016/j.soilbio.2016.04.005.

Johnson, D. R., Goldschmidt, F., Lilja, E. E., and Ackermann, M. (2012). Metabolic specialization and the assembly of microbial communities. *Isme J* 6, 1985–1991. doi: 10.1038/ismej.2012.46.

Jolliffe, I. T., and Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical Transactions Royal Soc Math Phys Eng Sci* 374, 20150202. doi: 10.1098/rsta.2015.0202.

Jun, W., Barahona, M., Yue-Jin, T., and Hong-Zhong, D. (2010). Natural Connectivity of Complex Networks. *Chinese Phys Lett* 27, 078902. doi: 10.1088/0256-307x/27/7/078902.

Kaiko, G. E., and Stappenbeck, T. S. (2014). Host–microbe interactions shaping the gastrointestinal environment. *Trends Immunol* 35, 538–548. doi: 10.1016/j.it.2014.08.002.

Kamo, T., Akazawa, H., Suzuki, J., and Komuro, I. (2017). Novel Concept of a Heart-Gut Axis in the Pathophysiology of Heart Failure. *Korean Circ. J.* 47, 663–669. doi: 10.4070/kcj.2017.0028.

Kaper, J. B., Nataro, J. P., and Mobley, H. L. T. (2004). Pathogenic Escherichia coli. *Nat Rev Microbiol* 2, 123–140. doi: 10.1038/nrmicro818.

Karmali, MohamedA., Petric, M., Steele, BrianT., and Lim, C. (1983). SPORADIC CASES OF HAEMOLYTIC-URAEMIC SYNDROME ASSOCIATED WITH FAECAL CYTOTOXIN AND CYTOTOXIN-PRODUCING ESCHERICHIA COLI IN STOOLS. *Lancet* 321, 619–620. doi: 10.1016/s0140-6736(83)91795-6.

Karnell, J. L., Dimasi, N., Karnell, F. G., Fleming, R., Kuta, E., Wilson, M., et al. (2014). CD19 and CD32b Differentially Regulate Human B Cell Responsiveness. *J. Immunol.* 192, 1480–1490. doi: 10.4049/jimmunol.1301361.

Karp, P. D., Midford, P. E., Caspi, R., and Khodursky, A. (2021). Pathway size matters: the influence of pathway granularity on over-representation (enrichment analysis) statistics. *BMC Genom.* 22, 191. doi: 10.1186/s12864-021-07502-8.

Karson, M. (1968). Handbook of Methods of Applied Statistics. Volume I: Techniques of Computation Descriptive Methods, and Statistical Inference. Volume II: Planning of Surveys and Experiments. I. M. Chakravarti, R. G. Laha, and J. Roy, New York, John Wiley; 1967, $9.00. *J Am Stat Assoc* 63, 1047–1049. doi: 10.1080/01621459.1968.11009335.

Kate, T., M., E, M., Shannon, N, S., Paul, Aamir, F., Frank, P., Kathryn, D., et al. (2006). Estimated Numbers of Community Cases of Illness Due to Salmonella, Campylobacter and Verotoxigenic Escherichia Coli: Pathogen-Specific Community Rates. *Can J Infect Dis Medical Microbiol* 17, 229–234. doi: 10.1155/2006/806874.

Kaupper, T., Mendes, L. W., Harnisz, M., Krause, S. M. B., Horn, M. A., and Ho, A. (2021a). Recovery of Methanotrophic Activity Is Not Reflected in the Methane-Driven Interaction Network after Peat Mining. *Appl Environ Microb* 87. doi: 10.1128/aem.02355-20.

Kaupper, T., Mendes, L. W., Lee, H. J., Mo, Y., Poehlein, A., Jia, Z., et al. (2021b). When the going gets tough: Emergence of a complex methane-driven interaction network during recovery

from desiccation-rewetting. *Soil Biology Biochem* 153, 108109. doi: 10.1016/j.soilbio.2020.108109.

Keen, J. E., Laegreid, W. W., Chitko-McKown, C. G., Durso, L. M., and Bono, J. L. (2010). Distribution of Shiga-Toxigenic Escherichia coli O157 in the Gastrointestinal Tract of Naturally O157-Shedding Cattle at Necropsy. *Appl. Environ. Microbiol.* 76, 5278–5281. doi: 10.1128/aem.00400-10.

Kim, H.-J., and Song, W.-J. (2022). Inactivation of Escherichia coli O157: H7 in foods by emerging technologies: a review. *Lett Appl Microbiol* 76. doi: 10.1093/lambio/ovac007.

Kimball, S. R. (1999). Eukaryotic initiation factor eIF2. *The Int. J. Biochem. Cell Biology* 31, 25–29. doi: 10.1016/s1357-2725(98)00128-9.

Kingsford, C., and Salzberg, S. L. (2008). What are decision trees? *Nat Biotechnol* 26, 1011–1013. doi: 10.1038/nbt0908-1011.

Koch, C., Hertwig, S., Lurz, R., Appel, B., and Beutin, L. (2001). Isolation of a Lysogenic Bacteriophage Carrying thestx 1OX3 Gene, Which Is Closely Associated with Shiga Toxin-Producing Escherichia coli Strains from Sheep and Humans. *J Clin Microbiol* 39, 3992–3998. doi: 10.1128/jcm.39.11.3992-3998.2001.

Kohonen, T. (1990). The self-organizing map. *P Ieee* 78, 1464–1480. doi: 10.1109/5.58325.

Korn, T., Bettelli, E., Oukka, M., and Kuchroo, V. K. (2009). IL-17 and Th17 Cells. *Immunology* 27, 485–517. doi: 10.1146/annurev.immunol.021908.132710.

Kotsiantis, S. B., Zaharakis, I. D., and Pintelas, P. E. (2006). Machine learning: a review of classification and combining techniques. *Artif Intell Rev* 26, 159–190. doi: 10.1007/s10462-007-9052-3.

Krause, M., Barth, H., and Schmidt, H. (2018). Toxins of Locus of Enterocyte Effacement-Negative Shiga Toxin-Producing Escherichia coli. *Toxins* 10, 241. doi: 10.3390/toxins10060241.

Krysenko, S., Matthews, A., Okoniewski, N., Kulik, A., Girbas, M. G., Tsypik, O., et al. (2019). Initial Metabolic Step of a Novel Ethanolamine Utilization Pathway and Its Regulation in Streptomyces coelicolor M145. *Mbio* 10. doi: 10.1128/mbio.00326-19.

Kuijpers, T. W., Bende, R. J., Baars, P. A., Grummels, A., Derks, I. A. M., Dolman, K. M., et al. (2010). CD20 deficiency in humans results in impaired T cell–independent antibody responses. *J. Clin. Investig.* 120, 214–222. doi: 10.1172/jci40231.

Kursa, M. B., and Rudnicki, W. R. (2010). Feature Selection with the Boruta Package. *J. Stat. Softw.* 36. doi: 10.18637/jss.v036.i11.

Larsen, J. M. (2017). The immune response to Prevotella bacteria in chronic inflammatory disease. *Immunology* 151, 363–374. doi: 10.1111/imm.12760.

Lecca, P., and Re, A. (2015). Detecting modules in biological networks by edge weight clustering and entropy significance. *Frontiers Genetics* 6, 265. doi: 10.3389/fgene.2015.00265.

Lee, J., Warner, E., Shaikhouni, S., Bitzer, M., Kretzler, M., Gipson, D., et al. (2022). Unsupervised machine learning for identifying important visual features through bag-of-words using histopathology data from chronic kidney disease. *Sci Rep-uk* 12, 4832. doi: 10.1038/s41598-022-08974-8.

Lee, M.-S., and Tesh, V. L. (2019). Roles of Shiga Toxins in Immunopathology. *Toxins* 11, 212. doi: 10.3390/toxins11040212.

Legros, N., Pohlentz, G., Steil, D., and Müthing, J. (2018). Shiga toxin-glycosphingolipid interaction: status quo of research with focus on primary human brain and kidney endothelial cells. *Int J Med Microbiol* 308, 1073–1084. doi: 10.1016/j.ijmm.2018.09.003.

Lehnert, S. A., Reverter, A., Byrne, K. A., Wang, Y., Nattrass, G. S., Hudson, N. J., et al. (2007). Gene expression studies of developing bovine longissimusmuscle from two different beef cattle breeds. *BMC Dev. Biol.* 7, 95. doi: 10.1186/1471-213x-7-95.

Leimbach, A., Hacker, J., and Dobrindt, U. (2013). Between Pathogenicity and Commensalism. *Curr Top Microbiol* 358, 3–32. doi: 10.1007/82_2012_303.

Leung, P. H. M., Peiris, J. S. M., Ng, W. W. S., Robins-Browne, R. M., Bettelheim, K. A., and Yam, W. C. (2003). A Newly Discovered Verotoxin Variant, VT2g, Produced by Bovine

Verocytotoxigenic Escherichia coli. *Appl Environ Microb* 69, 7549–7553. doi: 10.1128/aem.69.12.7549-7553.2003.

Levins, R. (n.d.). *Evolution in Changing Environments*. Princeton University Press doi: 10.1515/9780691209418.

Li, M., Penner, G. B., Hernandez-Sanabria, E., Oba, M., and Guan, L. L. (2009). Effects of sampling location and time, and host animal on assessment of bacterial diversity and fermentation parameters in the bovine rumen. *J. Appl. Microbiol.* 107, 1924–1934. doi: 10.1111/j.1365-2672.2009.04376.x.

Li, X., He, L., Luo, J., Zheng, Y., Zhou, Y., Li, D., et al. (2022). Paeniclostridium sordellii hemorrhagic toxin targets TMPRSS2 to induce colonic epithelial lesions. *Nat Commun* 13, 4331. doi: 10.1038/s41467-022-31994-x.

Liao, J., Cao, X., Zhao, L., Wang, J., Gao, Z., Wang, M. C., et al. (2016). The importance of neutral and niche processes for bacterial community assembly differs between habitat generalists and specialists. *Fems Microbiol Ecol* 92, fiw174. doi: 10.1093/femsec/fiw174.

Lidicker, W. Z. (1979). A Clarification of Interactions in Ecological Systems. *Bioscience* 29, 475–477. doi: 10.2307/1307540.

Lim, J. Y., Yoon, J., and Hovde, C. J. (2010). A brief overview of Escherichia coli O157:H7 and its plasmid O157. *J Microbiol Biotechn* 20, 5–14.

Lisboa, Szelewicki, Lin, Latonas, Li, Zhi, et al. (2019). Epidemiology of Shiga Toxin-Producing Escherichia coli O157 in the Province of Alberta, Canada, 2009–2016. *Toxins* 11, 613. doi: 10.3390/toxins11100613.

Liu, J., Meng, Z., Liu, X., and Zhang, X.-H. (2019). Microbial assembly, interaction, functioning, activity and diversification: a review derived from community compositional data. *Mar Life Sci Technology* 1, 112–128. doi: 10.1007/s42995-019-00004-3.

Liu, S., Lai, W., Shi, Y., Liu, N., Ouyang, L., Zhang, Z., et al. (2020). Annotation and cluster analysis of long noncoding RNA linked to male sex and estrogen in cancers. *npj Precis. Oncol.* 4, 5. doi: 10.1038/s41698-020-0110-5.

Liu, S., Yu, H., Yu, Y., Huang, J., Zhou, Z., Zeng, J., et al. (2022). Ecological stability of microbial communities in Lake Donghu regulated by keystone taxa. *Ecol. Indic.* 136, 108695. doi: 10.1016/j.ecolind.2022.108695.

Logares, R., Lindström, E. S., Langenheder, S., Logue, J. B., Paterson, H., Laybourn-Parry, J., et al. (2013). Biogeography of bacterial communities exposed to progressive long-term environmental change. *Isme J* 7, 937–948. doi: 10.1038/ismej.2012.168.

Lomana, A. L. G. de, Kusebauch, U., Raman, A. V., Pan, M., Turkarslan, S., Lorenzetti, A. P. R., et al. (2020). Selective Translation of Low Abundance and Upregulated Transcripts in Halobacterium salinarum. *mSystems* 5, e00329-20. doi: 10.1128/msystems.00329-20.

Łoś, M., Kuzio, J., McConnell, M. R., Kropinski, A. M., Węgrzyn, G., and Christie, G. E. (2020). Bacteriophages in the Control of Food- and Waterborne Pathogens. 157–198. doi: 10.1128/9781555816629.ch9.

Louis, P., and Flint, H. J. (2017). Formation of propionate and butyrate by the human colonic microbiota. *Environ. Microbiol.* 19, 29–41. doi: 10.1111/1462-2920.13589.

Lu, L., Li, X., Li, Z., Chen, Y., García, C. A. S. y, Yang, J., et al. (2021). Aerobic methanotrophs in an urban water cycle system: Community structure and network interaction pattern. *Sci Total Environ* 772, 145045. doi: 10.1016/j.scitotenv.2021.145045.

Ludwig, K., Karmali, M. A., Sarkim, V., Bobrowski, C., Petric, M., Karch, H., et al. (2001). Antibody Response to Shiga Toxins Stx2 and Stx1 in Children with Enteropathic Hemolytic-Uremic Syndrome. *J Clin Microbiol* 39, 2272–2279. doi: 10.1128/jcm.39.6.2272-2279.2001.

Luo, F., Yang, Y., Zhong, J., Gao, H., Khan, L., Thompson, D. K., et al. (2007). Constructing gene co-expression networks and predicting functions of unknown genes by random matrix theory. *Bmc Bioinformatics* 8, 299. doi: 10.1186/1471-2105-8-299.

Lynn, R. M., O'Brien, S. J., Taylor, C. M., Adak, G. K., Chart, H., Cheasty, T., et al. (2005). Childhood Hemolytic Uremic Syndrome, United Kingdom and Ireland - Volume 11, Number 4—April 2005 - Emerging Infectious Diseases journal - CDC. *Emerg Infect Dis* 11, 590–596. doi: 10.3201/eid1104.040833.

Ma, B., and Hottiger, M. O. (2016). Crosstalk between Wnt/β-Catenin and NF-κB Signaling Pathway during Inflammation. *Frontiers Immunol.* 7, 378. doi: 10.3389/fimmu.2016.00378.

MacFabe, D. F., Cain, N. E., Boon, F., Ossenkopp, K.-P., and Cain, D. P. (2011). Effects of the enteric bacterial metabolic product propionic acid on object-directed behavior, social behavior, cognition, and neuroinflammation in adolescent rats: Relevance to autism spectrum disorder. *Behav Brain Res* 217, 47–54. doi: 10.1016/j.bbr.2010.10.005.

Magee, D. A., Taraktsoglou, M., Killick, K. E., Nalpas, N. C., Browne, J. A., Park, S. D. E., et al. (2012). Global Gene Expression and Systems Biology Analysis of Bovine Monocyte-Derived Macrophages in Response to In Vitro Challenge with Mycobacterium bovis. *Plos One* 7, e32034. doi: 10.1371/journal.pone.0032034.

Mahecha, M. D., Martínez, A., Lischeid, G., and Beck, E. (2007). Nonlinear dimensionality reduction: Alternative ordination approaches for extracting and visualizing biodiversity patterns in tropical montane forest vegetation data. *Ecol. Inform.* 2, 138–149. doi: 10.1016/j.ecoinf.2007.05.002.

Mainil, J. G., Jacquemin, E. R., Kaeckenbeeck, A. E., and Pohl, P. H. (1993). Association between the effacing (eae) gene and the Shiga-like toxin-encoding genes in Escherichia coli isolates from cattle. *Am J Vet Res* 54, 1064–8.

Majowicz, S. E., Scallan, E., Jones-Bitton, A., Sargeant, J. M., Stapleton, J., Angulo, F. J., et al. (2014). Global Incidence of Human Shiga Toxin–Producing Escherichia coli Infections and

Deaths: A Systematic Review and Knowledge Synthesis. *Foodborne Pathog. Dis.* 11, 447–455. doi: 10.1089/fpd.2013.1704.

Maleki, F., Ovens, K., Hogan, D. J., and Kusalik, A. J. (2020). Gene Set Analysis: Challenges, Opportunities, and Future Research. *Frontiers Genet.* 11, 654. doi: 10.3389/fgene.2020.00654.

Malmuthuge, N., and Guan, L. L. (2017). Understanding host-microbial interactions in rumen: searching the best opportunity for microbiota manipulation. *J Anim Sci Biotechno* 8, 8. doi: 10.1186/s40104-016-0135-3.

Malmuthuge, N., Liang, G., Griebel, P. J., and Guan, L. L. (2019). Taxonomic and Functional Compositions of the Small Intestinal Microbiome in Neonatal Calves Provide a Framework for Understanding Early Life Gut Health. *Appl Environ Microb* 85. doi: 10.1128/aem.02534-18.

Maluta, R. P., Fairbrother, J. M., Stella, A. E., Rigobelo, E. C., Martinez, R., and Ávila, F. A. de (2014). Potentially pathogenic Escherichia coli in healthy, pasture-raised sheep on farms and at the abattoir in Brazil. *Vet Microbiol* 169, 89–95. doi: 10.1016/j.vetmic.2013.12.013.

Mao, S., Zhang, M., Liu, J., and Zhu, W. (2015). Characterising the bacterial microbiota across the gastrointestinal tracts of dairy cattle: membership and potential function. *Sci Rep-uk* 5, 16116. doi: 10.1038/srep16116.

Marcos-Zambrano, L. J., Karaduzovic-Hadziabdic, K., Turukalo, T. L., Przymus, P., Trajkovik, V., Aasmets, O., et al. (2021). Applications of Machine Learning in Human Microbiome

Studies: A Review on Feature Selection, Biomarker Identification, Disease Prediction and Treatment. *Front Microbiol* 12, 634511. doi: 10.3389/fmicb.2021.634511.

Martínez-Álvaro, M., Auffret, M. D., Stewart, R. D., Dewhurst, R. J., Duthie, C.-A., Rooke, J. A., et al. (2020). Identification of Complex Rumen Microbiome Interaction Within Diverse Functional Niches as Mechanisms Affecting the Variation of Methane Emissions in Bovine. *Front Microbiol* 11, 659. doi: 10.3389/fmicb.2020.00659.

Maslov, S., and Sneppen, K. (2002). Specificity and stability in topology of protein networks. *Arxiv*. doi: 10.1126/science.1065103.

MATTHEWS, L., McKENDRICK, I. J., TERNENT, H., GUNN, G. J., SYNGE, B., and WOOLHOUSE, M. E. J. (2005). Super-shedding cattle and the transmission dynamics of Escherichia coli O157. *Epidemiology Infect.* 134, 131–142. doi: 10.1017/s0950268805004590.

McCabe, E., Burgess, C. M., Lawal, D., Whyte, P., and Duffy, G. (2019). An investigation of shedding and super-shedding of Shiga toxigenic Escherichia coli O157 and E. coli O26 in cattle presented for slaughter in the Republic of Ireland. *Zoonoses Public Hlth* 66, 83–91. doi: 10.1111/zph.12531.

McDonald, K. G., McDonough, J. S., and Newberry, R. D. (2005). Adaptive Immune Responses Are Dispensable for Isolated Lymphoid Follicle Formation: Antigen-Naive, Lymphotoxin-Sufficient B Lymphocytes Drive the Formation of Mature Isolated Lymphoid Follicles. *J. Immunol.* 174, 5720–5728. doi: 10.4049/jimmunol.174.9.5720.

McPherson, M., Kirk, M. D., Raupach, J., Combs, B., and Butler, J. R. G. (2011). Economic Costs of Shiga Toxin–Producing Escherichia coli Infection in Australia. *Foodborne Pathog Dis* 8, 55–62. doi: 10.1089/fpd.2010.0608.

Melton-Celsa, A. R. (2014a). Shiga Toxin (Stx) Classification, Structure, and Function. *Microbiol Spectr* 2. doi: 10.1128/microbiolspec.ehec-0024-2013.

Melton-Celsa, A. R. (2014b). Shiga Toxin (Stx) Classification, Structure, and Function. *Microbiol Spectr* 2, EHEC-0024-2013.

Memmott, J., Waser, N. M., and Price, M. V. (2004). Tolerance of pollination networks to species extinctions. *Proc Royal Soc Lond Ser B Biological Sci* 271, 2605–2611. doi: 10.1098/rspb.2004.2909.

Menge, C. (2020). The Role of Escherichia coli Shiga Toxins in STEC Colonization of Cattle. *Toxins* 12, 607. doi: 10.3390/toxins12090607.

Menge, C., Stamm, I., Wuhrer, M., Geyer, R., Wieler, L. H., and Baljer, G. (2001). Globotriaosylceramide (Gb3/CD77) is synthesized and surface expressed by bovine lymphocytes upon activation in vitro. *Vet Immunol Immunop* 83, 19–36. doi: 10.1016/s0165-2427(01)00365-8.

Mills, K. H. G. (2023). IL-17 and IL-17-producing cells in protection versus pathology. *Nat. Rev. Immunol.* 23, 38–54. doi: 10.1038/s41577-022-00746-9.

Mir, R. A., Schaut, R. G., Allen, H. K., Looft, T., Loving, C. L., Kudva, I. T., et al. (2019). Cattle intestinal microbiota shifts following Escherichia coli O157:H7 vaccination and colonization. *Plos One* 14, e0226099. doi: 10.1371/journal.pone.0226099.

Mir, R. A., Weppelmann, T. A., Elzo, M., Ahn, S., Driver, J. D., and Jeong, K. C. (2016). Colonization of Beef Cattle by Shiga Toxin-Producing Escherichia coli during the First Year of Life: A Cohort Study. *Plos One* 11, e0148518. doi: 10.1371/journal.pone.0148518.

Miyazawa, K., Hondo, T., Kanaya, T., Tanaka, S., Takakura, I., Itani, W., et al. (2010). Characterization of newly established bovine intestinal epithelial cell line. *Histochem. Cell Biol.* 133, 125–134. doi: 10.1007/s00418-009-0648-3.

Mootha, V. K., Lindgren, C. M., Eriksson, K.-F., Subramanian, A., Sihag, S., Lehar, J., et al. (2003). PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet* 34, 267–273. doi: 10.1038/ng1180.

Morris, J. S., Stein, T., Pringle, M., Davies, C. R., Weber-Hall, S., Ferrier, R. K., et al. (2006). Involvement of axonal guidance proteins and their signaling partners in the developing mouse mammary gland. *J. Cell. Physiol.* 206, 16–24. doi: 10.1002/jcp.20427.

Mühldorfer, I., Hacker, J., Keusch, G. T., Acheson, D. W., Tschäpe, H., Kane, A. V., et al. (1996). Regulation of the Shiga-like toxin II operon in Escherichia coli. *Infect Immun* 64, 495–502.

Munns, K. D., Selinger, L. B., Stanford, K., Guan, L., Callaway, T. R., and McAllister, T. A. (2015). Perspectives on Super-Shedding of Escherichia coli O157:H7 by Cattle. *Foodborne Pathog Dis* 12, 89–103. doi: 10.1089/fpd.2014.1829.

Nadler, C., Baruch, K., Kobi, S., Mills, E., Haviv, G., Farago, M., et al. (2010). The Type III Secretion Effector NleE Inhibits NF-κB Activation. *Plos Pathog* 6, e1000743. doi: 10.1371/journal.ppat.1000743.

Nakata, R., Tanaka, F., Sugawara, N., Kojima, Y., Takeuchi, T., Shiba, M., et al. (2022). Analysis of autonomic function during natural defecation in patients with irritable bowel syndrome using real-time recording with a wearable device. *PLOS ONE* 17, e0278922. doi: 10.1371/journal.pone.0278922.

Nart, P., Naylor, S. W., Huntley, J. F., McKendrick, I. J., Gally, D. L., and Low, J. C. (2008). Responses of Cattle to Gastrointestinal Colonization by Escherichia coli O157:H7 ▿ . *Infect Immun* 76, 5366–5372. doi: 10.1128/iai.01223-07.

Nataro, J. P., and Kaper, J. B. (1998). Diarrheagenic Escherichia coli. *Clin Microbiol Rev* 11, 142–201.

Newman, M. E. J. (2006). Modularity and community structure in networks. *Proc National Acad Sci* 103, 8577–8582. doi: 10.1073/pnas.0601602103.

Newton, H. J., Sloan, J., Bulach, D. M., Seemann, T., Allison, C. C., Tauschek, M., et al. (2009). Shiga Toxin–producing Escherichia coli Strains Negative for Locus of Enterocyte Effacement. *Emerg Infect Dis* 15, 372–380. doi: 10.3201/eid1502.080631.

Ngo, V. L., Abo, H., Kuczma, M., Szurek, E., Moore, N., Medina-Contreras, O., et al. (2020). IL-36R signaling integrates innate and adaptive immune-mediated protection against enteropathogenic bacteria. *Proc. National Acad. Sci.* 117, 27540–27548. doi: 10.1073/pnas.2004484117.

Nicholson, J. K., Holmes, E., Kinross, J., Burcelin, R., Gibson, G., Jia, W., et al. (2012). Host-Gut Microbiota Metabolic Interactions. *Science* 336, 1262–1267. doi: 10.1126/science.1223813.

Nicholson, J. K., Holmes, E., Lindon, J. C., and Wilson, I. D. (2004). The challenges of modeling mammalian biocomplexity. *Nat. Biotechnol.* 22, 1268–1274. doi: 10.1038/nbt1015.

O'Hara, E., Neves, A. L. A., Song, Y., and Guan, L. L. (2020). The Role of the Gut Microbiome in Cattle Production and Health: Driver or Passenger? *Annu Rev Anim Biosci* 8, 199–220. doi: 10.1146/annurev-animal-021419-083952.

Onyeka, L. O., Adesiyun, A. A., Keddy, K. H., Manqele, A., Madoroba, E., and Thompson, P. N. (2021). Prevalence, risk factors and molecular characteristics of Shiga toxin-producing Escherichia coli in beef abattoirs in Gauteng, South Africa. *Food Control* 123, 107746. doi: 10.1016/j.foodcont.2020.107746.

Pan, Z., Chen, Y., McAllister, T. A., Gänzle, M., Plastow, G., and Guan, L. L. (2021a). Abundance and Expression of Shiga Toxin Genes in Escherichia coli at the Recto-Anal Junction Relates to Host Immune Genes. *Front Cell Infect Mi* 11, 633573. doi: 10.3389/fcimb.2021.633573.

Pan, Z., Chen, Y., Zhou, M., McAllister, T. A., and Guan, L. L. (2021b). Microbial interaction-driven community differences as revealed by network analysis. *Comput Struct Biotechnology J* 19, 6000–6008. doi: 10.1016/j.csbj.2021.10.035.

Pandit, S. N., Kolasa, J., and Cottenie, K. (2009). Contrasts between habitat generalists and specialists: an empirical extension to the basic metacommunity framework. *Ecology* 90, 2253–2262. doi: 10.1890/08-0851.1.

Panel, E. B., Koutsoumanis, K., Allende, A., Alvarez-Ordóñez, A., Bover-Cid, S., Chemaly, M., et al. (2020). Pathogenicity assessment of Shiga toxin-producing Escherichia coli (STEC) and the public health risk posed by contamination of food with STEC. *Efsa J* 18. doi: 10.2903/j.efsa.2020.5967.

Paton, A. W., Beutin, L., and Paton, J. C. (1995). Heterogeneity of the amino-acid sequences of Escherichia coli shiga-like toxin type-I operons. *Gene* 153, 71–74. doi: 10.1016/0378-1119(94)00777-p.

Paton, A. W., and Paton, J. C. (1996). Enterobacter cloacae producing a Shiga-like toxin II-related cytotoxin associated with a case of hemolytic-uremic syndrome. *J. Clin. Microbiol.* 34, 463–465. doi: 10.1128/jcm.34.2.463-465.1996.

Paton, A. W., Paton, J. C., and Manning, P. A. (1993). Polymerase chain reaction amplification, cloning and sequencing of variant Escherichia coli Shiga-like toxin type II operons. *Microb Pathogenesis* 15, 77–82. doi: 10.1006/mpat.1993.1058.

Paton, J. C., and Paton, A. W. (1998). Pathogenesis and Diagnosis of Shiga Toxin-Producing Escherichia coli Infections. *Clin Microbiol Rev* 11, 450–479. doi: 10.1128/cmr.11.3.450.

Perna, N. T., Mayhew, G. F., Pósfai, G., Elliott, S., Donnenberg, M. S., Kaper, J. B., et al. (1998). Molecular Evolution of a Pathogenicity Island from Enterohemorrhagic Escherichia coli O157:H7. *Infect Immun* 66, 3810–3817. doi: 10.1128/iai.66.8.3810-3817.1998.

Peroutka-Bigus, N., Nielsen, D. W., Trachsel, J., Mou, K. T., Sharma, V. K., Kudva, I. T., et al. (2022). Phenotypic and genomic comparison of human outbreak and cattle-associated Shiga toxin-producing Escherichia coli O157:H7. *Biorxiv*, 2022.09.30.510420. doi: 10.1101/2022.09.30.510420.

Persson, S., Olsen, K. E. P., Ethelberg, S., and Scheutz, F. (2007). Subtyping Method for Escherichia coli Shiga Toxin (Verocytotoxin) 2 Variants and Correlations to Clinical Manifestations ▿ . *J Clin Microbiol* 45, 2020–2024. doi: 10.1128/jcm.02591-06.

Petro, C. D., Trojnar, E., Sinclair, J., Liu, Z.-M., Smith, M., O'Brien, A. D., et al. (2019). Shiga Toxin Type 1a (Stx1a) Reduces the Toxicity of the More Potent Stx2a In Vivo and In Vitro. *Infect Immun* 87. doi: 10.1128/iai.00787-18.

Philpott, D. J., Ackerley, C. A., Kiliaan, A. J., Karmali, M. A., Perdue, M. H., and Sherman, P. M. (1997). Translocation of verotoxin-1 across T84 monolayers: mechanism of bacterial toxin penetration of epithelium. *Am J Physiol-gastr L* 273, G1349–G1358. doi: 10.1152/ajpgi.1997.273.6.g1349.

Piérard, D., Muyldermans, G., Moriau, L., Stevens, D., and Lauwers, S. (1998). Identification of new verocytotoxin type 2 variant B-subunit genes in human and animal Escherichia coli isolates. *J Clin Microbiol* 36, 3317–22.

Ponchel, F., Toomes, C., Bransfield, K., Leong, F. T., Douglas, S. H., Field, S. L., et al. (2003). Real-time PCR based on SYBR-Green I fluorescence: An alternative to the TaqMan assay for a relative quantification of gene rearrangements, gene amplifications and micro gene deletions. *Bmc Biotechnol* 3, 18. doi: 10.1186/1472-6750-3-18.

Poretsky, R., Rodriguez-R, L. M., Luo, C., Tsementzi, D., and Konstantinidis, K. T. (2014). Strengths and Limitations of 16S rRNA Gene Amplicon Sequencing in Revealing Temporal Microbial Community Dynamics. *Plos One* 9, e93827. doi: 10.1371/journal.pone.0093827.

Poudel, R., Jumpponen, A., Schlatter, D. C., Paulitz, T. C., Gardener, B. B. M., Kinkel, L. L., et al. (2016). Microbiome Networks: A Systems Framework for Identifying Candidate

Microbial Assemblages for Disease Management. *Phytopathology* 106, 1083–1096. doi: 10.1094/phyto-02-16-0058-fi.

Proença, J. T., Barral, D. C., and Gordo, I. (2017). Commensal-to-pathogen transition: One-single transposon insertion results in two pathoadaptive traits in Escherichia coli -macrophage interaction. *Sci Rep-uk* 7, 4504. doi: 10.1038/s41598-017-04081-1.

Pryde, S. E., Duncan, S. H., Hold, G. L., Stewart, C. S., and Flint, H. J. (2002). The microbiology of butyrate formation in the human colon. *Fems Microbiol Lett* 217, 133–139. doi: 10.1111/j.1574-6968.2002.tb11467.x.

Puente, J. L., and Finlay, B. B. (2001). Principles of Bacterial Pathogenesis. 387–456. doi: 10.1016/b978-012304220-0/50010-8.

Qin, L., Guo, J., Zheng, Q., and Zhang, H. (2016). BAG2 structure, function and involvement in disease. *Cell. Mol. Biology Lett.* 21, 18. doi: 10.1186/s11658-016-0020-2.

Raa, H., Grimmer, S., Schwudke, D., Bergan, J., Wälchli, S., Skotland, T., et al. (2009). Glycosphingolipid Requirements for Endosome-to-Golgi Transport of Shiga Toxin. *Traffic* 10, 868–882. doi: 10.1111/j.1600-0854.2009.00919.x.

Rangel, J. M., Sparling, P. H., Crowe, C., Griffin, P. M., and Swerdlow, D. L. (2005). Epidemiology of Escherichia coli O157:H7 Outbreaks, United States, 1982–2002 - Volume 11, Number 4—April 2005 - Emerging Infectious Diseases journal - CDC. *Emerg Infect Dis* 11, 603–609. doi: 10.3201/eid1104.040739.

Ravasz, E., Somera, A. L., Mongru, D. A., Oltvai, Z. N., and Barabási, A.-L. (2002). Hierarchical Organization of Modularity in Metabolic Networks. *Science* 297, 1551–1555. doi: 10.1126/science.1073374.

Rebala, G., Ravi, A., and Churiwala, S. (2019). An Introduction to Machine Learning. 1–17. doi: 10.1007/978-3-030-15729-6_1.

Reimand, J., Isserlin, R., Voisin, V., Kucera, M., Tannus-Lopes, C., Rostamianfar, A., et al. (2019). Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. *Nat Protoc* 14, 482–517. doi: 10.1038/s41596-018-0103-9.

Reinstein, S., Fox, J. T., Shi, X., and Nagaraja, T. G. (2007). Prevalence of Escherichia coli O157:H7 in Gallbladders of Beef Cattle. *Appl Environ Microb* 73, 1002–1004. doi: 10.1128/aem.02037-06.

Reusch, D. B., Alley, R. B., and Hewitson, B. C. (2005). Relative Performance of Self-Organizing Maps and Principal Component Analysis in Pattern Extraction from Synthetic Climatological Data. *Polar Geogr* 29, 188–212. doi: 10.1080/789610199.

Reynolds, C., Checkley, S., Chui, L., Otto, S., and Neumann, N. F. (2020). Evaluating the risks associated with Shiga-toxin-producing Escherichia coli (STEC) in private well waters in Canada. *Can J Microbiol* 66, 337–350. doi: 10.1139/cjm-2019-0329.

Rhoads, D. D., Cox, S. B., Rees, E. J., Sun, Y., and Wolcott, R. D. (2012). Clinical identification of bacteria in human chronic wound infections: culturing vs. 16S ribosomal DNA sequencing. *Bmc Infect Dis* 12, 321. doi: 10.1186/1471-2334-12-321.

Riley, L. W., Remis, R. S., Helgerson, S. D., McGee, H. B., Wells, J. G., Davis, B. R., et al. (1983). Hemorrhagic Colitis Associated with a Rare Escherichia coli Serotype. *New Engl J Medicine* 308, 681–685. doi: 10.1056/nejm198303243081203.

Rodríguez-Rubio, L., Haarmann, N., Schwidder, M., Muniesa, M., and Schmidt, H. (2021). Bacteriophages of Shiga Toxin-Producing Escherichia coli and Their Contribution to Pathogenicity. *Pathogens* 10, 404. doi: 10.3390/pathogens10040404.

Roussel, C., Cordonnier, C., Livrelli, V., de Wiele, T. V., and Blanquet-Diot, S. (2017). *Escherichia coli* - Recent Advances on Physiology, Pathogenesis and Biotechnological Applications. doi: 10.5772/intechopen.68309.

Roxas, J. L., Koutsouris, A., Bellmeyer, A., Tesfay, S., Royan, S., Falzari, K., et al. (2010). Enterohemorrhagic E. coli alters murine intestinal epithelial tight junction protein expression and barrier function in a Shiga toxin independent manner. *Lab. Investig.* 90, 1152–1168. doi: 10.1038/labinvest.2010.91.

Ruan, Q., Dutta, D., Schwalbach, M. S., Steele, J. A., Fuhrman, J. A., and Sun, F. (2006). Local similarity analysis reveals unique associations among marine bacterioplankton species and environmental factors. *Bioinformatics* 22, 2532–2538. doi: 10.1093/bioinformatics/btl417.

Russell, S. A., and Bashaw, G. J. (2018). Axon guidance pathways and the control of gene expression. *Dev. Dyn.* 247, 571–580. doi: 10.1002/dvdy.24609.

Ryu, J.-H., Kim, S., Park, J., and Choi, K.-S. (2020). Characterization of virulence genes in Escherichia coli strains isolated from pre-weaned calves in the Republic of Korea. *Acta Vet Scand* 62, 45. doi: 10.1186/s13028-020-00543-1.

Sakamoto, K., Karelina, K., and Obrietan, K. (2011). CREB: a multifaceted regulator of neuronal plasticity and protection. *J. Neurochem.* 116, 1–9. doi: 10.1111/j.1471-4159.2010.07080.x.

Samuel, J. E., Perera, L. P., Ward, S., O'Brien, A. D., Ginsburg, V., and Krivan, H. C. (1990). Comparison of the glycolipid receptor specificities of Shiga-like toxin type II and Shiga-like toxin type II variants. *Infect Immun* 58, 611–618. doi: 10.1128/iai.58.3.611-618.1990.

Sandhu, K. S., and Gyles, C. L. (2002). Pathogenic Shiga toxin-producing Escherichia coli in the intestine of calves. *Can J Vet Res Revue Can De Recherche Veterinaire* 66, 65–72.

Sandvig, K., Bergan, J., Dyve, A.-B., Skotland, T., and Torgersen, M. L. (2010). Endocytosis and retrograde transport of Shiga toxin. *Toxicon* 56, 1181–1185. doi: 10.1016/j.toxicon.2009.11.021.

Scallan, E., Hoekstra, R. M., Angulo, F. J., Tauxe, R. V., Widdowson, M.-A., Roy, S. L., et al. (2011). Foodborne Illness Acquired in the United States—Major Pathogens. *Emerg Infect Dis* 17, 7–15. doi: 10.3201/eid1701.p11101.

Scheutz, F., Teel, L. D., Beutin, L., Piérard, D., Buvens, G., Karch, H., et al. (2012). Multicenter Evaluation of a Sequence-Based Protocol for Subtyping Shiga Toxins and Standardizing Stx Nomenclature. *J Clin Microbiol* 50, 2951–2963. doi: 10.1128/jcm.00860-12.

Schierack, P., Kleta, S., Tedin, K., Babila, J. T., Oswald, S., Oelschlaeger, T. A., et al. (2011). E. coli Nissle 1917 Affects Salmonella Adhesion to Porcine Intestinal Epithelial Cells. *Plos One* 6, e14712. doi: 10.1371/journal.pone.0014712.

Schmidt, H. (2001). Shiga-toxin-converting bacteriophages. *Res Microbiol* 152, 687–695. doi: 10.1016/s0923-2508(01)01249-9.

Schmidt, H., Scheef, J., Morabito, S., Caprioli, A., Wieler, L. H., and Karch, H. (2000). A New Shiga Toxin 2 Variant (Stx2f) fromEscherichia coli Isolated from Pigeons. *Appl Environ Microb* 66, 1205–1208. doi: 10.1128/aem.66.3.1205-1208.2000.

Serra-Moreno, R., Jofre, J., and Muniesa, M. (2008). The CI Repressors of Shiga Toxin-Converting Prophages Are Involved in Coinfection of Escherichia coli Strains, Which Causes a Down Regulation in the Production of Shiga Toxin 2 ▿ . *J Bacteriol* 190, 4722–4735. doi: 10.1128/jb.00069-08.

Shafik, A. E., Shafik, A., EI-Sibai, O., and Ahmed, I. (2003). Role Of Sympathetic Innervation In The Defecation Mechanism: A Novel Concept Of Its Function. *The J. Spinal Cord Medicine* 26, 150–154. doi: 10.1080/10790268.2003.11753676.

Shafik, A., El-Sibai, O., and Ahmed, I. (2002). Parasympathetic extrinsic reflex: Role in defecation mechanism. *World J. Surg.* 26, 737–740. doi: 10.1007/s00268-002-6285-9.

Shahi, S. K., Freedman, S. N., and Mangalam, A. K. (2017). Gut microbiome in multiple sclerosis: The players involved and the roles they play. *Gut Microbes* 8, 607–615. doi: 10.1080/19490976.2017.1349041.

Shanks, O. C., Kelty, C. A., Archibeque, S., Jenkins, M., Newton, R. J., McLellan, S. L., et al. (2011). Community Structures of Fecal Bacteria in Cattle from Different Animal Feeding Operations†. *Appl Environ Microb* 77, 2992–3001. doi: 10.1128/aem.02988-10.

Sheng, H., Lim, J. Y., Knecht, H. J., Li, J., and Hovde, C. J. (2006). Role of Escherichia coli O157:H7 Virulence Factors in Colonization at the Bovine Terminal Rectal Mucosa. *Infect Immun* 74, 4685–4693. doi: 10.1128/iai.00406-06.

Sheng, H., Wang, J., Lim, J. Y., Davitt, C., Minnich, S. A., and Hovde, C. J. (2011). Internalization of Escherichia Coli O157:H7 by Bovine Rectal Epithelial Cells. *Front Microbiol* 2, 32. doi: 10.3389/fmicb.2011.00032.

Sherman, P. M., Johnson-Henry, K. C., Yeung, H. P., Ngo, P. S. C., Goulet, J., and Tompkins, T. A. (2005). Probiotics Reduce Enterohemorrhagic Escherichia coli O157:H7- and Enteropathogenic E. coli O127:H6-Induced Changes in Polarized T84 Epithelial Cell Monolayers by Reducing Bacterial Adhesion and Cytoskeletal Rearrangements. *Infect Immun* 73, 5183–5188. doi: 10.1128/iai.73.8.5183-5188.2005.

Shi, Y., Delgado-Baquerizo, M., Li, Y., Yang, Y., Zhu, Y.-G., Peñuelas, J., et al. (2020). Abundance of kinless hubs within soil microbial networks are associated with high functional potential in agricultural ecosystems. *Environ Int* 142, 105869. doi: 10.1016/j.envint.2020.105869.

Shrestha, N., Bahnan, W., Wiley, D. J., Barber, G., Fields, K. A., and Schesser, K. (2012). Eukaryotic Initiation Factor 2 (eIF2) Signaling Regulates Proinflammatory Cytokine Expression and Bacterial Invasion*. *J. Biological Chem.* 287, 28738–28744. doi: 10.1074/jbc.m112.375915.

Silva, Y. P., Bernardi, A., and Frozza, R. L. (2020). The Role of Short-Chain Fatty Acids From Gut Microbiota in Gut-Brain Communication. *Front Endocrinol* 11, 25. doi: 10.3389/fendo.2020.00025.

Smet, R. D., and Marchal, K. (2010). Advantages and limitations of current network inference methods. *Nat Rev Microbiol* 8, 717–729. doi: 10.1038/nrmicro2419.

Sogin, M. L., Morrison, H. G., Huber, J. A., Welch, D. M., Huse, S. M., Neal, P. R., et al. (2006). Microbial diversity in the deep sea and the underexplored “rare biosphere.” *Proc National Acad Sci* 103, 12115–12120. doi: 10.1073/pnas.0605127103.

Spooner, R. A., and Lord, J. M. (2011). Ricin and Shiga Toxins, Pathogenesis, Immunity, Vaccines and Therapeutics. *Curr Top Microbiol* 357, 19–40. doi: 10.1007/82_2011_154.

Sriram, K., Wiley, S. Z., Moyung, K., Gorr, M. W., Salmerón, C., Marucut, J., et al. (2019). Detection and Quantification of GPCR mRNA: An Assessment and Implications of Data from High-Content Methods. *ACS Omega* 4, 17048–17059. doi: 10.1021/acsomega.9b02811.

Stark, R., Grzelak, M., and Hadfield, J. (2019). RNA sequencing: the teenage years. *Nat Rev Genet* 20, 631–656. doi: 10.1038/s41576-019-0150-2.

Stegen, J. C., Lin, X., Konopka, A. E., and Fredrickson, J. K. (2012). Stochastic and deterministic assembly processes in subsurface microbial communities. *ISME J.* 6, 1653–1664. doi: 10.1038/ismej.2012.22.

Stein, R. A., and Katz, D. E. (2017). Escherichia coli, cattle and the propagation of disease. *Fems Microbiol Lett* 364, fnx050. doi: 10.1093/femsle/fnx050.

Stenkamp-Strahm, C., McConnel, C., Magzamen, S., Abdo, Z., and Reynolds, S. (2018). Associations between Escherichia coli O157 shedding and the faecal microbiota of dairy cows. *J Appl Microbiol* 124, 881–898. doi: 10.1111/jam.13679.

Stover, P., Zamora, M., Shostak, K., Gautam-Basak, M., and Schirch, V. (1992). Escherichia coli serine hydroxymethyltransferase. The role of histidine 228 in determining reaction specificity. *J. Biological Chem.* 267, 17679–17687. doi: 10.1016/s0021-9258(19)37096-6.

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-

wide expression profiles. *Proc National Acad Sci* 102, 15545–15550. doi: 10.1073/pnas.0506580102.

Sun, L., Jia, H., Li, J., Yu, M., Yang, Y., Tian, D., et al. (2019). Cecal Gut Microbiota and Metabolites Might Contribute to the Severity of Acute Myocardial Ischemia by Impacting the Intestinal Permeability, Oxidative Stress, and Energy Metabolism. *Front Microbiol* 10, 1745. doi: 10.3389/fmicb.2019.01745.

Tahamtan, Y., Hayati, M., and Namavari, M. (2010). Prevalence and distribution of the stx, stx genes in Shiga toxin producing E. coli (STEC) isolates from cattle. *Iran. J. Microbiol.* 2, 8–13.

Tamayo, P., Slonim, D., Mesirov, J., Zhu, Q., Kitareewan, S., Dmitrovsky, E., et al. (1999). Interpreting patterns of gene expression with self-organizing maps: Methods and application to hematopoietic differentiation. *Proc. National Acad. Sci.* 96, 2907–2912. doi: 10.1073/pnas.96.6.2907.

Tekaia, F. (2016). Genome Data Exploration Using Correspondence Analysis. *Bioinform. Biol. Insights* 10, BBI.S39614. doi: 10.4137/bbi.s39614.

Tenenbaum, J. B., Silva, V. de, and Langford, J. C. (2000). A Global Geometric Framework for Nonlinear Dimensionality Reduction. *Science* 290, 2319–2323. doi: 10.1126/science.290.5500.2319.

Tesh, V. L., Burris, J. A., Owens, J. W., Gordon, V. M., Wadolkowski, E. A., O'Brien, A. D., et al. (1993). Comparison of the relative toxicities of Shiga-like toxins type I and type II for mice. *Infect Immun* 61, 3392–402.

Tian, L., Zhang, Y., Zhang, L., Zhang, L., Gao, X., and Feng, B. (2023). Biogeographic Pattern and Network of Rhizosphere Fungal and Bacterial Communities in Panicum miliaceum Fields: Roles of Abundant and Rare Taxa. *Microorg* 11, 134. doi: 10.3390/microorganisms11010134.

Tian, S., Liu, Y., Wu, H., Liu, H., Zeng, J., Choi, M. Y., et al. (2020). Genome-Wide CRISPR Screen Identifies Semaphorin 6A and 6B as Receptors for Paeniclostridium sordellii Toxin TcsL. *Cell Host Microbe* 27, 782-792.e7. doi: 10.1016/j.chom.2020.03.007.

Toeuf, B. de, Soin, R., Nazih, A., Dragojevic, M., Jurėnas, D., Delacourt, N., et al. (2018). ARE-mediated decay controls gene expression and cellular metabolism upon oxygen variations. *Sci. Reports* 8, 5211. doi: 10.1038/s41598-018-23551-8.

Törönen, P., Kolehmainen, M., Wong, G., and Castrén, E. (1999). Analysis of gene expression data using self-organizing maps. *FEBS Lett.* 451, 142–146. doi: 10.1016/s0014-5793(99)00524-4.

Tsuji, M., Suzuki, K., Kitamura, H., Maruya, M., Kinoshita, K., Ivanov, I. I., et al. (2008). Requirement for Lymphoid Tissue-Inducer Cells in Isolated Follicle Formation and T Cell-Independent Immunoglobulin A Generation in the Gut. *Immunity* 29, 261–271. doi: 10.1016/j.immuni.2008.05.014.

Uchida, J., Lee, Y., Hasegawa, M., Liang, Y., Bradney, A., Oliver, J. A., et al. (2004). Mouse CD20 expression and function. *Int. Immunol.* 16, 119–129. doi: 10.1093/intimm/dxh009.

Van, T. D. le, Robinson, J. A., Ralph, J., Greening, R. C., Smolenski, W. J., Leedle, J. A. Z., et al. (1998). Assessment of Reductive Acetogenesis with Indigenous Ruminal Bacterium Populations and Acetitomaculum ruminis. *Appl Environ Microb* 64, 3429–3436. doi: 10.1128/aem.64.9.3429-3436.1998.

Vasco, K., Nohomovich, B., Singh, P., Venegas-Vargas, C., Mosci, R. E., Rust, S., et al. (2021). Characterizing the Cattle Gut Microbiome in Farms with a High and Low Prevalence of Shiga Toxin Producing Escherichia coli. *Microorg* 9, 1737. doi: 10.3390/microorganisms9081737.

Vedell, P. T., Svenson, K. L., and Churchill, G. A. (2011). Stochastic variation of transcript abundance in C57BL/6J mice. *BMC Genom.* 12, 167. doi: 10.1186/1471-2164-12-167.

VELDEN, V. D., and HULSMANN, A. R. (1999). Peptidases: structure, function and modulation of peptide-mediated effects in the human lung. *Clin. Exp. Allergy* 29, 445–456. doi: 10.1046/j.1365-2222.1999.00462.x.

Venegas-Vargas, C., Henderson, S., Khare, A., Mosci, R. E., Lehnert, J. D., Singh, P., et al. (2016). Factors Associated with Shiga Toxin-Producing Escherichia coli Shedding by Dairy and Beef Cattle. *Appl Environ Microb* 82, 5049–5056. doi: 10.1128/aem.00829-16.

Verstraete, K., Coillie, E. V., Werbrouck, H., Weyenberg, S. V., Herman, L., Del-Favero, J., et al. (2014). A qPCR Assay to Detect and Quantify Shiga Toxin-Producing E. coli (STEC) in

Cattle and on Farms: A Potential Predictive Tool for STEC Culture-Positive Farms. *Toxins* 6, 1201–1221. doi: 10.3390/toxins6041201.

Vidor, C. J., Bulach, D., Awad, M., and Lyras, D. (2019). Paeniclostridium sordellii and Clostridioides difficile encode similar and clinically relevant tetracycline resistance loci in diverse genomic locations. *Bmc Microbiol* 19, 53. doi: 10.1186/s12866-019-1427-5.

Vlisidou, I., Lyte, M., Diemen, P. M. van, Hawes, P., Monaghan, P., Wallis, T. S., et al. (2004). The Neuroendocrine Stress Hormone Norepinephrine Augments Escherichia coli O157:H7-Induced Enteritis and Adherence in a Bovine Ligated Ileal Loop Model of Infection. *Infect. Immun.* 72, 5446–5451. doi: 10.1128/iai.72.9.5446-5451.2004.

Wagner, P. L., Livny, J., Neely, M. N., Acheson, D. W. K., Friedman, D. I., and Waldor, M. K. (2002). Bacteriophage control of Shiga toxin 1 production and release by Escherichia coli. *Mol Microbiol* 44, 957–970. doi: 10.1046/j.1365-2958.2002.02950.x.

Walle, K. V., Vanrompay, D., and Cox, E. (2013). Bovine innate and adaptive immune responses against Escherichia coli O157:H7 and vaccination strategies to reduce faecal shedding in ruminants. *Vet Immunol Immunop* 152, 109–120. doi: 10.1016/j.vetimm.2012.09.028.

Wan, Y., Zuo, T., Xu, Z., Zhang, F., Zhan, H., CHAN, D., et al. (2021). Underdevelopment of the gut microbiota and bacteria species as non-invasive markers of prediction in children with autism spectrum disorder. *Gut*, gutjnl-2020-324015. doi: 10.1136/gutjnl-2020-324015.

Wang, H., Rogers, T. J., Paton, J. C., and Paton, A. W. (2014). Differential Effects of Escherichia coli Subtilase Cytotoxin and Shiga Toxin 2 on Chemokine and Proinflammatory Cytokine Expression in Human Macrophage, Colonic Epithelial, and Brain Microvascular Endothelial Cell Lines. *Infect Immun* 82, 3567–3579. doi: 10.1128/iai.02120-14.

Wang, J. (2012). Geometric Structure of High-Dimensional Data and Dimensionality Reduction. 151–180. doi: 10.1007/978-3-642-27497-8_8.

Wang, L., Chauliac, D., Moritz, B. E., Zhang, G., Ingram, L. O., and Shanmugam, K. T. (2019). Metabolic engineering of Escherichia coli for the production of butyric acid at high titer and productivity. *Biotechnol. Biofuels* 12, 62. doi: 10.1186/s13068-019-1408-9.

Wang, O., Liang, G., McAllister, T. A., Plastow, G., Stanford, K., Selinger, B., et al. (2016). Comparative Transcriptomic Analysis of Rectal Tissue from Beef Steers Revealed Reduced Host Immunity in Escherichia coli O157:H7 Super-Shedders. *Plos One* 11, e0151284. doi: 10.1371/journal.pone.0151284.

Wang, O., McAllister, T. A., Plastow, G., Stanford, K., Selinger, B., and Guan, L. L. (2017). Host mechanisms involved in cattle Escherichia coli O157 shedding: a fundamental understanding for reducing foodborne pathogen in food animal production. *Sci Rep-uk* 7, 7630. doi: 10.1038/s41598-017-06737-4.

Wang, O., McAllister, T. A., Plastow, G., Stanford, K., Selinger, B., and Guan, L. L. (2018). Interactions of the Hindgut Mucosa-Associated Microbiome with Its Host Regulate Shedding

of Escherichia coli O157:H7 by Cattle. *Appl Environ Microb* 84, e01738-17. doi: 10.1128/aem.01738-17.

Wang, O., Zhou, M., Chen, Y., McAllister, T. A., Plastow, G., Stanford, K., et al. (2021). MicroRNAomes of Cattle Intestinal Tissues Revealed Possible miRNA Regulated Mechanisms Involved in Escherichia coli O157 Fecal Shedding. *Front Cell Infect Mi* 11, 634505. doi: 10.3389/fcimb.2021.634505.

Waters, J. R., Sharp, J. C. M., and Dev, V. J. (1994). Infection Caused by Escherichia coli O157:H7 in Alberta, Canada, and in Scotland: A Five-Year Review, 1987-1991. *Clin Infect Dis* 19, 834–843. doi: 10.1093/clinids/19.5.834.

Watts, D. J., and Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature* 393, 440–442. doi: 10.1038/30918.

Weinstein, D. L., Jackson, M. P., Samuel, J. E., Holmes, R. K., and O'Brien, A. D. (1988). Cloning and sequencing of a Shiga-like toxin type II variant from Escherichia coli strain responsible for edema disease of swine. *J Bacteriol* 170, 4223–4230. doi: 10.1128/jb.170.9.4223-4230.1988.

Welle, S., Bhatt, K., and Thornton, C. A. (2000). High-abundance mRNAs in human muscle: comparison between young and old. *J. Appl. Physiol.* 89, 297–304. doi: 10.1152/jappl.2000.89.1.297.

Wells, J. G., Davis, B. R., Wachsmuth, I. K., Riley, L. W., Remis, R. S., Sokolow, R., et al. (1983). Laboratory investigation of hemorrhagic colitis outbreaks associated with a rare Escherichia coli serotype. *J Clin Microbiol* 18, 512–520. doi: 10.1128/jcm.18.3.512-520.1983.

Werner, M. S., Sieriebriennikov, B., Prabh, N., Loschko, T., Lanz, C., and Sommer, R. J. (2018). Young genes have distinct gene structure, epigenetic profiles, and transcriptional regulation. *Genome Res.* 28, 1675–1687. doi: 10.1101/gr.234872.118.

Wieler, L. H., Vieler, E., Erpenstein, C., Schlapp, T., Steinrück, H., Bauerfeind, R., et al. (1996). Shiga toxin-producing Escherichia coli strains from bovines: association of adhesion with carriage of eae and other genes. *J. Clin. Microbiol.* 34, 2980–2984. doi: 10.1128/jcm.34.12.2980-2984.1996.

WILLIAMS, K. J., WARD, M. P., DHUNGYEL, O. P., and HALL, E. J. S. (2015). Risk factors for Escherichia coli O157 shedding and super-shedding by dairy heifers at pasture. *Epidemiol Infect* 143, 1004–1015. doi: 10.1017/s0950268814001630.

Wilson, D. F. (2017). Oxidative phosphorylation: regulation and role in cellular and tissue metabolism. *The J. Physiol.* 595, 7023–7038. doi: 10.1113/jp273839.

Wolfe, A. W. (1997). Social Network Analysis: Methods and Applications. *Am Ethnol* 24, 219–220. doi: 10.1525/ae.1997.24.1.219.

Wu, W., Logares, R., Huang, B., and Hsieh, C. (2017). Abundant and rare picoeukaryotic sub-communities present contrasting patterns in the epipelagic waters of marginal seas in the northwestern Pacific Ocean. *Environ Microbiol* 19, 287–300. doi: 10.1111/1462-2920.13606.

WU, Y., HINENOYA, A., TAGUCHI, T., NAGITA, A., SHIMA, K., TSUKAMOTO, T., et al. (2010). Distribution of Virulence Genes Related to Adhesins and Toxins in Shiga Toxin-Producing Escherichia coli Strains Isolated from Healthy Cattle and Diarrheal Patients in Japan. *J Vet Med Sci* 72, 589–597. doi: 10.1292/jvms.09-0557.

Wu, Y., Li, Y., Zhang, C., A, X., Wang, Y., Cui, W., et al. (2014). S100a8/a9 Released by CD11b+Gr1+ Neutrophils Activates Cardiac Fibroblasts to Initiate Angiotensin II–Induced Cardiac Inflammation and Injury. *Hypertension* 63, 1241–1250. doi: 10.1161/hypertensionaha.113.02843.

Xia, C., Braunstein, Z., Toomey, A. C., Zhong, J., and Rao, X. (2018). S100 Proteins As an Important Regulator of Macrophage Inflammation. *Frontiers Immunol.* 8, 1908. doi: 10.3389/fimmu.2017.01908.

Xia, Y. (2020). Chapter Eleven Correlation and association analyses in microbiome study integrating multiomics in health and disease. *Prog Mol Biol Transl* 171, 309–491. doi: 10.1016/bs.pmbts.2020.04.003.

Xu, M., Huang, Q., Xiong, Z., Liao, H., Lv, Z., Chen, W., et al. (2021). Distinct Responses of Rare and Abundant Microbial Taxa to In Situ Chemical Stabilization of Cadmium-Contaminated Soil. *Msystems* 6, e01040-21. doi: 10.1128/msystems.01040-21.

Xu, Q., Vandenkoornhuyse, P., Li, L., Guo, J., Zhu, C., Guo, S., et al. (2022). Microbial generalists and specialists differently contribute to the community diversity in farmland soils. *J Adv Res* 40, 17–27. doi: 10.1016/j.jare.2021.12.003.

Xu, Y., Dugat-Bony, E., Zaheer, R., Selinger, L., Barbieri, R., Munns, K., et al. (2014). Escherichia coli O157:H7 Super-Shedder and Non-Shedder Feedlot Steers Harbour Distinct Fecal Bacterial Communities. *Plos One* 9, e98115. doi: 10.1371/journal.pone.0098115.

Xue, Y., Du, M., Sheng, H., Hovde, C. J., and Zhu, M.-J. (2017). Escherichia coli O157:H7 suppresses host autophagy and promotes epithelial adhesion via Tir-mediated and cAMP-independent activation of protein kinase A. *Cell Death Discov* 3, 17055. doi: 10.1038/cddiscovery.2017.55.

Xun, W., Li, W., Xiong, W., Ren, Y., Liu, Y., Miao, Y., et al. (2019). Diversity-triggered deterministic bacterial assembly constrains community functions. *Nat Commun* 10, 3833. doi: 10.1038/s41467-019-11787-5.

Yoon, S. C., Windham, W. R., Ladely, S., Heitschmidt, G. W., Lawrence, K. C., Park, B., et al. (2013). Differentiation of big-six non-O157 Shiga-toxin producing Escherichia coli (STEC) on

spread plates of mixed cultures using hyperspectral imaging. *J Food Meas Charact* 7, 47–59. doi: 10.1007/s11694-013-9137-4.

Yu, J., Feng, Q., Wong, S. H., Zhang, D., Liang, Q. yi, Qin, Y., et al. (2017). Metagenomic analysis of faecal microbiome as a tool towards targeted non-invasive biomarkers for colorectal cancer. *Gut* 66, 70. doi: 10.1136/gutjnl-2015-309800.

Yu, Z., and Morrison, M. (2004). Improved extraction of PCR-quality community DNA from digesta and fecal samples. *Biotechniques* 36, 808–812. doi: 10.2144/04365st04.

Zaheer, R., Dugat-Bony, E., Holman, D., Cousteix, E., Xu, Y., Munns, K., et al. (2017). Changes in bacterial community composition of Escherichia coli O157:H7 super-shedder cattle occur in the lower intestine. *Plos One* 12, e0170050. doi: 10.1371/journal.pone.0170050.

Zhang, B., and Horvath, S. (2005). A General Framework for Weighted Gene Co-Expression Network Analysis. *Stat Appl Genet Mol* 4, Article17. doi: 10.2202/1544-6115.1128.

Zhang, Y., Wu, G., Jiang, H., Yang, J., She, W., Khan, I., et al. (2018). Abundant and Rare Microbial Biospheres Respond Differently to Environmental and Spatial Factors in Tibetan Hot Springs. *Front Microbiol* 9, 2096. doi: 10.3389/fmicb.2018.02096.

Zhou, H., Beltrán, J. F., and Brito, I. L. (2022). Host-microbiome protein-protein interactions capture disease-relevant pathways. *Genome Biol* 23, 72. doi: 10.1186/s13059-022-02643-9.

Zhou, J., Deng, Y., Luo, F., He, Z., Tu, Q., and Zhi, X. (2010). Functional Molecular Ecological Networks. *Mbio* 1, e00169-10. doi: 10.1128/mbio.00169-10.

Zhou, J., Deng, Y., Luo, F., He, Z., and Yang, Y. (2011). Phylogenetic Molecular Ecological Network of Soil Microbial Communities in Response to Elevated CO2. *Mbio* 2, e00122-11. doi: 10.1128/mbio.00122-11.

Zhou, J., Liu, W., Deng, Y., Jiang, Y.-H., Xue, K., He, Z., et al. (2013). Stochastic Assembly Leads to Alternative Communities with Distinct Functions in a Bioreactor Microbial Community. *Mbio* 4, e00584-12. doi: 10.1128/mbio.00584-12.

Zhou, J., and Ning, D. (2017). Stochastic Community Assembly: Does It Matter in Microbial Ecology? *Microbiol Mol Biol R* 81, e00002-17. doi: 10.1128/mmbr.00002-17.

Zhu, M., Qi, X., Yuan, Y., Zhou, H., Rong, X., Dang, Z., et al. (2023). Deciphering the distinct successional patterns and potential roles of abundant and rare microbial taxa of urban riverine plastisphere. *J Hazard Mater* 450, 131080. doi: 10.1016/j.jhazmat.2023.131080.

Zhunina, O. A., Yabbarov, N. G., Grechko, A. V., Starodubova, A. V., Ivanova, E., Nikiforov, N. G., et al. (2021). The Role of Mitochondrial Dysfunction in Vascular Disease, Tumorigenesis, and Diabetes. *Frontiers Mol. Biosci.* 8, 671908. doi: 10.3389/fmolb.2021.671908.

Ziegler, M., Eguíluz, V. M., Duarte, C. M., and Voolstra, C. R. (2018). Rare symbionts may contribute to the resilience of coral–algal assemblages. *Isme J* 12, 161–172. doi: 10.1038/ismej.2017.151.