Identification of Hypermutator Enterohemorrhagic *Escherishia coli*
(EHEC) Using a High Throughput Screening Method to Inform Food
Safety Investigations

by

Alexa McCarthy

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs in partial
fulfillment of the requirements for the degree of

Master of Science

in

Biology

Carleton University
Ottawa, Ontario

© 2020, Alexa McCarthy

Abstract

Enterohemorrhagic *E. coli* (EHEC) are one of the top bacteria causing foodborne illnesses in Canada. During an outbreak investigation, bacterial clusters associated with high-mutators may be missed due to their ability to accumulate mutations. This study compared three approaches for mutation rate estimation: fluctuation assay, mutation accumulation, and a high-throughput screening assay based on the fluctuation assay. The screening assay successfully detected a known high-mutator and was then used to screen 100 verotoxigenic strains for high mutation rates. Two high-mutators were successfully detected, indicating that although the phenotype is rare, it is still detectable. Whole-genome sequences of these strains were investigated for genes associated with high mutation rates. The two high-mutators were found to be lacking the *mutS* gene, known to be associated with hypermutators. Rapid identification of high-mutators would be valuable to inform whole-genome sequencing data and aid in the determination of the scope of a foodborne investigation.

Acknowledgements

I would like to thank Dr. Catherine Carrillo, Dr. Alex Wong, and Dr. Burton Blais for giving me the necessary guidance and support needed to complete my thesis. I am forever appreciative of the opportunity to work with them, which has helped me learn and grow into a worthy member of the scientific community. I gratefully acknowledge the technical support provided by Paul Manninger and Susan Van Zanten, as well as assistance with bioinformatic analysis provided by Andrew Low and Dr. Adam Koziol. Thank you to Dr. Daniel Plante at Health Canada for his statistical expertise, and to Dr. Catherine Carrillo and Dr. Alex Wong for the critical review of my manuscript. I gratefully acknowledge Martine Dixon, Mylene Deschênes, as well as my fellow students in the lab for helpful input and continuous encouragement. Lastly, I would like to gratefully acknowledge Dr. Martha Mullally and Dr. Min Lin for being part of my committee and giving me additional guidance in my project.

Table of Contents

Abstract	ii
Acknowledgements	iii
Table of Contents	iv
List of Tables	vii
List of Figures	viii
List of Appendices	ix
Preface	x
Chapter 1: Introduction	1
1.1. Foodborne Illness in Canada	1
1.1.1. Salmonella enterica	2
1.1.2. Listeria monocytogenes	3
1.1.3. Escherichia coli	3
1.1.3.1. Enterohemorrhagic <i>E. coli</i> (EHEC)	4
1.1.3.2. Clinical Relevance of EHEC Serotypes	5
1.1.4. Hemolytic Uremic Syndrome	7
1.2. The Structure of an Outbreak Investigation	8
1.2.1. The Foodborne Illness Outbreak Response Protocol (FIORP)	8
1.3. Current Genotyping Methods	9
1.3.1. Pulse-Field Gel Electrophoresis (PFGE)	10
1.3.2. Multi-locus Variable Tandem Repeat Analysis (MLVA)	11
1.4. Foodborne Disease Outbreak Detection Based on Whole-Genome Sequencing	11
1.4.1. Post-Sequencing Strain Discrimination	12
1.4.2. Data Interpretation	14
1.4.3. Whole-Genome Sequencing (WGS) Shortcomings	14
1.5. Mutation Rates	15
1.5.1. Single Nucleotide Variants (SNVs)	15
1.5.2. Mutation Rates	17
1.6. Factors Affecting Mutation Rates	17
1.7. Methods of Mutation Rate Estimation	18

1.7.1.	Mutation Accumulation (MA) & Whole-Genome Sequencing (WGS).....	18
1.7.2.	Fluctuation Assay.....	19
1.7.2.1.	Nalidixic Acid Resistance	20
1.8.	Experimental Aim.....	21
Chapter 2: Materials & Methods		22
2.1.	Bacterial Strains and Growth Conditions	22
2.1.1.	Nalidixic Acid Plates.....	23
2.2.	Mutation Accumulation Experiment	23
2.2.1.	Estimation of Generations in Colonies.....	24
2.3.	Whole-Genome Sequencing	25
2.3.1.	DNA Isolation, Sequencing, and Assembly.....	25
2.3.2.	SNV Analysis.....	26
2.4.	Fluctuation Assay	26
2.4.1.	Cell Enumeration.....	27
2.5.	Screening Assay	27
2.6.	Agar Microdilution	29
2.7.	Estimation of Mutation Rates	31
2.7.1.	Mutation Accumulation.....	31
2.7.2.	Fluctuation Assay	32
2.8.	Mutator Allele Identification	32
Chapter 3: Results		33
3.1.	Determination of Mutation Rate in <i>E. coli</i>	33
3.1.1.	Fluctuation Assay	33
3.1.2.	Mutation Accumulation & Whole-Genome Sequencing	36
3.2.	Screening Assay	42
3.3.	Screening 100 VTEC Strains.....	44
3.3.1.	Agar Microdilution.....	45
3.4.	Mutator Alleles	47
Chapter 4: Discussion.....		48
4.1.	<i>E. coli</i> Mutation Rate Determination.....	48
4.2.	High Throughput Screening Assay.....	51

4.3.	Identification of Genes Associated with High Mutation Rate.....	53
4.4.	Other Factors Affecting Mutation Rate	56
4.5.	Mutation Detection in The Context of Food Safety Investigation	57
4.6.	Limitations & Future Directions.....	58
Appendices.....		61
References.....		82

List of Tables

Table 1. Antibiotic dilution chart for agar microdilution assay	31
Table 2. The mutation rate for <i>E. coli</i> strains O111:NM (OLC 0682), O157:H7 (OLC 0809), and O103:H2 (OLC 0969) as determined by three fluctuation assay replicates.	35
Table 3. Parameters of the mutation accumulation and whole-genome sequencing experiments for <i>E. coli</i> strains O111:NM (OLC 0682), O157:H7 (OLC 0809), and O103:H2 (OLC 0969). ..	38
Table 4. Comparison of the average mutation rate for strains O111:NM (OLC 0682), O157:H7 (OLC 0809), and O103:H2 (OLC 0969) as calculated by a fluctuation assay and a mutation accumulation experiment.	38
Table 5. Parameters of the mutation accumulation and whole-genome sequence experiments compared to the percentage of wells containing viable mutants identified by the screening assay for strains O165:H25 (OLC 0787) and O22:H8 (OLC 0984).....	45
Table 6. Minimal inhibitory concentrations (MICs) of nalidixic acid and rifampicin for resistant and susceptible variants of <i>E. coli</i> strains O26:H11 (OLC 0683), O113:H21 (OLC 0686), O45:H2 (OLC 1020), a susceptible positive control (OLC 1614), and a negative control containing sterile NaCl.	46

List of Figures

Figure 1. Schematic demonstrating the workflow for a mutation accumulation assay	24
Figure 2. Schematic demonstrating the workflow for a high throughput screening assay used to detect strains with possible high mutation rates	29
Figure 3. Percentage of plates demonstrating growth of 10 or more viable mutants (out of 30) $\pm$ standard error for three fluctuation assay replicates	35
Figure 4. Mutation rates as estimated by the average of three fluctuation assays $\pm$ 95% confidence limits	36
Figure 5. Maximum likelihood phylogenetic tree of <i>E. coli</i> O111:NM (OLC0682).....	40
Figure 6. Maximum likelihood phylogenetic tree of <i>E. coli</i> O157:H7 (OLC 0809)	41
Figure 7. Maximum likelihood phylogenetic tree of <i>E. coli</i> O103:H2 (OLC0969)	42
Figure 8. Percentage of wells containing viable mutants $\pm$ standard error for three replicates of the screening method for strains O111:NM (OLC 0682), O157:H7 (OLC 0809), and O103:H2 (OLC 0969).....	44

List of Appendices

Appendix A: List of Strains Screened Using Screening Assay.....	61
Appendix B: Screening Assay Protocol	65
Appendix C: Fluctuation Assay Culture Enumeration & Viable Mutant Count.....	67
Appendix D: Raw Data for Screening Assay.....	75
D.1 Raw Data for Screening Assay with 96 Wells.....	75
D.2 Raw Data for Screening Assay with 24 Wells.....	76
Appendix E: Raw Data for Screening Assay of 100 VTEC Strains.....	77
Appendix F: Quality Metrics for WGS Sequences.....	81

Preface

Whole-genome sequencing is now being used by public agencies worldwide to detect clusters of related strains, potentially involved in outbreaks. Presence of a high mutator strain could impact cluster detection. Currently there are no methods to detect high mutators within the time frame of a food outbreak. To address this, we wanted to develop laboratory and genomic approaches for rapidly identifying high mutators. This was completed by first comparing two methods of mutation rate estimation: (1) the fluctuation assay and (2) the mutation accumulation assay. A high-throughput screening assay derived from the fluctuation assay was then developed and used to screen a panel of 100 verotoxigenic *E. coli* strains for high mutators. Upon the detection of high mutators, genes and alleles involved in mutation rate were investigated.

All the laboratory experiments and DNA extraction and quantification were completed by me. Whole-genome sequencing was completed by Paul Manninger, Ashley Cooper, Susan Van Zanten, and Bridgette Kelly at The Canadian Food Inspection Agency (CFIA), Ottawa Laboratory Carling (OLC). Adam Koziol assisted with bioinformatics analyses (AlleleFinder), protein sequence alignment (CLUSTAL) was performed by Dr. Catherine Carrillo, SNV analysis (SNVPhyl) and interpretation of the results was done by me with assistance from Dr. Catherine Carrillo.

Results of this thesis are currently being prepared for publication.

Chapter 1: Introduction

1.1. Foodborne Illness in Canada

Every year 1 in 8 Canadians are affected by a foodborne illness. Of these cases, there are approximately 11,800 hospitalizations and 238 deaths (PHAC, 2016a). Foodborne illness can be caused by several pathogens including bacteria, viruses, and parasites. The most common causes are norovirus, *Listeria monocytogenes*, *Salmonella enterica*, enterohemorrhagic *Escherichia coli* (EHEC) -often O157:H7- and *Campylobacter* spp. (PHAC, 2016a). Contamination can be found in raw foods, such as meat, poultry, seafood, eggs, dairy products, and fresh produce (National Institute of Health, 2014). Food can be contaminated during various stages of production, such as growth, harvesting, slaughtering, storage, processing, and during food preparation (National Institute of Health, 2014). Common symptoms of such foodborne illnesses include fever, vomiting, diarrhea, abdominal pain (National Institute of Health, 2014; PHAC 2016a), and in more severe cases of illness, hemolytic uremic syndrome (HUS) and/or death.

Tracking foodborne illness across the country is a top priority in maintaining public health. To achieve this, Canada has several surveillance systems that are used to detect and monitor outbreaks and identify risk factors with the help of local, provincial, and national public-health laboratories. These include the Canadian Notifiable Disease Surveillance system (Totten *et al.*, 2019), which collects the number of confirmed cases annually, the National Enteric Surveillance Program (Nesbitt *et al.*, 2012), which collects weekly numbers at the species level, FoodNet Canada (PHAC, 2018), which collects information on enteric illness and sources of exposure in communities throughout Canada, National Studies on Acute Gastrointestinal Illnesses (Thomas *et al.*, 2017), which gathers information from the population regarding enteric

illness symptoms, PulseNet Canada (Nadon *et al.*, 2017), which is a subtyping network that collects information on foodborne disease outbreaks, amongst many others.

1.1.1. *Salmonella enterica*

Salmonella enterica are rod-shaped, gram-negative bacteria frequently associated with serious illness in humans and animals. Non-typhoidal *Salmonella* are responsible for 24% of hospitalizations and 16% of deaths attributed to foodborne illness across Canada (PHAC, 2016a). Infection caused by *Salmonella* results in salmonellosis which is characterized by abdominal cramps, diarrhea, vomiting, and fever (PHAC, 2016b). The preferred niche for these organisms is the gastrointestinal tract of humans, as well as animals such as reptiles, birds, amphibians, cats, dogs, and even insects (PHAC, 2016b). While norovirus is the leading cause of foodborne illness in Canada, *Salmonella* is responsible for 1 in 4 foodborne-related illnesses each year, with the highest number of outbreaks being related to raw poultry (CDC, 2013; Morton *et al.*, 2019; PHAC, 2020). More than 2,500 salmonella serotypes have been discovered, however, less than 100 cause illness in humans (CDC, 2020). Salmonellosis is largely caused by two serotypes; Enteritidis, which was responsible for 56% of foodborne illness in Canada in 2017, and Typhimurium, which was responsible for 7% of illnesses (Herikstad *et al.*, 2002; Jones *et al.*, 2008; PHAC, 2017; Morton *et al.*, 2019). In 2018, however, it was reported that Enteritidis was responsible for 43% of foodborne illness in Canada, followed by Heidelberg with 9%, and Typhimurium with 6% (PHAC, 2019). Questions remain regarding the relationship between serotypes and the severity of the illness. Some factors that may be involved include plasmids containing virulence genes, pathogenicity islands, and surface cell structure (Andino & Hanning, 2015).

1.1.2. *Listeria monocytogenes*

Listeria monocytogenes are gram-positive bacteria responsible for Listeriosis, a very serious illness in humans. *Listeria* is the leading cause of death attributed foodborne illness nationwide, however, is only responsible for 178 illnesses, and 150 hospitalizations every year (PHAC, 2016a). The most common outcome of listeriosis is mild gastrointestinal symptoms, similar to that of other foodborne illness, such as salmonellosis. More serious symptoms of listeriosis include seizures, meningitis, ataxia, high fever, pneumoniae, and fetal death in pregnant women (PHAC, 2016c; Pizarro-Cerdá & Cossart, 2019). The reason for such severe illness is due to its ability to traverse several of the body's barriers, which are supposed to protect against pathogens of this nature. Such barriers include the blood-brain barrier, the placental barrier, and the intestinal barrier (Pizarro-Cerdá & Cossart, 2019). All 13 *L. monocytogenes* serotypes have the potential to cause listeriosis, however, the most common serotypes to cause illness in humans and animals are serotypes 4b, 1/2b and 1/2a (Ward *et al.*, 2004; PHAC, 2016c; Shamloo *et al.*, 2019). Recent discoveries have shed light on pathogenicity by demonstrating that certain clonal complexes are associated with different clinical and foodborne isolates. Four clonal complexes have been identified within 4b, and two have been identified in 1/2b and 1/2a (Maury *et al.*, 2016).

1.1.3. *Escherichia coli*

Escherichia coli are symbiotic organisms that often reside in the gastrointestinal tract of mammals. The relationship is characterized by the mammal offering the nutrients and niches needed for survival, while the bacteria of the gut (microbiota) contribute to nutrient metabolism and aids in immune system maintenance by protecting against the colonization of pathogens to the intestinal wall (Kau *et al.*, 2011; Belkaid & Hand, 2014). While *E. coli* can be beneficial to

our health, some acquire virulence factors resulting in different pathotypes. There are still challenges that remain when determining the genomic changes responsible for a strain to be pathogenic. Some factors speculated to be involved in the process include genomic rearrangement, mutation accumulation in the genome, and horizontal gene transfer (Tenaillon *et al.*, 2010).

1.1.3.1. Enterohemorrhagic *E. coli* (EHEC)

Enterohemorrhagic *E. coli* (EHEC) (also referred to as verotoxigenic *E. coli* (VTEC) or shiga toxin-producing *E. coli* (STEC)) are the sixth most common cause for reported foodborne illnesses in Canada (PHAC, 2016a), where they are responsible for 6% of hospitalizations and 8% of deaths attributed to foodborne pathogens (PHAC, 2016a). The degree of pathogenicity of these virulent strains is determined by taking into account the following; (1) The infectious dose, where typically less than 100 cells is considered a low infectious dose (Nguyen & Sperandio, 2012; Pacheco & Sperandio, 2012), (2) risk of severe disease development and possible fatality (Gill, 2018; PHAC 2016a), and (3) treatability of infection (Karpman & Stahl, 2014; Pacheco & Sperandio, 2012). Shiga toxins (Stx), also called verotoxins (VT) interchangeably, are potent toxins of the AB5 family found in *Shigella dysenteriae* and among many *E. coli* serogroups. They act by cleaving a specific adenosine residue found in the 28S RNA of the 60S ribosomal subunit, resulting in protein synthesis termination (Melton-Celsa, 2014). *Stx1* and *Stx2* are the two toxins associated with *E. coli* strains (Nakao & Takeda, 2000), *Stx2* being epidemiologically associated with more severe disease conditions, such as HUS (Friedrich *et al.*, 2002; Karch, 2001; Karch *et al.*, 2006; Melton-Celsa, 2014). In addition to these verotoxins and their derivatives, many other virulence factors contribute to the pathogenicity of an isolate. Binding factors, such as the protein intimin, and aggregative regulators, such as that encoded by *aggR*,

are common virulence factors associated with EHEC. Intimin is encoded by the chromosomal gene *eae* and is responsible for the adherence of EHEC to the intestinal epithelial cells, which leads to possible lesions in the intestinal mucosa (Blanco *et al.*, 2003; Fan *et al.*, 2019). *aggR* is responsible for the regulation of virulence factors, such as genes involved in toxin production itself, as well as aggregative adherence fimbriae biogenesis (Elias *et al.*, 1999; Morin *et al.*, 2013). These virulence factors have been implicated in illness related to foodborne outbreaks worldwide (Beutin *et al.*, 2004; CDC, 2013; Hebbelstrup Jensen *et al.*, 2014; Kinnula *et al.*, 2018).

1.1.3.2. Clinical Relevance of EHEC Serotypes

Serotyping is a traditional method used to type bacterial species based on distinct variations involving their surface antigens. These antigens include the O- polysaccharide antigens, flagellar H-antigens, and capsular K-antigens (Fratamico *et al.*, 2016). Certain serotypes have been proposed as a factor involved in the likelihood of illness development. *E. coli* have an outer membrane that consists of lipopolysaccharides (LPS), including lipid A, core oligosaccharides, and the O-antigen (Lugtenberg, 1981). Currently, there are upward of 180 identified O-groups, with more being discovered to this day (Lang *et al.*, 2019). In addition, there are 53 H-groups (Fratamico *et al.*, 2016), and over 80 K-groups (Yang *et al.*, 2018). The O antigen is attached to the core oligosaccharide, making it the outermost portion of the LPS. This surface antigen is responsible for the serotype diversity amongst *E. coli*, making it clinically important for the identification and classification of strains. The two genes that are targeted for O-antigen serotyping are *wzx*, encoding the O antigen flippase, and *wzy*, encoding the O antigen polymerase (DebRoy *et al.*, 2016). The presence or absence of this antigen determines whether it is penetrable by hydrophobic antibiotics (Tsujimoto *et al.*, 1999), which can be considered

during the clinical treatment of pertaining strains. The H-antigen is encoded by the *fliC* gene and is an important aspect of the flagella, which is involved in *E. coli* mobility (Wang *et al.*, 2003). The absence of this antigen results in a non-motile strain of *E. coli*. The flagella are important for invasion of and, in some cases, adhesion to the intestinal wall which is a critical step in colonization and survival (Haiko & Westerlund-Wikström, 2013).

The K-antigen is found on the surface of the bacterium, under a thick, mucous-like capsule made of acidic polysaccharides (Whitfield, 2006). The importance of this molecule lies within the protection that it serves from environmental stress and desiccation (Yang *et al.*, 2018). These antigens can largely be subdivided into four groups; groups 1 and 4 K-antigens can sometimes be attached to the lipid-A core of the LPS and are only found with certain O-antigens (such as O8, O9, O20, and O101) (Whitfield, 2006). Group 2 and 3 K-antigens resemble those found on gram-positive bacteria and can be subdivided based on acidic constituents (Yang *et al.*, 2018). These antigens have been associated with extraintestinal infections (Whitfield, 2006). Different biosynthesis pathways and genes are required for different K-antigen groups and antigen molecules. More information on these pathways and gene clusters is outlined in Yang *et al.* (2018). Since there are limited capabilities to type the K-antigens in the laboratory, the use of the H- and O-antigen are the antigens of choice in serotyping (Banjo *et al.*, 2018; Fratamico *et al.*, 2016; Prager *et al.*, 2003). The famous EHEC in food safety is *E. coli* O157, as it has been the focus of numerous studies (Gao *et al.*, 2018; Pianciola & Rivas, 2018; Li *et al.*, 2017) and has been associated with numerous outbreaks (Gyles, 2007; FAO/WHO STEC EXPERT GROUP, 2019; Hussein & Omaye, 2003; Currie *et al.*, 2018; Currie *et al.*, 2019; Gaulin *et al.*, 2015; Tataryn *et al.*, 2014), along with serotypes O111 and O103 (Woodward *et al.*, 2002; Hussein & Omaye, 2003; Hussein, 2007; Gill & Gill, 2010).

1.1.4. Hemolytic Uremic Syndrome

A worrisome condition that can result from foodborne disease outbreaks is hemolytic uremic syndrome (HUS) which belongs to a group of diseases known as thrombotic microangiopathies. These diseases are characterized by conditions affecting the small blood vessels of the body (Sheerin & Glover, 2019). The initial symptoms of HUS include bloody diarrhea, vomiting, fever, and general weakness. The disease can then progress to more severe symptoms, which includes the destruction of red blood cells and platelets, anemia, and kidney failure (as a result of the destruction of the small blood vessels in the kidney) (Fakhouri *et al.*, 2017; Karpman & Stahl, 2014; Karpman *et al.*, 2017). Future complications can include heart failure and neurological issues including encephalopathy (brain dysfunction), seizures, paralysis, and in severe cases, coma, and death. Since HUS can emerge from different origins, it can either be referred to as typical HUS (HUS) or atypical HUS (aHUS). Typical HUS (commonly referred to, simply, as HUS) is caused by bacterial pathogens, which is the most common of the two, while aHUS is only responsible for 5-10% of cases (Norris & Remuzzi, 2009). Rather than being caused by pathogens, aHUS is characterized by complement dysregulation caused by mutations or autoantibodies (Jokiranta, 2017).

Typical HUS is commonly caused by EHEC, as well as non-typhoidal *Salmonella*, *Shigella*, and *Streptococcus pneumoniae*. While *E. coli* O157:H7 is the most common EHEC implicated in HUS (Bell *et al.*, 1994; Center for Disease Control and Prevention, 1996; Shefer *et al.*, 1996; Rowe *et al.*, 1998; Karmali *et al.*, 2009; Wendel *et al.*, 2009; Bustamante *et al.*, 2012; Salvadori & Bertoni, 2013; Canpolat, 2015; Lisboa *et al.*, 2019), non-O157 EHEC have also been implicated around the globe (Buchholz *et al.*, 2011; Coombes *et al.*, 2011; Couturier *et al.*, 2011; Bustamante *et al.*, 2012; Wasilenko *et al.*, 2014), as well as non-EHEC including

Salmonella spp. (Helms *et al.*, 2006; Ternhag *et al.*, 2008; Keithlin *et al.*, 2015; Lisboa *et al.*, 2019; Joseph *et al.*, 2020), *Shigella* spp. (Taylor, 2008; Canpolat, 2015; Lamba *et al.*, 2016; Adams *et al.*, 2017), and *Streptococcus pneumoniae* (Vanderkooi *et al.*, 2003; Copelovitch & Kaplan, 2010; Spinale *et al.*, 2013; Weintraub *et al.*, 2014; Cody & Dixon, 2019). The possibility of developing HUS varies depending on the strain associated with the illness; the rate of reported HUS development following an enteric illness is between 2 and 22% (Vallis *et al.*, 2018). This possibly fatal illness is one of many reasons attributing to the importance of timely outbreak investigations.

1.2. The Structure of an Outbreak Investigation

A foodborne outbreak is defined as an event in which several people have developed a foodborne illness with similar symptoms from the same contaminated food. Public health surveillance is used to monitor and detect clusters, which are defined as a collection of similar health events. These events are usually compiled over some time or across a geographical area, using formal and informal illness reports, in addition to resources such as PulseNet (Nadon *et al.*, 2017). An investigation is launched upon detection of human enteric illnesses thought to be caused by a common source, or detection of a hazard that could potentially cause human enteric illness.

1.2.1. The Foodborne Illness Outbreak Response Protocol (FIORP)

The Public Health Agency of Canada, Health Canada, and the Canadian Food Inspection Agency have collectively produced a protocol to follow in these cases, The Foodborne Illness Outbreak Response Protocol (FIORP) (Public Health Agency of Canada, 2017). The purpose of the FIORP is to promote coordination between the three government agencies for a prompt response to an outbreak occurring in more than one province. Upon the identification of a

potential threat, all partners involved in outbreak management are notified. The compiled information involving these health events is reviewed by the Outbreak Investigation Coordinating Committee (OICC), who then collectively determine if the investigation will progress. In the case where further measures are initiated, investigations of food safety, laboratory, and epidemiological nature are coordinated to detect a potential source and cause of the outbreak. The OICC compiles the findings of these basal investigations to collectively decide the need for further investigation and to determine the level of health risk that this outbreak poses to the public. Public health actions are then implemented to manage the outbreak and prevent future outbreaks of the same nature. This often includes actions such as, recalling and disposing of the contaminated food product, inspection and sanitation of the facility where the outbreak originated, communication with the public to ensure the control and prevention of possible enteric illness associated with the outbreak, amongst others. The OICC monitors the progression of the outbreak until the number of reported outbreak-related illnesses have returned to a baseline level, which then signifies the conclusion of the outbreak. The final step pertains to the post-outbreak implementations. The OICC releases a post-outbreak debrief outlining specific goals to allow transparency with the public. These goals can include the underlying cause of the outbreak, policy revision for recurrence prevention, assessment of the control measures taken following the outbreak, among others. More in-depth information on the structure of an outbreak investigation can be found in the FIORP (Public Health Agency of Canada, 2017).

1.3. Current Genotyping Methods

Strain identification and characterization is a crucial step in recognizing the cause of an outbreak. This is commonly done through genotyping methods, which operate by examining the DNA sequence of a biological organism and comparing it to another. There are three broad

categories of bacterial genotyping: DNA banding pattern via digestion and/or amplification, DNA hybridization with known probes, and DNA sequencing. Several methods can classify bacteria based on fragment size generated by PCR amplification or DNA digestion by restriction enzymes (Li *et al.*, 2009). Historically, the most common methods for detecting clusters in foodborne outbreak surveillance were pulsed-field gel electrophoresis (PFGE) and multiple-locus variable-number tandem repeat analysis (MLVA). Now, with technological advances, whole-genome sequencing (WGS) is the leading method.

1.3.1. Pulse-Field Gel Electrophoresis (PFGE)

PFGE is a method that separates DNA fragments using a constant electrical pulse with the rationale that larger fragments will move through the gel at a slower pace whereas smaller fragments, due to their size, will be able to move through the gel at a faster rate (Centers for Disease Control and Prevention, 2016a). Bacterial cells are first combined with melted agarose, which is then poured into a plug mold. The cells are lysed, allowing the DNA to be free-floating in the agarose. Restriction enzymes are used to cut the DNA, which is then loaded into a gel for the rest of the PFGE process to occur. PFGE differs from standard gel electrophoresis because instead of running a constant voltage in one direction, it is run in three different directions periodically. This is because PFGE is typically used for larger DNA fragments that reach a threshold where they all run at the same rate, making them appear as one large band. The change of voltage direction allows these bands to appear separate on the gel depending on the differing rates at which the realignment of their charges occurs. The resulting gel is then stained and photographed under ultraviolet light.

1.3.2. Multi-locus Variable Tandem Repeat Analysis (MLVA)

MLVA is another method used by the CDC (Centers for Disease Control and Prevention, 2016b) that combines PCR and bands of repetitive sequences. The DNA of the bacterial cells is extracted (usually via a boiling method), then several loci are carefully selected to be amplified by PCR, considering the mutation rate and diversity of the loci selected. The sizes of the amplified loci are determined by using capillary electrophoresis, which runs an electric supply through a gel matrix that contains the PCR products. An electropherogram (MLVA data output graph) then demonstrates the comparison of the DNA standards with the PCR products. Using computer software, the PCR products are converted to allele types which determines the evolutionary relationship between the isolates.

1.4. Foodborne Disease Outbreak Detection Based on Whole-Genome Sequencing

Whole-genome sequencing (WGS) has been a disruptive technology in the science community that provides the entire genetic blueprint of organisms. This method has recently been implemented in the monitoring and detection of foodborne infections in several countries (Holmes *et al.*, 2015; Butcher *et al.*, 2016; Ferdus *et al.*, 2016; Crowe *et al.*, 2017; Rumore *et al.*, 2018; Jenkins *et al.*, 2019). Currently, there are two main networks used to monitor foodborne infections: PulseNet and GenomeTrakr. PulseNet is an international network that uses molecular subtyping for foodborne infection surveillance and includes 86 countries (Swaminathan *et al.*, 2001). In addition to contributing to the international network, each country has its own network for national monitorization. In its infancy, PulseNet used the subtyping method PFGE and later added MLVA as a method (Timme *et al.*, 2019). Although widely used, PFGE is not consistent in demonstrating optimal discrimination (particularly for highly clonal strains) (Holmes *et al.*, 2015; Rumore *et al.*, 2018). To enhance resolution, MLVA was implemented and has been

successful in discriminating closely related isolates which would not be distinguished by PFGE (Noller *et al.*, 2003; Keys & Keim, 2005; Hyytiä-Trees *et al.*, 2006). WGS was later implemented for the surveillance of *Listeria* and *Campylobacter* in 2014 and of EHEC and *Salmonella* in 2019 in the United States, and has been increasingly implemented in Canada (Nadon *et al.*, 2017; Rumore *et al.*, 2018; Brown *et al.*, 2019). This method of detection and investigation of pathogenic bacteria has added accuracy in surveillance and thus acceleration and efficiency in the decision-making response to outbreak investigations (Brown *et al.*, 2019). In 2013, the FDA established GenomeTrakr, a global, publicly available network containing the genetic makeup of thousands of foodborne disease-causing bacteria. It consists of the genomic and geographic data from foodborne pathogens shared by public health and university laboratories and is available in public databases at the National Center for Biotechnology Information (NCBI).

WGS analyses have been conducted on foodborne pathogens recovered by the Canadian Food Inspection Agency (CFIA) since 2014 and is now replacing several traditional analyses for bacterial identification and characterization since it can be used to serotype, detect virulence, and conduct antimicrobial resistance profiling (Lambert *et al.*, 2015; Carrillo *et al.*, 2016; Cooper *et al.*, 2020). Not only does it provide a level of precision and strain discrimination that is better than traditional subtyping technologies, but the turnaround time is days compared to weeks.

1.4.1. Post-Sequencing Strain Discrimination

High-resolution strain discrimination can be achieved by two methods of post-sequencing analysis: (1) Single Nucleotide Variant (SNV) analysis, where the isolates are compared to a closely related reference, and (2) multilocus sequence typing (MLST), which works by assessing sequence variations in the coding region of genes. Following read mapping to a

reference genome, the basis for SNV-based genotyping involves the identification of variants for each isolate, which are then combined into a population-wide SNV matrix. The identified SNVs from the matrix are then used for the construction of a phylogenetic tree. This can all be accomplished using a pipeline, such as SNVPhyl (Petkau *et al.*, 2017). MLST, on the other hand, has three levels of discrimination applied by PulseNet; (1) 7-gene MLST, which focuses on seven housekeeping genes, (2) core genome MLST (cgMLST) which focuses on genes present in nearly all strains of the species, and (3) whole-genome MLST (wgMLST) which focuses on nearly all genes present (Nadon *et al.*, 2017). wgMLST provides the highest discriminatory power, making it the most useful for cluster investigations. In contrast to SNV analysis, wgMLST is based on allelic variation rather than single base mutations, which signifies that the idea of a single evolutionary event includes recombinations and indels of multiple positions. In comparison, SNV analysis is more discriminatory because of its ability to assess variations in noncoding regions (Brown *et al.*, 2019). This method is preferentially used by the FDA and GenomeTrakr because it is deemed more fit to identify products associated with an outbreak (Timme *et al.*, 2018). On the other hand, wgMLST is more computationally available and less bioinformatically demanding, especially for inexperienced scientists (Brown *et al.*, 2019). Because of this, PulseNet and the CDC use MLST as their primary method but use SNV analysis for confirmation after inconclusive results or more complex outbreaks (Ribot *et al.*, 2019). A WGS pilot project by Brown *et al.* (2019) was conducted in parallel with PFGE-based surveillance, which found that WGS provides higher resolution and more precision than PFGE. This suggests that more outbreaks could be detected, investigated, and controlled in a timelier manner. Other studies comparing PFGE and MLVA with WGS for EHEC typing concluded that WGS was as sensitive as MLVA in respect to detecting outbreaks, and was significantly faster in

respect to clustering (Joensen *et al.*, 2014; Holmes *et al.*, 2015; Rusconi *et al.*, 2016); making it the superior technique.

1.4.2. Data Interpretation

During sequence data interpretation, it is not only crucial to ensure that the results make epidemiological sense, but also to define a SNV/allele cut-off which determines which organisms are closely related. There are still questions that remain regarding how many SNV differences are found among clusters of epidemiologically related isolates. Federal agencies typically suggest that isolates are closely related if they contain 0-20 SNVs (Brown *et al.*, 2019) however, many studies demonstrate a tentative cutoff of ≤ 10 SNVs (Underwood *et al.*, 2013; Joensen *et al.*, 2014; Dallman *et al.*, 2015; Butcher *et al.*, 2016; Holmes *et al.*, 2015; Rumore *et al.*, 2018). The only issue with this rough cutoff is that it ignores the fact that some strains can be high-mutators, and accumulate more than 10 SNVs while still being epidemiologically-related, which, for an outbreak investigation, would mean that these isolates would be excluded from the cluster as they appear to be unrelated.

1.4.3. Whole-Genome Sequencing (WGS) Shortcomings

While WGS has provided many advances in foodborne outbreak surveillance, it contains some minor disadvantages. It is not currently possible to determine MLVA types from WGS with the use of short read lengths (Holmes *et al.*, 2015), which is commonly used. However, with advances in technology such as single-molecule real-time sequencing and Oxford Nanopore Technologies (ONT) sequencing, longer read lengths capable of spanning tandem repeat regions are provided. With these longer read lengths, MLVA type information can be obtained to define the evolutionary relationships of the isolates.

Converting WGS data into an interpreted output that is useful for public health professionals has also been deemed a challenge. This is because most of these individuals do not possess the required bioinformatics background to interpret the output data. To overcome this, several bioinformatic-savvy teams have developed accessible and user-friendly tools for WGS analysis. In addition to the bioinformatics team at the Canadian Food Inspection Agency, which have made custom tools for automated processing of diagnostic WGS data, and to support research activities, The Center for Genomic Epidemiology has developed over 30 web-based tools for the analysis of bacterial WGS, including SerotypeFinder (Joensen *et al.*, 2015), Resfinder (Zankari *et al.*, 2012), VirulenceFinder (Joensen *et al.*, 2014), and many more. More recently, the Public Health Agency of Canada have also developed a useful tool for WGS analysis; IRIDA (<https://github.com/phac-nml/irida/releases>).

1.5. Mutation Rates

1.5.1. Single Nucleotide Variants (SNVs)

Single nucleotide variants (SNVs) are used to define the substitution of a single nucleotide at a specific position in the genome. SNVs are one of the main mechanisms in which novel phenotypes can occur, such as antimicrobial resistance (Shastry, 2009; Li *et al.*, 2019a; Li *et al.*, 2019b), biofilm formation (Martinez-Medina *et al.*, 2009; Lajhar *et al.*, 2018; Sharma *et al.*, 2019), virulence factors (Guy *et al.*, 2013; Li *et al.*, 2019a; Tang *et al.*, 2019), amongst many others. As such, these underlying differences in susceptibility and resistance to certain conditions can be advantageous for the bacteria in some cases. The categorization of SNVs occurs based on their location in the genome. They can occur in a non-coding sequence, which is a segment that does not encode a protein sequence, resulting in a lack of modification in the phenotype. There are several fates if the SNV occurs in a coding sequence instead. A synonymous SNV results in

no change to the amino acid sequence thanks to the degenerate nature of the genetic code (Bryant, 2012; Shastry, 2009). On the other hand, a non-synonymous SNV results in a change to the amino acid sequence, usually due to a missense mutation, which results in a different amino acid, or a nonsense mutation, which encodes a premature stop codon, cutting the protein sequence short (Bryant, 2012; Shastry, 2009). A missense mutation can have a large or little effect on protein structure and function depending on the amino acid change that has occurred. There is usually little predicted impact on the protein if the amino acid is of similar size and properties to the wildtype amino acid (Sneath, 1966). For example, the substitution of a hydrophobic phenylalanine residue with another similar hydrophobic residue, such as tyrosine. Since the only difference is an OH molecule, it would be expected that there would not be a significant impact on the structure and functioning of the protein. On the other hand, a significant impact can be predicted when an amino acid is replaced by a residue with remarkably different properties (Sneath, 1966). For example, the substitution of a large hydrophobic residue like phenylalanine to a small hydrophilic residue like cysteine. To gain insight on the impact of these substitutions, a percent accepted mutation (PAM) matrix can be used to determine how common certain substitutions are (Mount, 2008). SNVs are not always detrimental; they can also be beneficial to the bacteria, thus allowing for adaptation to occur. For example, antimicrobial resistance can be a result of SNV occurrences, which can be lucrative to the survival of bacteria. These polymorphisms can occur spontaneously due to faulty DNA metabolism or induced by environmental oppression (Friedberg *et al.*, 2006). The density of these SNVs in a given bacterial strain is affected by the mutation rate of the strain.

1.5.2. Mutation Rates

The mutation rate refers to the estimated probability of a mutation to occur per cell division (Pope et al., 2008). Organisms with high mutation rates are often called “high-mutator” strains. These strains usually have a deficiency in mismatch repair processes and play a role in the spread of antimicrobial resistance (Miller et al., 2004). The importance of mutation rate measurement lies in the implication it has on food safety and public health. High-mutator strains can quickly accumulate mutations at a high rate in adverse environments such as food processing environments and clinical settings, which could ultimately lead to adaptation (Dragosits, & Mattanovich, 2013; Swings *et al.*, 2017; Matic, 2019). These strains pose an issue during outbreak investigations as it becomes difficult to identify clusters of epidemiologically related cases due to their ability to accumulate SNVs in a short time.

1.6. Factors Affecting Mutation Rates

The rate at which an organism accumulates mutations differs between species, and even strains of the same species. Mutation rate can change in response to environmental stress, nutrition deficiency, and can even differ within regions of the genome (Elena & Visser, 2003; Galhardo *et al.*, 2012; Lukačišinová *et al.*, 2017; Maharjan & Ferenci, 2018). One of the most studied aspects involving high mutation rates is mutator phenotypes in certain genes (Lee *et al.*, 2012). All organisms require the function of DNA and mismatch repair to prevent mutations and preserve integrity. Most genes that are involved in DNA and mismatch repair almost certainly affect the mutation rate (Galhardo *et al.*, 2012). The *mutS* gene encodes one of these mismatch repair proteins and is widely studied for its effect on mutation rate as mutations and knockouts of this gene results in suppression of mismatch repair during replication, ultimately resulting in significant increases in mutation rates (LeClerc *et al.*, 1997; Willems *et al.*, 2003; Zhu *et al.*,

2014; Ishizawa *et al.*, 2015; Kang *et al.*, 2019). Other mismatch and DNA repair genes that have been implicated in hypermutability include *mutL*, *mutY*, *mutM*, and *mutT* (Fowler *et al.*, 1997; Onda *et al.*, 1999; Notley-McRobb *et al.*, 2002; Willems *et al.*, 2003). *mutD*, which encodes a proofreading polymerase III holoenzyme, *dinB*, which encodes an error-prone DNA Polymerase IV, and *DAM*, which encodes DNA adenine methylase, have also been implicated in spontaneous mutations and hypermutability (Glickman *et al.*, 1978; Echols *et al.*, 1983; Scheuermann *et al.*, 1983; Wagner *et al.*, 1999; McKenzie *et al.*, 2003).

1.7. Methods of Mutation Rate Estimation

1.7.1. Mutation Accumulation (MA) & Whole-Genome Sequencing (WGS)

Presently, the most advantageous method of mutation rate estimation is the combination of a mutation accumulation (MA) experiment and whole-genome sequencing (WGS) (Lynch *et al.*, 2016). The MA method is characterized by the persistent bottlenecking of a given population so that the evolution of the population is driven by nearly pure genetic drift (Kondrashov & Kondrashov, 2010). Under these conditions, any mutations that occur within the single cell that is chosen to continue the line will be fixed in subsequent descendants, thus resulting in a condition of weak selection (Kondrashov & Kondrashov, 2010). Under these circumstances, all viable mutations will accumulate in the bacterial genome at the approximate same rate at which they occur in the case of endogenous factors (such as DNA replication and damage repair) (Kondrashov & Kondrashov, 2010). Following WGS of the final mutant, the number of accumulated SNVs can be identified using a pipeline such as SNVPhyl (Petkau *et al.*, 2017). The number of SNVs can then be used to calculate the estimated mutation rate. The mutation rate per generation can be calculated by taking into account the total number of generations (N_t), the number of SNVs at time point 2 (N_2), and the number of SNVs at time point 1 (N_1):

$$\text{Equation (1) Mutation Rate/Generation} = \frac{(N_2 - N_1)}{N_t}$$

1.7.2. Fluctuation Assay

The fluctuation assay, first developed by Luria and Delbrück (1943), is another accepted method for mutation rate estimation. The original experiment set out to discover whether the theory of natural selection proposed by Charles Darwin applied to bacteria as well, more specifically, whether a mutation occurs before selection. In agreement with the theory, they found that mutations can occur spontaneously in the absence of a selective condition, rather than occurring as a result of selective pressure (Luria & Delbrück, 1943). The execution of a fluctuation assay begins with the inoculation of parallel cultures in a non-selective medium using a modest number of cells (Foster, 2006). Following growth to equal saturation, the number of mutants per culture is calculated by plating the cells on a selective medium. Dilutions of the cultures are then plated on a nutrient-rich medium to obtain the number of viable cells (Foster, 2006). The obtained data can be inputted into a tool such as FALCOR (Hall *et al.*, 2009) or Bz-Rates (Gillet-Markowska *et al.*, 2015) for mutation rate estimation. FALCOR takes into account the number of mutants (r) and the total number of viable cells (N_t) with the implementation of the mutation rate calculation of your choosing, such as the Lea-Coulson method of the median (Lea and Coulson, 1949) and the Ma-Sandri-Sarkar maximum likelihood estimator (Zheng, 2002). In addition to including the number of mutants and the total number of viable cells, Bz-Rates gives the option to take into account the plating efficiency (z), which considers a deviation from the Luria-Delbrück distribution, and the relative fitness of mutant versus wildtype cells (b). In addition to the mean number of mutations per culture and the mutation rate per cell per division (both corrected and not corrected by the plating efficiency), the Bz-Rates output

contains confidence limits, the average number of plated cells per culture and its standard deviation, as well as Pearson's chi-square value and p-value (Gillet-Markowska *et al.*, 2015).

1.7.2.1. Nalidixic Acid Resistance

Nalidixic acid is a synthetic quinolone with a narrow spectrum of activity traditionally prescribed for the treatment of urinary tract infections (Johnson *et al.*, 2003). It is a bacteriostatic antimicrobial primarily used against gram negative bacteria by targeting two enzymes which are essential for cell survival: DNA gyrase, and DNA topoisomerase IV. Inhibition of these two enzymes, in turn, inhibits DNA synthesis (Madigan *et al.*, 2015). DNA gyrase is only found in bacteria and uses energy released from ATP hydrolysis to induce negative supercoils, making it a great target for antimicrobials. The enzyme is made of an A₂B₂ tetramer, in which the A subunit encodes the gene *gyrA*, and the B subunit encodes the gene *gyrB*. Topoisomerase IV is also made of a tetramer comprised of two A subunits and two B subunits where subunits A is encoded by the gene *parC* and the B subunits is encoded by the gene *parE*. For nalidixic acid resistance, it is thought that several mutations are needed for a high level of resistance to be produced. The mutations in *E. coli* that are most implicated in acquiring resistance are in the *gyrA* gene at positions Ser83, and Asp87 (Nakano *et al.*, 2013; Fabrega *et al.*, 2008). Lower levels of resistance in comparison have been associated with Ser80 and Glu84 in the *parC* gene (Nakano *et al.*, 2013; Fabrega *et al.*, 2008). This antibiotic was chosen as the selective condition for the fluctuation assays in this study due to the rarity of the mutations conferring resistance, and because the minimal inhibitory concentration of the antibiotic is greatly increased when resistance is acquired. Nalidixic acid susceptible strains typically have a minimal inhibitory concentration (MIC) of $\leq 1\mu\text{g/mL}$ - $16\mu\text{g/mL}$, whereas nalidixic acid resistant strains typically have an MIC of $32\mu\text{g/mL}$ – $1,024\mu\text{g/mL}$ (Johnson *et al.*, 2003; Sáenz *et al.*, 2003; Cavaco &

Aarestrup, 2009). This allows the clear interpretation of viable mutants, which in turn allows mutation rate interpretation.

1.8. Experimental Aim

The purpose of this project was to develop a time-efficient method which can be used to identify high-mutator strains of *E. coli* in the context of a food outbreak investigation. This was done based on the detection of low frequency mutations in the *gyrA* and *parC* genes, conferring nalidixic acid resistance. The hypothesis for this study is that hypermutator foodborne pathogens could impact cluster detection. A sub-hypothesis for this study includes the notion that hypermutators could impact cluster detection frequently enough to have an impact on public health. A secondary sub-hypothesis for this study is that genes involved in DNA repair are responsible for the high mutator phenotype. The first goal of the project compared two methods of mutation rate determination: (1) the fluctuation assay and (2) the mutation accumulation assay, using strains with known mutation rates. The second goal of the project involved the development of a high-throughput screening assay derived from the fluctuation assay. The third goal of the project was to screen a panel of 100 EHEC strains for high mutators. The detected high-mutator strains were then confirmed with a 10-passage mutation accumulation experiment. The final goal of the project was to evaluate the genomes of high and low mutators to identify alleles associated with high mutation rates. This will enable future development of bioinformatic tools to identify and flag high mutators. The developed method will be useful to inform whole-genome sequence data during an outbreak investigation, as well as improve the time efficiency of high-mutator detection in EHEC with the possibility of being applied to other Enterobacteriaceae, such as *Salmonella*.

Chapter 2: Materials & Methods

2.1. Bacterial Strains and Growth Conditions

The three experimental strains (OLC0682, OLC0969, OLC0809) used to construct a functional screening method for high-mutator detection were selected as a follow-up of the thesis presented by Markell (2017). Based on previous results, OLC0682 was predicted to have a high mutation rate, whereas other strains were expected to have typical mutation rates. These strains are genetically similar to high priority VTEC but lack the verotoxin genes. The *E. coli* O111:NM strain (OLC0682) was originally obtained from the National Microbiology Laboratory (NML, Winnipeg). The O103:H2 (OLC0969) was originally isolated by the Laboratory for Foodborne Zoonosis (LFZ) in Guelph, Ontario. In the original study conducted by Markell, a nalidixic acid resistant variant of the ATCC700728 (NCTC 12900) O157:H7 strain was used (OLC0811). In the present study the parent ATCC700728 strain (OLC0809) was used as the method relies on nalidixic acid sensitivity. The positive control strain used was a derivative of *E. coli* K12 substr. BW25113 with a *mutS* knockout, causing a high mutation rate and the negative control strain has a *yqiG* (pseudogene) knockout mutation of this same strain, resulting in no effect on mutation rate. The two control strains were obtained from the Keio collection (Baba *et al.* 2006). A list of the 100 VTEC strains that were screened using the developed screening assay (outlined in section 2.6) can be found in Appendix A. The stock cultures from the OLC culture collection were maintained in tryptic soy broth (TSB, Oxoid, Ltd., Basingstoke, UK) containing 10% glycerol and stored at -80°C. The working cultures were plated on nutrient agar (OXOID, Nepean, ON, Canada) and kept at 4°C.

2.1.1. Nalidixic Acid Plates

The 25µg/mL nalidixic acid plates were prepared by first making a 25 mg/mL stock solution by adding 100 µg of nalidixic acid sodium salt (SIGMA) in 40mL of MQ water. The solution was then filtered using a 0.22µm vacuum filter (EMD Milipore Sterilflip™ Sterile Disposable Vacuum Filter Units; Thermo Fisher Scientific, Ottawa, ON, Canada). A 2.5 mg/mL working solution was made by adding 5mL of the stock solution to 45 mL of sterile water. One mL of the working solution was added to every 100 mL of nutrient agar prior to pouring. The 50µg/mL nalidixic acid plates used in the screening assay for the strains giving inconclusive results were prepared by adding 2mL of the working solution to every 100 mL of nutrient agar prior to pouring. Nalidixic acid concentrations used in the agar microdilution experiment can be found in Table 1.

2.2. Mutation Accumulation Experiment

Mutation accumulation (MA) lines originated from a frozen stock culture. *E. coli* strains O111:NM (OLC0682), O157:H7 (OLC0809), and O103:H2 (OLC0969) were plated on nutrient agar. Each line was streaked each day for single colonies onto a nutrient agar plate and incubated at 37 °C overnight (20-24 hours). Five replicates were created and passaged for 100 days. Confirmation of high-mutators detected by the developed screening assay was achieved by passaging for 10 days. The isolated colony from each plate was arbitrarily chosen to avoid bias. The plates were then stored at 4 °C for up to 3 months. A schematic demonstrating the workflow of a MA experiment is demonstrated in Figure 1.

Figure 1. Schematic demonstrating the workflow for a mutation accumulation assay. The ancestral line was subdivided into 100 MA lines for each experimental strain (OLC 0682, OLC 0809, and OLC 0969). Mutations were allowed to accumulate in the genome (as depicted below the plates) for a total of 2,900 generations, and then compared to the preserved ancestral line as a control.

2.2.1. Estimation of Generations in Colonies

The number of generations found in a colony per passage was determined by plating dilutions of four different colonies on nutrient agar and incubating them at 37 °C overnight (20-24 hours). The number of generations was calculated using the following equation, where the number of cells at the beginning of the incubation was one:

$$\text{Equation (2) Number of Generations} = \frac{(\text{Log}_{10}(\text{Cells at the end of incubation})) - (\text{Log}_{10}(\text{Cells at the beginning of incubation}))}{(\text{Log}_{10}(2))}$$

2.3. Whole-Genome Sequencing

2.3.1. DNA Isolation, Sequencing, and Assembly

Whole-genome sequencing is a laboratory technique which provides a precise DNA fingerprint of organisms, which can help link cases during an outbreak investigation. Passages 50 and 100 for each MA line for strains O111:NM (OLC 0682), O157:H7 (OLC0809), and O103:H2 (OLC 0969) were grown overnight on nutrient agar (Oxoid) at 37°C and preserved for DNA extraction. Genomic DNA (gDNA) extraction was performed on a single bacterial colony using the Maxwell 16 SEV Cell DNA Purification Kit (Promega, Madison, WI, USA) according to manufacturer's instructions. Quantification of the DNA was completed using the Qubit High Sensitivity Assay kit (Life Technologies Inc., Burlington, ON, Canada). DNA sequencing libraries were prepared from 1 ng of gDNA using the Nextera XT DNA Sample Preparation Kit (Illumina, Inc., San Diego, CA, USA) and the Nextera XT Index Kit (Illumina, Inc.) according to manufacturer's instructions. Paired-end sequencing was then performed using a 600-cycle MiSeq Reagent Kit v3 (Illumina, Inc.) on the Illumina MiSeq Platform (Illumina Inc.) with 6% PhiX control (Illumina). where the nucleotide sequence of each fragment is determined.

FastQC (version 0.11.8) (Andrews, 2010) was used to assess the quality of the raw reads and BBDuk from BBTools (version 38.22) (Bushnell, 2014) was used to trim the reads using a trim quality of 10, and the removal of reads that were less than 50 bp. The "correct" mode in tadpole (version 8.22), also from BBTools (Bushnell, 2014), was used with default parameters for error correction. The resulting contigs were then assembled with SKESA (version 2.3.0) (Souvorov *et al.*, 2018) with the vector percent argument option disabled. Qualimap (version 2.2) (García-Alcalde *et al.*, 2012; Okonechnikov *et al.*, 2016) and quast (version 2.3) (Gurevich *et al.*, 2013) were then used to assess the resulting assembly quality. ConFindr (version 0.5.0) (Low

et al., 2019) was used to detect possible contamination using default settings. Custom Python scripts were used to conduct isolate characterization (<https://github.com/OLC-Bioinformatics/COWBAT/tree/v0.1.5>).

2.3.2. SNV Analysis

Identification of high quality SNVs and the generation of maximum likelihood phylogenetic trees was performed using SNVPhyl (Petkau *et al.*, 2017) with the recommended default settings, where the minimum coverage was set to 10, the minimum mean mapping quality score was set to 30, and the alternative allele proportion was set to 0.75. Reads of each isolate was compared to that of a closed reference genome of each strain. Variant calling is performed using FreeBayes (version 0.9.20), where insertions/deletions are removed and complex variant calls are split, and the SAMtools and BCFtools packages which confirm the variant calls generated by FreeBayes, as well as generate base calls for non-variants (Petkau *et al.*, 2017). Base calls are filtered through the set parameters previously set, where base calls that did not pass the criteria are removed, and the remaining base calls that did pass the criteria are referred to as high-quality SNVs. These variant calls for each isolate are combined into a population wide SNV matrix, which is then used to construct a phylogenetic tree estimated using PhyML's approximate likelihood ratio test. More detailed information can be found in Petkau *et al.* (2017).

2.4. Fluctuation Assay

The protocol utilized was based on the method described by Luria & Delbrück (1943). Glycerol-stock cells were streaked for single colonies then inoculated in 5mL of nutrient broth and left to grow overnight (20-24 hours) at 37 °C. A 10-fold serial dilution was then performed to obtain a 10⁻⁵ dilution, of which was further diluted in fresh nutrient broth to obtain an approximate concentration of 10-100 CFU. 10 mL of the inoculum was then dispersed into each

of 30 tubes and left to grow overnight (20-24 hours) at 37 °C. Ten µl of this inoculum was used for enumeration (see section 2.4.1). The cultures were then centrifuged at 10000xg RPM for 2 minutes at 20 °C to pellet the cells. The supernatant was then discarded, and the pellets were resuspended into 100µL of nutrient broth. The entirety of the inoculum was spread plated onto round 25µg/mL nalidixic acid nutrient agar and left to grow at 37 °C for 48 hours. The colonies on each plate were then counted.

2.4.1. Cell Enumeration

Cells were enumerated via a drop-plate enumeration method where 96-well plates were prepared with 90 µl of nutrient broth in all columns and rows. 10 µl of each culture were dispensed into the first column of separate rows, making the 10⁻¹ dilution. A multi-channel pipette was used to transfer 10 µl of the inoculum from column 1 into column 2, making the 10⁻² dilution. This was continued until the 10⁻¹² dilution was obtained. 10 µl of each dilution was drop-plated on a square nutrient agar plate. The plates were left to dry, uncovered, in the biological safety cabinet for 15-30 minutes, then incubated overnight at 37°C. Colony forming units (CFU) were calculated using equation 3, which was used to determine the concentration of the colony. The average colony count for the dilution is divided by the volume plated (in mL), multiplied by the dilution factor (DF).

$$\text{Equation (3) Starting CFU} = \frac{\text{Avg. Colony Count}}{(\text{Vol. (mL)}) \times (\text{DF})}$$

2.5. Screening Assay

The screening method was developed as a modified fluctuation assay. An overnight culture was prepared for each strain to obtain approximately 10⁸ to 10⁹ CFU/mL. A serial dilution of these overnight cultures was then prepared for cell enumeration (see section 2.4.1)

and to obtain a 10^{-5} dilution, containing approximately 1000 to 10 000 CFU/mL. The cells were further diluted to make enough inoculum for 1mL in 24 wells of a 96 deep-well plate (Bel-Art™, Wayne, NJ, United States) making the founding population of each well approximately 10-100 CFU. The use of 24 replicates was chosen because it was the lowest number of replicates that can be used without compromising accuracy (see section 2.5.1). The deep well plate was covered using a breathable film (Corning™, Corning, NY, United States) then incubated overnight at 37°C. The plates were centrifuged at 4350 RPM for 30 minutes to form pellets at the bottom of the wells. The supernatant was removed, and the cells were resuspended in 10 µl of nutrient broth. The entirety of the inoculum in each well was drop-plated on a square 25 µg/mL nalidixic acid agar plate using a multichannel pipette and incubated at 37°C for 24-48 hours. The wells containing growth were counted and converted to percentage growth. Strains with growth in 40% or more of the wells we determined to be possible high mutators. In cases of growth in 60% or more of the replicates, strains were determined to be high mutator strain. A schematic demonstrating the screening assay protocol can be found in Figure 2, and a detailed protocol can be found in Appendix B.

Figure 2. Schematic demonstrating the workflow for a high throughput screening assay used to detect strains with possible high mutation rates. An overnight culture is diluted in nutrient broth to obtain a 10^{-5} dilution. This inoculum is further diluted 1000X then distributed into 24 wells of a 96 deep-well plate (1 mL per well) to achieve approximately 10-100 CFU per well. The plate is incubated overnight at 35°C, then centrifuged for 30 minutes at 4350 RPM. The supernatant is discarded, and the pellets are resuspended in 10 µl of nutrient broth, then plated on nalidixic acid agar plates (25µg/mL). The plates are incubated for 24-48 hours at 35°C, and the quadrants containing viable mutants are counted and converted to percentage growth.

2.6. Agar Microdilution

The protocol that was followed is similar to that proposed by the Clinical and Laboratory Standards Institute (CLSI) (2015). The minimal inhibitory concentration (MIC) of nalidixic acid and rifampicin was determined for resistant and susceptible variants of three experimental *E. coli* strains giving inconclusive screening assay results; O26:H11 (OLC0683), O113:H21 (OLC0686), and O45:H2 (OLC1020), alongside a positive susceptible control, *E. coli* ATCC 25922 (OLC 1614), and a negative control containing only NaCl. The positive control (OLC

1614) is used by the National Antimicrobial Resistance Monitoring System and the Canadian Integrated Program for Antimicrobial Resistance Surveillance for quality control of antibiotic plates. Resistant variants were selected for by plating on nutrient agar containing the corresponding antibiotic (25 µg/mL of nalidixic acid or 50µg/mL of rifampicin). Antibiotic concentrations were prepared and used as stated in Table 1. The corresponding volume of antibiotic stock solution was added to a centrifuge tube (50 mL). Mueller-Hinton agar (MHA) was added to the centrifuge tube until 50 mL is obtained. Agar with antibiotic was dispersed between two round petri plates to make two replicates for that dilution. Mueller-Hinton agar (MHA) was used as it supports satisfactory growth for most pathogens, making it the standard for MIC experiments (CLSI, 2012). One colony of each strain was used to inoculate 3mL of NaCl so that the density is 0.5 McFarland as determined by DensiCHEK TM. 10 µL of each suspension was further diluted in 90 µL of NaCl in a 96-well plate so that each 1uL aliquot plated has approximately 10^4 cfu. Aliquots were plated in duplicate using a multichannel pipette and left to dry in a biological safety cabinet for 15 minutes. Plates were then incubated overnight at 37 °C. MICs were recorded as the lowest concentration of antimicrobial agent that completely inhibits growth.

Table 1. Antibiotic dilution chart for agar microdilution assay. Volume of antibiotic stock is added to 50mL of agar to make two round agar plates with the corresponding antibiotic concentration.

Antibiotic Stock Concentration (mg/mL)	Volume of Antibiotic Stock Solution (μl)	Final Plate Concentration for each plate (μg/mL)
0.1	62.50	0.125
0.1	125	0.25
0.1	250	0.50
0.1	500	1
1	100	2
1	200	4
1	400	8
10	80	16
10	160	32
10	320	64
10	640	128
100	128	256
100	256	512
100	512	1,024

2.7. Estimation of Mutation Rates

2.7.1. Mutation Accumulation

The number of high quality SNVs were identified using SNVPhyl, as described above. In addition to the number of SNVs, the mutation rate calculation takes into account the total number of generations (n_t), where N_2 is the number of SNVs at time point 2, and N_1 is the number of SNVs at time point 1. The total number of generations (n_t) is calculated by multiplying the number of generations per passage (see section 2.2.1 and equation 4) by the number of passages completed in the MA line. For a mutation rate per nucleotide, the mutation rate per generation is divided by the genome size (bp).

$$\text{Equation (4) Mutation Rate per Generation} = \frac{(N_2 - N_1)}{n_t}$$

2.7.2. Fluctuation Assay

Mutation rate estimation for the fluctuation assay was performed using bz-rates (<http://www.lcqb.upmc.fr/bzrates>) which integrates the number of mutant cells on the plate and the total number of plated cells. The settings were left at the default; the relative fitness of mutant versus wild-type cells is not known, and the plating efficiency is 1.

2.8. Mutator Allele Identification

Alleles were identified for the following genes, previously associated with high mutation rates *dam* (NCBI gene ID: 947893), *dinB* (NCBI gene ID: 944922), *dnaQ* (NCBI gene ID: 946441), *mutL* (NCBI gene ID: 948691), *mutS* (NCBI gene ID: 947206), *mutT* (NCBI gene ID: 944824), and *mutY* (NCBI gene ID: 947447). The genes from the DNA sequences for the 100 screened strains were aligned with sequences of the reference genes identified above. This analysis was performed using AlleleFinder (<https://github.com/adamkoziol/AlleleFinder>). Protein sequence were aligned using CLUSTAL (version 2.1) (<ftp://ftp.ebi.ac.uk/pub/software/clustalw2/2.1/>) (Larkin *et al.*, 2007) to identify variations associated with high mutator strains.

Chapter 3: Results

3.1. Determination of Mutation Rate in *E. coli*

In the first part of this study, two approaches for mutation rate determination were compared: Fluctuation assay and mutation accumulation. This was accomplished using a panel of strains including a high mutator control (*mutS*), three experimental strains, one of which was a suspected high mutator (O111:NM), and a negative control with a pseudogene knockout (*yqiG*).

3.1.1. Fluctuation Assay

A 30-tube fluctuation assay was completed for strains O111:NM (OLC0682), O157:H7 (OLC0809), and O103:H2 (OLC0969), alongside a high mutator control with a *mutS* knockout, and a negative control with a *yqiG* knockout. Parallel 10 mL cultures were grown overnight and the number of mutants per culture was calculated by plating the cells on nutrient agar containing 25µg/mL of nalidixic acid. Typically, this assay is conducted using smaller volumes, however, the volume was increased due to rarity of the mutations known to confer nalidixic acid resistance. The fluctuation assay was repeated three times to generate three biological replicates for each strain. Nalidixic acid resistance was consistently observed in greater than 60% of cultures for high mutator strains (O111:NM and the *mutS* knockout), and in fewer than 40% of cultures for the low mutator strains (O157:H7, O103:H2, and the *yqiG* knockout) (Figure 3). Positive O111:NM cultures contained up to 445 resistant colonies whereas positive O157:H7 cultures generally contained up to 50 colonies, with a few containing up to 101 colonies. Positive O103:H2 cultures generally contained up to 50 colonies, however, several samples contained up to 454. The total number of cells per culture were similar among strains, ranging from 2.00×10^{10} to 1.46×10^{11} in a 10mL volume. Raw data on culture enumeration and viable mutant count can be found in Appendix C. A t-test demonstrated a significant difference in

percentage of plates containing viable mutants between *E. coli* O111:NM and O157:H7 ($P = 1.35 \times 10^{-2}$), as well as O103:H2 ($P = 1.08 \times 10^{-2}$), however, there was no significant difference when compared to the *mutS* knockout strain ($P = 1.97 \times 10^{-1}$). *E. coli* O157:H7 and O103:H2 were not significantly different when compared to each other ($P = 4.40 \times 10^{-1}$) and to the *yqiG* knockout ($P = 4.02 \times 10^{-1}$ and $P = 1.91 \times 10^{-1}$), however, are significantly different when compared to the positive control ($P = 3.51 \times 10^{-5}$ and $P = 6.99 \times 10^{-5}$) and the suspected high mutator ($P = 1.35 \times 10^{-2}$ and $P = 1.08 \times 10^{-2}$). Mutation rates ($\pm 95\%$ CL) were estimated for the experimental strains and both control strains (Figure 4). The average mutation rate for *E. coli* O111:NM, as estimated from the fluctuation assay, is 7.27×10^{-11} , with O157:H7 following with 4.77×10^{-12} , and O103:H3 with 3.40×10^{-12} . Overall, *E. coli* O111:NM is demonstrating a high rate of mutation, however *mutS* knockout had a mutation rate approximately four times that of the O111:NM strain (3.02×10^{-10}). *E. coli* O111:NM had a mutation rate 15-21x higher than strains O157:H7, O103:H2, and the *yqiG* knockout. Mutation rates had little variation between replicates for strains O157:H7, O103:H2, and the *mutS* knockout, however, the replicates for O111:NM and the *yqiG* knockout differed up to 10x (Table 2). *E. coli* O157:H7 and O103:H2 demonstrate low rates of mutation, similar to that of the strain with the *yqiG* knockout (8.93×10^{-12}).

Figure 3. Percentage of plates demonstrating growth of 10 or more viable mutants (out of 30) $\pm$ standard error for three fluctuation assay replicates (each color represents a replicate). Strains used are *E. coli* O111:NM (OLC 0682), O157:H7 (O157:H7), O103:H2 (O103:H2), a positive control containing a *mutS* knockout, and a negative control containing a pseudogene knockout (*yqiG*).

Table 2. The mutation rate for three replicates of *E. coli* strains O111:NM (OLC 0682), O157:H7 (OLC 0809), and O103:H2 (OLC 0969) determined by fluctuation assay.

Strain	Mutation Rate (per base pair per generation) $\times 10^{-12}$
O111:NM (OLC 0682)	10.7
	34.3
	173
O157:H7 (OLC 0809)	1.90
	5.40
	7.00
O103:H2 (OLC 0969)	2.40
	1.30
	6.50
MutS	263
	323
	320
YqiG	10.1
	14.1
	2.60

Figure 4. Mutation rates estimated by the average of three fluctuation assays for *E. coli* strains O111:NM (OLC 0682), O157:H7(OLC 0809), O103:H2 (OLC 0969), a positive control (*mutS* knockout), and a negative control (*yqiG* knockout). The values are a measure of mutation rate per generation as calculated by bz-rates (<http://www.lcqb.upmc.fr/bzrates>), with 95% confidence limit (CL) error bars. Note the logarithmic scale.

3.1.2. Mutation Accumulation & Whole-Genome Sequencing

Mutation rates were also determined using a mutation accumulation experiment. The whole genomes of the three experimental strains were sequenced following 100 single-cell bottlenecks (quality metrics for the sequences can be found in Appendix F). A colony enumeration experiment was performed using four separate colonies for each strain. The average number of viable cells per colony for *E. coli* O111:NM was 3.94×10^8 , whereas for *E. coli* O157:H7 the average CFU was 5.66×10^8 , and lastly, *E. coli* O103:H2 had an average CFU of 4.53×10^8 . Calculations revealed that *E. coli* O111:NM, O157:H7, and O103:H2 underwent 28.6, 29.1, and 28.8 generations, respectively. The results revealed very little variation between strains and demonstrated that each colony underwent approximately 29 generations, equaling a total of 2,900 cell divisions per strain over 100 passages. The number of

SNVs that accumulated in the genomes of the suspected high-mutator (O111:NM) varied between 185 and 355, giving a mutation rate of between 1.17×10^{-8} and 2.25×10^{-8} mutations per base pair per generation (Table 3). Most of the O111:NM lines accumulated approximately 200 SNVs, however, line B accumulated almost twice as many with 355 SNVs. *E. coli* O157:H7 accumulated no more than 8 SNVs, resulting in mutations rates ranging between 1.99×10^{-10} and 3.32×10^{-10} . Similarly, *E. coli* O103:H2 accumulated between 2 and 5 SNVs, however, line A did not accumulate any SNVs. The resulting mutation rate for this strain ranged between 0 and 3.47×10^{-10} . The average mutation rates estimated by the fluctuation assay and the mutation accumulation experiment for each experiment strain is compared in Table 4. The mutation rates estimated by the two methods are consistent with each other: the mutation rates estimated for *E. coli* O111:NM are very similar with little variability, while the mutation rates estimated for O157:H7 and O103:H2 both differ from only one order of magnitude.

Table 3. Parameters of the mutation accumulation and whole-genome sequencing experiments. All strains were passaged for 100 days with an estimated 29 generations occurring per passage. The number of accumulated SNVs were determined using SNVPhyl, and the mutation rates were determined using equation 4 in section 2.7.1. The mutation rate for each line is an average mutation rate calculated between passages 0 – 50, 50 – 100, and 0 – 100.

Strain	Description	Line	SNVs	Genome Size (MBP)	Mutation Rate (per generation per genome) $\times 10^{-3}$	Mutation Rate (per base pair per generation) $\times 10^{-10}$
O111:NM (OLC 0682)	Suspected high-mutator	A	210	5.47	73.4	134
		B	355		123	225
		C	208		72.1	132
		D	214		74.1	136
		E	185		63.8	117
O157:H7 (OLC 0809)	Parent of nalidixic-acid resistant strain.	A	5	5.19	1.72	3.32
		B	3		1.03	1.99
		C	3		1.03	1.99
		D	4		1.38	2.66
		E	4		1.38	2.66
O103:H2 (OLC 0969)	Suspected low-mutator	A	0	4.97	0	0
		B	2		0.69	1.39
		C	5		1.72	3.47
		D	2		0.69	1.39
		E	3		1.03	2.08

Table 4. Comparison of the average mutation rate for strains O111:NM (OLC 0682), O157:H7 (OLC 0809), and O103:H2 (OLC 0969) as calculated by a fluctuation assay and a mutation accumulation experiment.

Strain	Fluctuation Assay (10^{-12})	Mutation Accumulation (10^{-10})
O111:NM (OLC 0682)	72.7	148
O157:H7 (OLC 0809)	4.77	2.52
O103:H2 (OLC 0969)	3.40	1.66

Maximum likelihood trees were generated for strains O111:NM, O157:H7, and O103:H2 based on high quality SNVs identified with SNVPhyl (Figures 5, 6, and 7, respectively). The phylogenetic tree of the *E. coli* O111:NM lines demonstrates what appears to be distant relationships between the isolates despite being derived from the same reference isolate. The O157:H7 and O103:H2 lines have a much closer evolutionary relationship and all appear to be closely related to the reference isolate, which is representative of their low number of accumulated SNVs. *E. coli* O111:NM lines gained 99-212 SNVs in the first 50 passages, then accumulated 62-143 SNVs between passages 50 and 100. *E. coli* O157:H7 accumulated between 1–3 in the first 50 passages, then accumulated 0-4 SNVs between passages 50 and 100. *E. coli* O103:H2 line A did not accumulate any SNVs following 100 passages, however, the other lines accumulated 0-2 SNVs in the first 50 passages and 0-3 SNVs between passages 50 and 100.

Figure 5. Maximum likelihood phylogenetic tree of *E. coli* O111:NM (OLC0682) following a mutation accumulation experiment with the corresponding number of SNVs accumulated. The tree was produced based on high-quality SNVs identified using SNVPhyl with default parameters. Each line (A – E) was sequenced at passage 50 and 100 and compared to a closed reference genome for the strain. No data on passage 50C. There are 99 - 355 SNVs among isolates in this cluster.

Figure 6. Maximum likelihood phylogenetic tree of *E. coli* O157:H7 following a mutation accumulation experiment with the corresponding number of SNVs accumulated. The tree was produced using SNVPhyl with default parameters. Each line (A – E) was sequenced at passage 50 and 100 and compared to a closed reference genome. There are 3 - 5 SNVs among isolates in this cluster.

Figure 7. Maximum likelihood phylogenetic tree of *E. coli* O103:H2 (OLC0969A) following a mutation accumulation experiment with the corresponding number of SNVs accumulated. The tree was produced using SNVPhyl with default parameters. Each line (A – E) was sequenced at passage 50 and 100 and compared to a reference genome that was not passaged. There were 0 - 5 SNVs among isolates in this cluster.

3.2. Screening Assay

While the fluctuation assay was a relatively quick approach for assessing strain mutation rate, it was not practical for high-throughput screening. To address this, a high-throughput screening method was derived from the fluctuation assay. The screening assay is less intensive and uses 24 one milliliter cultures rather than 30 larger cultures and can be conducted in a 96 well plate format. To assess performance for identifying high mutators relative to the fluctuation and mutation accumulations assays, the set of strains used in the

previous experiments were tested with the screening assay. To start, two biological replicates containing four technical replicates each was performed using 94 wells for each in order to test the method (see Appendix D.1 for raw data). Once this was deemed a success, three biological replicates were completed using 24 wells, as described above (see Appendix D.2 for raw data). Nalidixic acid resistance was consistently observed in greater than 60% of *E. coli* O111:NM cultures, whereas *E. coli* O157:H7 and O103:H2 consistently demonstrated nalidixic acid resistance in less than 40% of the cultures (Figure 8), as also demonstrated by the fluctuation assay (Figure 3). A t-test demonstrated a significant difference between *E. coli* O111:NM and the O157:H7 ($P = 6.80 \times 10^{-4}$), O103:H2 ($P = 2.80 \times 10^{-3}$), as well as the *mutS* knockout strain ($P = 8.94 \times 10^{-3}$). *E. coli* O157:H7 and O103:H2 were not significantly different when compared to each other ($P = 5.88 \times 10^{-1}$) and the *yqiG* knockout ($P < 3.25 \times 10^{-1}$), however, were significantly different when compared to the positive control and the suspected high mutator ($p\text{-value} < 1.57 \times 10^{-4}$). Overall, the screening assay depicted results in accordance with that of the fluctuation assay.

Figure 8. Percentage of wells containing viable mutants $\pm$ standard error as determined by three screening assay replicates (each color represents a replicate). Strains used are *E. coli* O111:NM (OLC 0682), O157:H7 (OLC 0809), and O103:H2 (OLC 0969), a positive control containing a *mutS* knockout, and a negative control containing a pseudogene knockout (*yqiG*).

3.3. Screening 100 VTEC Strains

The developed screening assay was used to screen 100 VTEC strains. The complete dataset of screened VTEC strains can be found in Appendix E. The screening assay detected two high-mutator strains (Table 5). Upon detection, the suspected high-mutator strains were further investigated with two more replicates of the screening assay. *E. coli* strains O165:H25 (OLC0787) and O22:H8 (OLC 0984) demonstrated similar results in terms of percentage growth; 2 out of 3 replicates demonstrated nalidixic acid resistance in 100% of the wells, and 1 replicate demonstrated nalidixic acid resistance in 95.83% of wells (Table 5). These results were verified with a 10-passage mutation accumulation experiment (Table 5). *E. coli* O165:H25 accumulated 238 SNVs, whereas the O22:H8 strain accumulated 67 SNVs in 10 passages

(quality metrics for the sequences can be found in Appendix F). The resulting mutation rates for strains O165:H25 and O22:H8 was estimated to be 1.65×10^{-8} and 4.32×10^{-9} , respectively. The results from the mutation accumulation experiment are therefore in accordance with the results from the screening assay.

Additionally, three strains gave inconclusive screening assay results due to growth of spreading colonies, therefore an agar microdilution experiment was performed to determine if these strains were naturally more resistant to nalidixic acid (see section 3.5).

Table 5. Parameters of the mutation accumulation and whole-genome sequence experiments. Mutation rate calculated via a mutation accumulation experiment, compared to the percentage of wells containing viable mutants identified by three screening assays. Strains were identified as potential high-mutators by the developed screening method then passaged for 10 days, before being sequenced and performing a SNV analysis using SNVPhyl. Each passage contains 29 generations, as calculated using equation 2.

Strain	Wells with viable mutants (%)			Mutation Accumulation/Whole-Genome Sequencing			
	1	2	3	SNVs	Genome Size (MBP)	Mutation Rate (per generation) $\times 10^{-2}$	Mutation Rate (per base pair per generation) $\times 10^{-9}$
O165:H25 (OLC 0787)	100	95.83	100	238	4.96	82.1	16.5
O22:H8 (OLC 0984)	95.83	100	100	67	5.34	23.1	4.32

3.3.1. Agar Microdilution

An agar microdilution experiment was conducted to determine the minimal inhibitory concentration (MIC) of nalidixic acid and rifampicin for three strains giving inconclusive

screening assay results due to low levels of confluent growth. The experiment was conducted on nalidixic acid (Nal)/rifampicin (Rif) susceptible and resistant variants of the strains. All three strains demonstrate the same trend; susceptible isolates had low MICs (2-4µg/mL of Nal, 4-8µg/mL of Rif), whereas resistant variants had high MICs (256µg/mL Nal, >1024µg/mL Rif), and thus full resistance (Table 6). These strains were compared to a susceptible control (OLC 1614) which demonstrated a lower MIC than the susceptible strains when plated on nalidixic acid and double the MIC of the susceptible strains when plated on rifampicin. The concentration of nalidixic acid used in the screening assay was then increased to 50µl/mL, which allowed clear interpretation of the results. Nalidixic acid resistance was found in 4.17 % of wells from the screening assay for both O26:H11 (OLC 0683) and O113:H21 (OLC 0686), and 0% of wells for O45:H2 (OLC1020), suggesting a low mutation rate (see Appendix F).

Table 6. Minimal inhibitory concentrations (MICs) of nalidixic acid and rifampicin for resistant and susceptible variants of *E. coli* strains O26:H11 (OLC 0683), O113:H21 (OLC 0686), and O45:H2 (OLC 1020). MICs were compared to that of a susceptible positive control (OLC 1614), and a negative control containing sterile NaCl.

Strain	Resistance	MIC (µg/mL)	
		Nalidixic Acid	Rifampicin
Positive Control (OLC 1614)	N/A	1	8
Negative Control	N/A	-	-
<i>E. coli</i> O26:H11 (OLC0683)	Susceptible	2	4
	Resistant	256	>1024
<i>E. coli</i> O113:H21 (OLC0686)	Susceptible	2	4
	Resistant	256	>1024
<i>E. coli</i> O45:H2 (OLC1020)	Susceptible	4	4
	Resistant	256	>1024

3.4. Mutator Alleles

Given that WGS was available for strains used in this study, genes known to be associated with differences in mutation rate were characterized: *dam*, *dinB*, *dnaQ*, *mutL*, *mutS*, *mutT*, and *mutY*. The two strains identified as high mutators by the screening assay, O165:H25 (OLC 0787) and O22:H8 (OLC 0984), were found to be missing the *mutS* gene.

The O111:NM isolate had unique alleles in the *mutT* and *mutL* genes relative to the strains deemed to have typical mutation rates. The P352T mutation in MutL and G37S in MutT were not observed in any of the non-mutator strains included in this study. In the course of the mutation accumulation study of this strain, several mutations were acquired within the *mutS* gene resulting in 5 different nucleotide sequences for this gene. Line B of the mutation accumulation assay for the isolate O111:NM, which had a higher mutation rate relative to the others, had nonsynonymous mutations in MutS and DinB. There was F40C mutation in MutS, relative to the parent strain, and a Glu98Ala mutation in the *dinB* gene.

Chapter 4: Discussion

The purpose of this study was to find an accurate and timely approach for estimating strain mutation rates so that high mutators could be rapidly identified. Three approaches were evaluated: a fluctuation assay, a mutation accumulation assay in combination with WGS, and a rapid screening method. It was demonstrated that mutation rate calculations via mutation accumulation and fluctuation assay both provided similar results. Furthermore, a rapid screening method could be used to accurately identify mutator isolates in a timely manner. The screening method successfully distinguished the high mutator strain from two low mutator strains, in addition to identifying two novel high-mutators by screening 100 VTEC strains.

4.1. *E. coli* Mutation Rate Determination

Fluctuation assays typically require smaller volumes (~100-200µl) than what was used in this study (Lee *et al.*, 2012; Bhatnagar & Wong, 2019; Ramiro *et al.*, 2020), however, because the mutations conferring nalidixic acid resistance are rare, larger volumes were required to generate populations of bacteria that were large enough (approximately 10^{10} cells) to enumerate mutants. Mutations must occur at one of four amino acids in the entire genome to confer resistance to the concentrations being used. The four amino acids include Ser83 and Asp 87 in *gyrA*, and Ser80 and Glu84 in *parC* (Nakano *et al.*, 2013). The fluctuation assay and mutation accumulation experiment both demonstrated similar results in terms of distinguishing high mutators from strains with a typical mutation rate (Table 4), however, the mutation rates estimated by the two methods differed by approximately 100x. The mutation rate for O111:NM calculated from the fluctuation assay ranged between 1.07×10^{-11} and 1.73×10^{-10} whereas the mutation rate calculated from the mutation accumulation experiment varied between 1.17×10^{-8} and 2.25×10^{-8} . The fluctuation assay estimated mutation rates ranging between 1.90×10^{-12} and 7.00×10^{-12}

for O157:H7 and between 6.50×10^{-12} and 1.30×10^{-12} for O103:H2, whereas the mutation rates calculated via the mutation accumulation experiment ranged from 1.99×10^{-10} to 3.32×10^{-10} and 0 to 3.47×10^{-10} , respectively. Studies have shown that the typical mutation rate for a wild-type *E. coli* strain normally ranges between 8.90×10^{-11} and 2.20×10^{-10} mutations per base pair per generation (Drake, 1991; Wielgoss *et al.*, 2011; Lee *et al.*, 2012). This is consistent with the mutation rates estimated for the two low mutators by the mutation accumulation experiment, however, the mutation rate estimates for the wild-type strains are 10-100x higher than the mutation rates estimated by the fluctuation assay. The fluctuation assay estimated a mutation rate of 3.02×10^{-10} for the *mutS* knockout which is 1000x lower than the mutation rates estimated in other studies investigating *mutS* knockouts, which ranged between 2.71×10^{-7} and $\sim 5.00 \times 10^{-7}$ (Turrientes *et al.*, 2013; Jee *et al.*, 2016; Swings *et al.*, 2017). Although the results are not entirely consistent with other studies, the mutation rates estimated in these studies give an idea of the expected range of mutation rates in *E. coli*.

The explanation as to why these two methods have differing mutation rates is quite simple: the mutation accumulation experiment detects every non-deleterious mutation in the genome, whereas the fluctuation assay only detects a small number of mutations conferring nalidixic acid resistance. Additionally, several other factors may have contributed to the discrepancies between the two methods. The calculation used to estimate mutation rates following a fluctuation assay assumes that mutants will grow as efficiently as their parents (Pope *et al.*, 2008). Although the mutations being selected for in the fluctuation assay are associated with little fitness cost, this may be a contributing factor, especially to high mutators (Gillespie *et al.*, 2002). Another factor includes the incomplete detection of mutants. Mutations can sometimes occur later, which can result in phenotypic lag (Luria, & Delbrück, 1943; Sun *et al.*, 2018;

Carballo-Pacheco *et al.*, 2020). This can lead to some mutants not being counted, and thus not being included in the mutation rate calculation. To account for this, correct initial and final plating volumes must be used. Other possible contributing factors include the concentration of antibiotic used, as well as carbon source availability (Pope *et al.*, 2008; Foster, 2007), and the ability for a mutation rate to fluctuate in a population of cells (see section 4.4). Several attempts have been made to improve the accuracy of mutation rate estimates since the development of the fluctuation assay in 1943 (Asteris & Sarkar, 1996; Kepler & Oprea, 2001; Zheng, 2002; Zheng, 2005; Gerrish, 2008; Zheng, 2017). Currently, some of the simplest and most accurate methods to estimate mutation rates are online web tools such as FALCOR (Hall *et al.*, 2009), Bz-rates (Gillet-Markowska *et al.*, 2015), and flan (Mazoyer *et al.*, 2017).

Currently, the most accurate method of mutation rate estimation is a mutation accumulation experiment. Several studies have used 100 passages in their experiment (Lee *et al.*, 2012; Alvarez-Ponce *et al.*, 2016), however, the time frame required to complete this experiment is often not practical. While this may be necessary for strains with relatively low mutation rates, high mutator variants of *E. coli* could be determined with fewer passages. *E. coli* O111:NM accumulated a large number of SNVs in the first 50 passages of the mutation accumulation assay, whereas there was not a difference between the 50 and 100 passages for the low mutators (Figure 5, 6, and 7). This study was able to accurately estimate the mutation rate of two high mutator strains detected using the screening assay (see section 3.4) in ten passages. Enough mutations accumulated in the first 10 passages of high mutators to estimate mutation rates, whereas low mutators would not typically accumulate any SNVs in the first few passages. In a food safety investigation, the most passages expected would be less than 5, since this is the approximate number of passages a strain undergoes in preparation for WGS. On the other hand,

while it may not be the most accurate, the fluctuation assay is the most time-efficient and low-cost method. This method could still be used for identification of strains with abnormally high mutation rates.

4.2. High Throughput Screening Assay

The screening assay was developed as a less intensive and more time efficient fluctuation assay in order to inform WGS data in the case of an outbreak investigation and to enable more high throughput screening of historical isolates. The method investigates 24 replicates per strain, and 4 strains per 96 deep-well plate, allowing the investigation of more strains at once. To test this method, the same three experimental and two control strains used in the fluctuation assay and MA experiment were screened using the method. The results are similar to that of the fluctuation assay (Figure 3) and are consistent with the mutation accumulation assay results in terms of distinguishing high mutators (Table 3), making this a viable option for identification of high mutators. The screening method was then used to screen 100 VTEC strains to identify any possible high mutator strains. Two of the strains selected for testing had the high mutator phenotype, indicating that the phenotype is relatively rare, but detectable even in small collections.

While screening 100 VTEC strains three of the drop-plates were giving inconclusive results. These strains (OLC 0683, OLC 0686, and OLC 1020) demonstrated background growth via the appearance of very small colonies which developed in all quadrants of the drop-plate, whether large colonies were also present or not, while other strains seemed to have a foggy film in the place where the inoculum was dispensed. The foggy film is thought to be caused by biofilm production, which was later confirmed by comparing the plates to that of other studies describing and demonstrating the appearance of a biofilm (DePas *et al.*, 2014; Schiebel *et al.*, 2017; Lajhar

et al., 2018). Biofilms are made up of cells with the ability to adhere to one another, as well as a surface. In this case, the surface would be the agar plate with nalidixic acid. The cells embed themselves into a slimy matrix which is made from extracellular polymeric substances produced by the cells (López *et al.*, 2010). The production of biofilm is accomplished by a phenotypic switch controlled by the regulation of a number of genes which allows the organism to grow in such circumstances (An & Parsek, 2007). Biofilms manifest in a way to protect the cells. They offer resilience in times of environmental stress, such as growth in the presence of antibiotics (Beloin *et al.*, 2008). Several mechanisms can be responsible for increased resilience against antibiotics. Due to a nutrient gradient that exists from the top to the bottom of the biofilm, there is an increase in cell division and a decrease in metabolic activity (Høiby *et al.*, 2010). This yields somewhat dormant cells, which may play a part in antimicrobial resilience. There is also an increased mutation rate associated with biofilm growth, which often results in multi-drug resistance (Driffield *et al.*, 2008; Conibear *et al.*, 2009; Ryder *et al.*, 2012).

An agar microdilution experiment was performed on the three strains giving inconclusive results where the MIC of resistant and susceptible colonies of each strain were determined for nalidixic acid and rifampicin. The susceptible isolates demonstrated an MIC of between 2 and 4 µg/mL for nalidixic acid and 4 µg/mL for rifampicin which is consistent with other studies (Cambeau *et al.*, 1993; Sjölund *et al.*, 2009). The resistant isolates demonstrated full resistance (Table 6), which suggests that once an isolate accumulates the required mutations for resistance, they can survive on even the highest concentrations of antibiotics. These results also suggested the possibility that even the susceptible isolates may be more resistant to nalidixic acid than what is typical for *E. coli*. In order to eliminate the background growth and the biofilm, the nalidixic acid concentration in the agar plates was doubled to 50 µg/mL and a new criterion was created

when counting the wells with viable mutants; only large, clear colonies are considered (Appendix B). This allowed clear interpretation of the plates, and confirmation that none of the strains in question were high mutators, as all of them demonstrated viable mutants in less than 40% of the wells (see Appendix D).

4.3. Identification of Genes Associated with High Mutation Rate

The thesis by Markell (2017) predicted that the O111:NM strain used in this study is a high mutator. In this study, the O111:NM isolate demonstrated mutations in the *mutS* gene, as well as unique alleles in the *mutT* and *mutL* genes, however, this requires further investigation. Additionally, 100 VTEC for which WGS was available were characterized to determine if there were gene alleles more highly associated with high mutators. The two identified high mutators (O165:H25 and O22:H8) appeared to be missing the *mutS* gene, which was surprising since these two isolates were naturally occurring clinical strains that were not modified in the lab. It was unexpected for a conserved gene to be missing in such strains. Interestingly, these two strains were provided as reference strains for method testing. A high mutation rate is undesirable in a reference panel, and possibility should be considered when selecting reference strains for validating a method.

The mutation accumulation experiment also demonstrated a large range in accumulated SNVs between *E. coli* O111:NM lines (Table 3). Most of the lines accumulated between 185 and 214 SNVs, however, one outlier was detected. Line B accumulated almost twice the number of SNVs in the same time frame. Upon investigation, it was found that this line has a mutation in the DNA repair gene, *dinB*, resulting in an amino acid change at residue 98 in the DinB protein. Glutamate, an acidic residue, was replaced with Alanine, an aliphatic residue. It is suspected that this may have influenced the mutation rate of this line. DNA polymerase IV, which is the protein

encoded by the *dinB* gene, is involved in translesion repair and untargeted mutagenesis (Wagner *et al.*, 1999; Tang *et al.*, 2000). It is known to be poorly processive, error prone, and exhibits no proofreading mechanisms (Wagner *et al.*, 1999), and studies have shown that this protein is involved in mutagenesis (Strauss *et al.*, 2000; McKenzie *et al.*, 2003; Galhardo *et al.*, 2009).

At first glance, it was speculated that this line accumulated a mutation in a gene that is involved in DNA or mismatch repair. Since this strain is already a high evolving strain, it is probable to have accumulated a mutation of sorts, which increases the mutation rate of the organism. For a gene to influence mutation rate, it is typically involved in mismatch or DNA repair. One of the most studied genes for high mutation rates is *mutS*, a DNA repair gene, which upon deletion results in a drastic increase in mutation rate (LeClerc *et al.*, 1997; Willems *et al.*, 2003; Lee *et al.*, 2005; Zhu *et al.*, 2014; Ishizawa *et al.*, 2015; Kang *et al.*, 2019), which is why a *mutS* knockout was used as a positive control for this study. In one of the *E. coli* O111:NM (OLC 0682) lines, a substitution mutation was developed at amino acid position 40 of the *mutS* gene where phenylalanine was substituted for cysteine. This substitution appears to be novel, as it has not been observed in the literature to date. Phenylalanine is thought to be fairly non-reactive as it is rarely involved in protein function directly, but rather has a structural importance (Betts & Russell, 2003). It can be substituted with other aromatic or hydrophobic amino acids, preferably tyrosine due to its structural similarity. On the other hand, cysteine is a small and neutral amino acid. The role of this amino acid is dependent on cellular location, making it difficult for interpretation (Betts & Russell, 2003). Cysteine has been known to be involved in disulphide bonds outside of the cell, however, disulphide bonds are unlikely inside of the cell due to its reducing environment. The presence of this amino acid in the wrong environment could affect protein folding (Betts & Russell, 2003). Upon alignment, it was demonstrated that

the mutation is not in a conserved region. Kurusu *et al.* (2000) found that there are five conserved regions in the *mutS* gene in *E. coli*; the first is the amino acid sequence GRHPVVE (AA sequence 583-589), the second is TLFATHYFELT (663-673), the third is IITGPNMGGKSTYMRQ (550-566), the fourth is GRSTFMVEM (606-614), and the last one is SLVLMDE (628-634). The mutation is considered to be in a non-conserved region at position 40. Similarly, upon inspection of the sequence using the Conserved Domain Search (Marchler-Bauer, 2017), the location of the mutation seems to be in a location in which there is a probability for mutations to occur. With this evidence it is possible that this mutation may have had an effect on the mutation rate for that isolate.

The VTEC strains that were screened were first investigated for nalidixic acid resistance. Two strains were found to have resistance to nalidixic acid and were thus excluded from the experiment. Another antibiotic that is widely used with fluctuation assays and mutation rate estimation is rifampicin (Pope *et al.*, 2008; Lee *et al.*, 2012; Krašovec *et al.*, 2014; Swings *et al.*, 2017; Ramiro *et al.*, 2020). Rifampicin resistance is achieved via mutations in the *rpoB* gene. This gene encodes the beta subunit of RNA polymerase, which, once altered by a mutation, has a reduced affinity to rifampicin (Goldstein, 2014). The majority of the mutations leading to resistance occur in three distinct clusters known as the rifampicin resistance-determining region (RRDR) (Xu *et al.*, 2005). The first cluster includes 13 out of 17 alleles conferring rifampicin resistance and spans the amino acid sites 507-533. The second cluster spans amino acid sites 563–572 and includes 3 of the alleles, whereas the third cluster is at amino acid site 687, which includes 1 allele. In addition to these mutations, other less common mutations have also been identified (Landick *et al.*, 1990; Severinov *et al.*, 1993) to confer resistance. Very few mutations are needed for nalidixic acid resistance, all of which significantly increase the MIC to nalidixic

acid. This, however, seems to also be the case with rifampicin (Table 6). Since they can both be used for fluctuation assays, rifampicin could be used on nalidixic acid resistant strains, and vice versa. However, use of this alternative antibiotic should be verified with the reference panel used in this study.

4.4. Other Factors Affecting Mutation Rate

Although having a high mutation rate is beneficial in times of environmental and genetic stress, there is speculation as to whether it can also be detrimental in times of non-stress. When referring to non-neutral mutations, advantageous mutations are rare, making deleterious mutations the most common (Sprouffske *et al.*, 2018). A high mutator strain is more likely to accumulate a number of deleterious mutations, which would ultimately result in lower fitness. High mutation rates can be quite detrimental since selection tends to reduce the mutation rate in the population. This might be overcome due to mutation rate plasticity in certain sites in the genome, depending on the environment. The mutation rate also has an inverse relationship with population density because the larger a population, the more competition exists, and vice versa (Krašove *et al.*, 2014; Aanen & Debets, 2019). Any factor, whether exogenous or endogenous, that affects the relationship between mutagenesis and DNA repair seems to affect the fluctuation in mutation rate of a given cell. Several factors which may affect mutation rate plasticity include dNTP pool sizes (Gon *et al.*, 2011), factors affecting the expression of repair proteins, such as antibiotics (Gutierrez *et al.*, 2013), as well as factors affecting the presence, or lack thereof, of repair proteins (Uphoff *et al.*, 2016).

Although some mutations can occur randomly, non-random mutations are an important factor in driving evolution. In other words, mutational hotspots have an effect on the mutation rate of a given organism. Many studies have shown mutational hotspots and coldspots in *E. coli*

(Martincorena *et al.*, 2012; Shee *et al.*, 2012). Hotspots may be present in areas of the genome which promote growth and resilience in times of stress, and possibly facilitating multiple non-deleterious mutations in small zones throughout the genome to facilitate protein evolution (Ninio, 1996; Romero & Arnold, 2009). There doesn't seem to be a significant relationship between hotspot mutability and the overall mutation rate of the organism, as the only sites associated with hypermutability in *E. coli* are short tandem repeats (Deutsch *et al.*, 1997; Chattopadhyay *et al.*, 2009) compared to SNVs. There is, however, a relationship in the opposite direction. High-mutation rates, specifically caused by mutations in mismatch repair genes, cause a higher degree of mutability in mutational hotspots (Chattopadhyay *et al.*, 2009).

4.5. Mutation Detection in The Context of Food Safety Investigation

Mutation rate is not typically considered in the context of a food safety investigation, however, use of this information could increase sensitivity of cluster detection. Ideally mutation rate could be predicted by analysis of genes associated with mutation rate in WGS data. As seen with the O111:NM strain, where no genetic variations associated with mutation were detected, this is not always possible given current state of knowledge. Application of high throughput screening assay will help to identify a larger set of high mutator strains which could then be used to identify gene alleles associated with high mutation rate. Even though all the isolates in Figure 5 are derived from the same strain, they have accumulated such a high number of SNVs that they appear to be unrelated. This poses an issue in food safety because isolates demonstrating a phylogenetic tree of this nature with more than 20 SNVs separating the isolates would be excluded from a cluster. The phylogenetic trees of *E. coli* O157:H7 and O103:H2 (Figure 6 & 7) demonstrates that all lines are part of one large clade with very few evolutionary events between them. Both strains demonstrated less than 10 SNVs, which is a tentative cutoff for

inclusion/exclusion of cluster investigations (Underwood *et al.*, 2013; Joensen *et al.*, 2014; Dallman *et al.*, 2015; Butcher *et al.*, 2016; Holmes *et al.*, 2015; Rumore *et al.*, 2018; Brown *et al.*, 2019). They would thus be included in the cluster and further investigated. Insertion/deletion mutations were not accounted for in this strain investigation, which is similar to what would occur in a cluster investigation. The screening method developed in this project could be used within the timeframe of a food safety investigation to rapidly screen for high mutator phenotypes. Upon detection of nalidixic acid resistance in more than 40% of replicates, a full fluctuation assay can be performed on the respective strain for mutation rate calculation. This estimate can then be used to determine the range of mutations expected in a cluster of related isolates.

4.6. Limitations & Future Directions

The screening method was developed and tested with only *E. coli* strains; however, it is also presumed to work with other *Enterobacteriaceae* such as *Salmonella* and *Shigella*. Nalidixic acid has a narrow spectrum of activity. It targets two enzymes that are essential for cell survival; DNA gyrase and topoisomerase IV, which in turn inhibits DNA synthesis (Madigan *et al.*, 2015). DNA gyrase is only found in bacteria and uses energy released from ATP hydrolysis to induce negative supercoils, making it a great target. The most important mutations in *E. coli* are in the *gyrA* gene at positions Ser83, and Asp87, which have also been found in *Salmonella* (Nakano *et al.*, 2013; Fabrega *et al.*, 2008). Lower levels of resistance in comparison have been associated with Ser80 and Glu84 in the *parC* gene (Nakano *et al.*, 2013; Fabrega *et al.*, 2008). Therefore, the method is expected to work on any *Enterobacteriaceae* that can acquire nalidixic acid resistance through mutations in the *gyrA* and *parC* genes, such as the ones previously mentioned.

Although the screening method is much more time-efficient than a MA/WGS experiment, five days is still quite long when working on the clock of a food safety investigation. It is possible to reduce this time by two days by utilizing real-time PCR (RT-PCR). This technology has been used in combination with SNV genotyping for studies in many disciplines (Oliver *et al.*, 2000; Lager *et al.*, 2016; Broccanello *et al.*, 2018; Ayalew *et al.*, 2019). A recently characterized RT-PCR technology pairs PCR with RNase H (rhPCR), which uses blocked primers instead of conventional PCR primers and requires an additional enzyme; RNase H2 (Beltz *et al.*, 2018). The rhPrimers used with this technology differs from traditional PCR primers because they contain a single RNA base and a 3' blocking group which must be removed by RNase H2 before DNA Polymerase extension occurs. This results in increased target specificity and amplification yield (Dobosy *et al.*, 2011). In studies where rhPCR is compared to TaqMan and KASP PCR, it was found that rhAmp had the highest call rate, concordance rate, and successful design rates, and the most amplified data points (Broccanello *et al.*, 2018; Ayalew *et al.*, 2019). This method uses a novel mutant Taq DNA Polymerase and contains double enzyme mismatch detection which allows it to be more sensitive when detecting rare variants and in SNV detection compared to standard allele-specific PCR (Ayalew *et al.*, 2019). rhPCR would be a great technology to use as part of the screening assay. The plating of the strains on nalidixic acid could be replaced by adding the cells to a liquid media containing nalidixic acid, then using rhPCR to genotype SNVs in the *gyrA* and *parC* genes.

Lastly, while this study was able to identify mutations and unique alleles that may potentially be involved in conferring the high mutation rate found in *E. coli* O111, there is still further investigation that must be done. Protein modelling using a software such as Phyre (<http://www.sbg.bio.ic.ac.uk/phyre2>) (Kelley *et al.*, 2015) would be useful to identify the

location of a mutation within the protein, and the effect of the mutation on the structure. This would provide further information on the potential mutations involved in mutation rate.

Application of the screening method to more *E. coli* will enable the identification of more high mutators, and will allow the detection of more mutations associated with high mutation rates.

This can then be implemented in bioinformatic pipelines to detect these high mutators, which would do so in a much timelier manner. It is recommended that the *mutS* gene be included in automated analysis aimed at detecting high mutators. This achievement could potentially improve food safety investigations in a timely manner which will, in turn, contribute to effective regulatory action and public safety

Appendices

Appendix A: List of Strains Screened Using Screening Assay

OLC #	Organism	Serotype	VT Profile	Donor Laboratory
455	E. coli	O111:H11	VT1 (vtx1a), eae, hlyA	C. Poppe (LFZ, PHAC)
457	E. coli	O111:H11	VT1 (vtx1a), eae, hlyA	C. Poppe (LFZ, PHAC)
458	E. coli	O111:H8	VT1 (vtx1a), eae, hlyA	C. Poppe (LFZ, PHAC)
461	E. coli	O111:NM	VT1 (vtx1a)	CFIA (OLC)
464	E. coli	O26:H11	VT1 (vtx1a), eae, hlyA	C. Poppe (LFZ, PHAC)
465	E. coli	O26:H11	VT1 (vtx1a), eae, hlyA	C. Poppe (LFZ, PHAC)
466	E. coli	O26:H11	VT1 (vtx1a), eae, hlyA	C. Poppe (LFZ, PHAC)
469	E. coli	O157:H7	VT1 (vtx1a), VT2 (vtx2c), eae, hlyA	C. Poppe (LFZ, PHAC)
470	E. coli	O157:H7	VT2 (vtx2c), eae, hlyA	C. Poppe (LFZ, PHAC)
471	E. coli	O157:H7	VT2 (vtx2c), eae, hlyA	C. Poppe (LFZ, PHAC)
472	E. coli	O157:H7	VT1 (vtx1a), VT2 (vtx2a), eae, hlyA	C. Poppe (LFZ, PHAC)
473	E. coli	O157:H7	VT1 (vtx1a), VT2 (vtx2c), eae, hlyA	C. Poppe (LFZ, PHAC)
474	E. coli	O157:H7	VT1 (vtx1a), VT2 (vtx2c), eae, hlyA	C. Poppe (LFZ, PHAC)
475	E. coli	O157:H7	VT1 (vtx1a), VT2 (vtx2c), eae, hlyA	C. Poppe (LFZ, PHAC)
476	E. coli	O157:H7	eae, hlyA	C. Poppe (LFZ, PHAC)
477	E. coli	O157:H7	VT2 (vtx2c), eae, hlyA	C. Poppe (LFZ, PHAC)
478	E. coli	O157:H7	VT2 (vtx2a), eae, hlyA	C. Poppe (LFZ, PHAC)
479	E. coli	O157:H7	VT1 (vtx1a), VT2 (vtx2c), eae, hlyA	C. Poppe (LFZ, PHAC)
480	E. coli	O157:H7	VT1 (vtx1a), VT2 (vtx2c), eae, hlyA	C. Poppe (LFZ, PHAC)
481	E. coli	O157:H7	VT1 (vtx1a), VT2 (vtx2c), eae, hlyA	C. Poppe (LFZ, PHAC)
482	E. coli	O157:H7	VT1 (vtx1a), VT2 (vtx2c), eae, hlyA	C. Poppe (LFZ, PHAC)
483	E. coli	O157:H7	VT1 (vtx1a), VT2 (vtx2c), eae, hlyA	C. Poppe (LFZ, PHAC)
484	E. coli	O157:H7	VT2 (vtx2a), eae, hlyA	C. Poppe (LFZ, PHAC)
609	E. coli	O157:H7	vtx1a	Health Canada

611	E. coli	O26:H11	VT1 (vtx1a), eae, hlyA	S. Bidawid (BMH, Health Canada)
622	E. coli	O78:H11	-	ATCC35401
629	E. coli	O111:H11	VT1 (vtx1a), eae, hlyA	C. Poppe (LFZ, PHAC)
630	E. coli	O111:H11	VT1 (vtx1a), eae, hlyA	C. Poppe (LFZ, PHAC)
631	E. coli	O111:H11	VT1 (vtx1a), eae, hlyA	R. Johnson (LFZ, PHAC)
632	E. coli	O111:NM	VT1 (vtx1a), VT2(vtx2a), eae	C. Poppe (LFZ, PHAC)
633	E. coli	O111:NM	VT1 (vtx1a), eae, hlyA	C. Poppe (LFZ, PHAC)
634	E. coli	O111:NM	VT1 (vtx1a), eae, hlyA	R. Johnson (LFZ, PHAC)
635	E. coli	O26:H11	VT1 (vtx1a), eae, hlyA	C. Poppe (LFZ, PHAC)
636	E. coli	O26:H11	VT1 (vtx1a), eae, hlyA	R. Johnson (LFZ, PHAC)
637	E. coli	O26:H11	VT1 (vtx1a), eae, hlyA	C. Poppe (LFZ, PHAC)
638	E. coli	O26:H11	VT1 (vtx1a), eae, hlyA	C. Poppe (LFZ, PHAC)
639	E. coli	O26:H11	VT1 (vtx1a), eae, hlyA	C. Poppe (LFZ, PHAC)
640	E. coli	O26:H11	VT1 (vtx1a), eae, hlyA	R. Johnson (LFZ, PHAC)
667	E. coli	O119:H25	VT1 (vtx1a), eae, hlyA	C. Poppe (LFZ, PHAC)
669	E. coli	O76:H19	VT1 (vtx1d)	C. Poppe (LFZ, PHAC)
675	E. coli	O145:NM	VT1 (vtx1a), eae, hlyA	M. Gilmour (NML, PHAC)
676	E. coli	O145:NM	VT1 (vtx1a), eae, hlyA	M. Gilmour (NML, PHAC)
677	E. coli	O121:NM	VT2 (vtx2a), eae, hlyA	M. Gilmour (NML, PHAC)
678	E. coli	O121:H19	VT2 (vtx2a), eae, hlyA	M. Gilmour (NML, PHAC)
679	E. coli	O103:H2	VT1 (vtx1a), eae, hlyA	M. Gilmour (NML, PHAC)
680	E. coli	O103:H25	VT1 (vtx1a), eae, hlyA	M. Gilmour (NML, PHAC)
683	E. coli	O26:H11	eae	M. Gilmour (NML, PHAC)
684	E. coli	O145:NM	eae, hlyA	M. Gilmour (NML, PHAC)
686	E. coli	O113:H21	VT2 (vtx2a)	V. Arling (Calgary, CFIA)
707	E. coli	O165:H25	VT2 (vtx2a, vtx2c), eae, hlyA	M. Gilmour (NML, PHAC)
708	E. coli	O115:H18	VT1 (vtx1a), VT2 (vtx2a), hlyA	M. Gilmour (NML, PHAC)
709	E. coli	O5:NM	VT1 (vtx1a), eae, hlyA	M. Gilmour (NML, PHAC)
710	E. coli	O121:H19	VT2 (vtx2a), eae, hlyA	M. Gilmour (NML, PHAC)
711	E. coli	O103:H2	VT1 (vtx1a), eae, hlyA	M. Gilmour (NML, PHAC)
712	E. coli	O103:H2	VT1 (vtx1a), eae, hlyA	M. Gilmour (NML, PHAC)
714	E. coli	O111:NM	VT1 (vtx1a), eae	M. Gilmour (NML, PHAC)
715	E. coli	O111:NM	VT1 (vtx1a), VT2 (vtx2a), eae	M. Gilmour (NML, PHAC)
716	E. coli	O45:H2	VT1 (vtx1a), eae, hlyA	M. Gilmour (NML, PHAC)
717	E. coli	O55:H7	VT1 (vtx1a), eae	M. Gilmour (NML, PHAC)
718	E. coli	O157:H7	VT1 (vtx1a), VT2 (vtx2a), eae, hlyA	M. Gilmour (NML, PHAC)
720	E. coli	O91:H21	VT2 (vtx2d), hlyA	M. Gilmour (NML, PHAC)
725	E. coli	O26:H11	VT1 (vtx1a), eae, hlyA	M. Gilmour (NML, PHAC)
727	E. coli	O103:H25	VT1 (vtx1a), eae, hlyA	M. Gilmour (NML, PHAC)

728	E. coli	O103:H11	VT1 (vtx1a), eae, hlyA	M. Gilmour (NML, PHAC)
731	E. coli	O26:H11	VT1 (vtx1a), VT2 (vtx2a), eae, hlyA	M. Gilmour (NML, PHAC)
733	E. coli	O85:H1	VT2	M. Gilmour (NML, PHAC)
789	E. coli	O121:H19	VT2 (vtx2a), eae, hlyA	M. Gilmour (NML, PHAC)
791	E. coli	O121:NM	VT2 (vtx2a), eae, hlyA	M. Gilmour (NML, PHAC)
792	E. coli	O121:NM	VT2 (vtx2a), eae, hlyA	M. Gilmour (NML, PHAC)
797	E. coli	O157:H7	VT1 (vtx1a), VT2 (vtx2a), eae, hlyA	H. Cohen (OLC, CFIA)
816	E. coli	O104:H7	VT2 (vtx2a)	M. Gilmour (NML, PHAC)
982	E. coli	O121:H7	VT1 (vtx1d)	M. Gilmour (NML, PHAC)
983	E. coli	O145	eae	M. Gilmour (NML, PHAC)
984	E. coli	O22:H8	VT1 (vtx1c), VT2 (vtx2b), hlyA	F. Scheutz (Statens Serum Institut)
985	E. coli	O145:H34	VT2(vtx2f), eae	F. Scheutz (Statens Serum Institut)
986	E. coli	O157:H7	VT2 (vtx2a), eae, hlyA	F. Scheutz (Statens Serum Institut)
987	E. coli	O103:H2	VT1 (vtx1a), eae, hlyA	F. Scheutz (Statens Serum Institut)
988	E. coli	O91:H21	VT2 (vtx2d), hlyA	F. Scheutz (Statens Serum Institut)
989	E. coli	O107:H2	VT2 (vtx2b, vtx2c, vtx2d)	F. Scheutz (Statens Serum Institut)
990	E. coli	O139:H1	VT2 (vtx2e)	F. Scheutz (Statens Serum Institut)
991	E. coli	O2:H25	VT2 (vtx2g)	F. Scheutz (Statens Serum Institut)
992	E. coli	O154:H31	VT1 (vtx1d)	F. Scheutz (Statens Serum Institut)
993	E. coli	O146:H21	VT1 (vtx1c), VT2 (vtx2a), hlyA	F. Scheutz (Statens Serum Institut)
994	E. coli	O48:H21	VT1 (vtx1a), VT2 (vtx2a), hlyA	F. Scheutz (Statens Serum Institut)
995	E. coli	O174:H21	VT2 (vtx2b, vtx2c)	F. Scheutz (Statens Serum Institut)
996	E. coli	O157:H7	VT1 (vtx1a), VT2 (vtx2a), eae, hlyA	F. Scheutz (Statens Serum Institut)
997	E. coli	O118:H12	VT2 (vtx2b)	F. Scheutz (Statens Serum Institut)
998	E. coli	O73:H18	VT2 (vtx2d)	F. Scheutz (Statens Serum Institut)

999	E. coli	O2:H25	VT2 (vtx2g)	F. Scheutz (Statens Serum Institut)
1000	E. coli	O8:K85ab:H rough	VT1 (vtx1d), eae	F. Scheutz (Statens Serum Institut)
1001	E. coli	O128ac:[H2]	VT2 (vtx2f), eae	F. Scheutz (Statens Serum Institut)
1002	E. coli	O174:H8	VT1 (vtx1c), VT2 (vtx2b)	F. Scheutz (Statens Serum Institut)
1003	E. coli	O139:K12:H1	VT2 (vtx2e)	F. Scheutz (Statens Serum Institut)
1019	E. coli	O45:H2	VT1 (vtx1a), eae, hlyA	C. Poppe (LFZ, PHAC)
1020	E. coli	O45:H2	VT1 (vtx1a), eae, hlyA	R. Johnson (LFZ, PHAC)
1021	E. coli	O45:H2	VT1 (vtx1a), eae, hlyA	C. Poppe (LFZ, PHAC)
1022	E. coli	O45:H2	VT1 (vtx1a), eae, hlyA	C. Poppe (LFZ, PHAC)
1031	E. coli	O121:H19	VT2 (vtx2a), eae, hlyA	R. Johnson (LFZ, PHAC)
1036	E. coli	O145:H25	VT2 (vtx2a), eae, hlyA	R. Johnson (LFZ, PHAC)
1059	E. coli	O157:H7	VT2 (vtx2c), eae, hlyA	F. Scheutz (Statens Serum Institut)

Appendix B: Screening Assay Protocol

High Mutator Screening Assay

Purpose: To screen for high mutator strains of *E. coli* using nalidixic acid resistance.

Materials:

- Test Tubes
- Microcentrifuge Tubes
- Centrifuge Tubes (Or 96-well plate if there are numerous samples)
- Reagent Reservoir
- 96 Deep-Well Plate (VWR® 96-Well Polypropylene Plates; Catalog No. 89237-514)
- Breathable Sealing Film (Corning™ Breathable Sealing Tape for Microplates; Catalog No. 13-660-100).
- Nutrient Broth (Thermo Scientific™ Nutrient Broth (Dehydrated); Catalog No. OXCM0001B)
- Nutrient Agar (Thermo Scientific™ Nutrient Agar (Dehydrated); Catalog No. OXCM0003B)
- Nalidixic Acid Sodium Salt (Sigma Aldrich Nalidixic Acid Sodium Salt; Catalog No. N4382)

Procedure:

1. Prepare 50 µg/mL Nalidixic acid agar plates by adding 2 mL of 2.5 mg/mL nalidixic acid solution for every 100 mL of nutrient agar prior to pouring.
2. Prepare an overnight culture by transferring one full colony of bacteria from a plate to a test tube containing 5 mL of nutrient broth. Repeat for each sample. Incubate overnight at 37 °C.

NOTE: For <i>E. coli</i> , overnight cultures should be 10^8 to 10^9 CFU/mL.
--

3. Perform a serial dilution to obtain a 10^{-5} dilution:
 - a. Label 5 microcentrifuge tubes 10^{-1} to 10^{-5} and add 900 µl of nutrient broth to each tube.

NOTES:

If needed, prepare 8 microcentrifuge tubes to obtain a 10^{-8} dilution if enumeration is desired.
--

If there are numerous samples, use a 96-well plate to perform a serial dilution and following the drop-plate protocol instead of step 3.
--

- b. Add 100 µl of the overnight culture to the 10^{-1} microcentrifuge tube. Mix by pipetting up and down 10 times.
 - c. Make the 10^{-2} dilution by adding 100 µl of the 10^{-1} to the 10^{-2} tube. Continue until 10^{-5} dilution is obtained.
4. Using a 50 mL centrifuge tube, suspend 26 µl of the 10^{-5} dilution into 26 mL of nutrient broth. Vortex for 10 seconds to mix.

NOTE: This dilution should be approximately 1000 to 10 000 CFU/mL, making the founding population in each well 10-100 CFU.

5. Transfer the inoculum into a 50 mL reagent reservoir.
6. Using a multichannel pipette, transfer 1 mL of the inoculum into two rows of a 96 deep-well plate, making 24 wells. Repeat for each sample. 4 samples can fit in one 96 deep-well plate.
7. Cover the plate by placing a breathable membrane on the top of the plate. Incubate overnight at 37°C.
8. Centrifuge the plate at 4350 RPM for 30 minutes to form pellets in each well.

NOTE: Remember to balance the centrifuge.

9. Discard the supernatant from each well using a pipette.

NOTE: Place the pipette tip on the side of the well to avoid disturbing the pellet.

10. Using a multichannel pipette, resuspend each pellet in 10 µl of nutrient broth. Pipet up and down 10 times to mix. Change tips in between.
11. Making sure that the entirety of the inoculum is at the bottom of the well, drop-plate the entire volume onto a 50 µg/mL Nalidixic acid agar plate using a multichannel pipette.
12. Allow plates to dry uncovered in the biological safety cabinet for 15-30 minutes.
13. Incubate the plates inverted, at 37°C for 24-48 hours.
14. Count the squares containing growth out of 24. If growth is found in more than 40% of the squares, the strain must be further investigated. Growth in 60% or more of the squares suggests a high mutator strain.

Appendix C: Fluctuation Assay Culture Enumeration & Viable Mutant Count

Strain	Rep.	10 ⁻⁷	10 ⁻⁸	10 ⁻⁹	Mutants	10 ⁻⁷	10 ⁻⁸	10 ⁻⁹	Mutants	10 ⁻⁷	10 ⁻⁸	10 ⁻⁹	Mutants
O111:NM (OLC 0682)	1	12 15	0 1	0 0	18	11 14	1 7	0 2	55	13 30	3 2	0 0	59
	2	14 15	3 3	0 0	0	13 15	0 3	0 0	8	0 0	0 0	0 0	46
	3	16 17	4 0	0 0	19	16 9	2 1	0 0	21	13 14	2 1	0 0	362
	4	20 15	5 5	0 0	4	17 11	2 0	0 0	10	12 9	2 1	0 0	434
	5	15 18	4 2	1 0	11	14 21	2 1	0 1	5	11 15	1 2	1 0	250
	6	23 21	4 5	1 1	95	16 13	2 3	0 0	13	7 14	2 1	0 1	149
	7	0 0	0 0	0 0	15	17 8	1 1	0 0	35	10 12	2 1	0 0	203
	8	18 23	4 7	1 0	22	13 15	2 4	0 0	22	15 8	2 0	0 0	447
	9	9 12	2 1	0 0	0	15 20	0 0	0 0	17	8 15	1 0	0 0	316
	10	10 16	0 2	0 0	22	11 13	1 1	0 0	20	9 9	1 2	0 2	54
	11	13 11	1 4	0 1	16	18 15	1 2	0 0	137	30 30	2 4	0 1	58
	12	11 12	3 3	0 0	27	15 16	1 2	0 0	3	36 33	3 3	0 1	237
	13	13 TNTC	0 7	0 0	1	16 16	3 1	0 0	24	14 14	2 0	1 0	315
	14	19 21	3 4	0 0	7	17 28	3 3	0 0	8	12 10	3 4	0 0	151
	15	22 17	5 3	0 2	7	24 27	3 1	0 0	19	16 13	1 2	0 1	297
	16	13 12	1 3	0 0	0	27 22	3 7	0 0	63	5 15	2 1	0 0	40
	17	0 0	0 0	0 0	7	22 26	1 4	0 2	61	14 9	1 3	0 1	49
	18	16 12	2 2	2 0	138	22 21	2 1	1 0	112	8 10	1 0	0 0	42
	19	15	1	0	73	15	1	1	19	16	4	0	138

O157:H7 (OLC 0809)

	21	2	0		15	1	0		22	2	1	
20	11	1	1	70	22	5	3	33	16	3	0	156
	22	1	0		28	6	2		8	3	0	
21	16	1	0	1	TNTC	9	1	3	18	2	0	238
	11	2	0		TNTC	12	2		7	2	0	
22	26	3	0	5	28	7	1	18	15	1	0	173
	29	4	1		24	6	0		13	1	0	
23	12	2	0	0	0	0	0	152	19	1	0	277
	15	1	0		0	0	0		15	3	0	
24	22	1	0	0	0	0	0	129	17	2	0	208
	18	2	1		0	0	0		17	2	0	
25	12	0	0	36	0	0	0	106	9	3	0	179
	15	1	0		0	0	0		19	3	0	
26	TNTC	6	0	0	0	0	0	125	14	1	0	101
	TNTC	5	0		0	0	0		20	1	0	
27	21	2	0	0	0	0	0	169	9	0	0	130
	18	0	0		0	0	0		20	2	0	
28	20	5	0	6	0	0	0	94	15	1	0	207
	14	5	0		0	0	0		14	2	0	
29	15	4	1	122	0	0	0	93	14	2	0	409
	18	2	0		0	0	0		12	2	0	
30	24	4	0	49	0	0	0	128	16	0	0	290
	20	5	1		0	0	0		14	2	0	
1	42	8	2	0	35	3	1	21	21	4	0	1
	45	9	2		29	3	0		30	3	0	
2	39	10	1	0	30	3	1	0	31	4	1	1
	39	11	2		37	4	0		34	5	1	
3	46	14	3	0	23	3	0	10	14	2	0	4
	40	23	1		31	2	2		12	3	0	
4	42	5	1	0	19	5	0	1	10	2	0	10
	31	4	1		34	8	2		16	0	0	
5	27	8	0	0	28	5	0	2	12	6	0	0
	29	3	0		23	6	0		16	0	0	
6	0	0	0	0	29	7	3	0	19	1	0	0
	0	0	0		30	6	0		15	3	0	
7	12	2	1	0	30	7	0	8	26	3	1	0
	18	3	1		24	4	0		22	5	0	
8	23	6	1	0	29	2	0	0	19	3	0	6
	15	2	0		30	6	1		24	2	0	
9	17	2	0	0	31	2	3	1	18	1	0	10
	25	4	1		35	6	1		16	0	0	
10	19	4	1	0	24	6	2	33	8	0	0	17

O1
O3:

	16	6	2		27	1	1		16	0	0	
11	19	3	0	0	35	6	0	6	22	2	2	3
	20	2	0		26	2	1		22	1	0	
12	18	2	2	0	22	5	0	7	12	0	0	7
	22	4	0		26	5	1		10	0	0	
13	68	19	1	0	31	3	0	6	14	3	0	2
	65	17	1		34	4	0		14	1	0	
14	58	23	4	0	18	5	3	4	20	2	0	33
	63	16	2		20	5	1		12	2	0	
15	59	14	7	1	22	2	0	3	18	0	0	56
	60	17	3		20	3	0		16	3	0	
16	73	11	4	13	16	6	0	2	26	5	1	1
	68	15	2		18	5	0		19	1	0	
17	77	17	1	5	26	5	0	0	6	1	0	23
	81	20	5		22	0	0		11	3	0	
18	69	16	6	6	23	6	0	5	13	2	1	89
	76	21	8		20	4	1		11	2	1	
19	52	16	6	2	30	4	1	1	24	2	0	5
	54	21	5		33	8	1		21	2	0	
20	66	10	2	30	26	7	0	9	13	2	1	2
	57	5	1		25	9	1		16	1	0	
21	30	6	0	2	27	1	0	1	23	3	0	0
	24	5	1		32	7	0		20	1	1	
22	58	12	4	1	31	8	1	1	13	0	0	0
	50	10	2		30	5	0		18	1	0	
23	33	10	2	21	29	7	1	11	14	22	0	2
	37	10	1		27	6	1		20	16	1	
24	33	9	4	2	29	5	1	2	17	3	0	0
	49	9	2		27	3	0		14	3	1	
25	0	0	0	2	30	4	0	0	19	2	0	0
	0	0	0		22	5	0		15	4	0	
26	20	2	0	1	24	3	2	0	22	2	0	2
	25	2	0		28	5	0		20	6	0	
27	26	3	1	1	34	6	1	6	14	0	0	0
	18	2	0		31	3	1		19	5	1	
28	0	0	0	5	28	7	0	4	23	1	1	8
	0	0	0		27	5	0		21	3	1	
29	0	0	0	34	36	6	1	0	20	8	0	101
	0	0	0		34	4	0		13	4	0	
30	0	0	0	9	31	7	1	87	14	1	0	0
	0	0	0		32	8	1		16	3	0	
1	23	6	2	3	15	3	1	0	8	0	0	1

	25	6	0		19	1	0		19	2	0	
2	42	6	3	0	29	3	1	0	26	4	0	7
	37	5	2		32	1	0		22	1	0	
3	45	5	0	0	27	8	0	0	14	2	0	12
	39	11	1		34	7	1		20	2	1	
4	21	3	0	6	36	2	0	0	27	3	0	0
	31	8	1		35	5	1		13	1	1	
5	26	5	0	7	38	11	1	0	30	2	1	13
	34	5	0		39	9	1		29	3	0	
6	37	6	0	0	26	1	0	0	28	4	0	5
	33	8	2		26	2	1		30	2	2	
7	14	2	0	0	40	9	0	5	20	3	0	5
	18	1	1		39	11	1		14	2	1	
8	20	1	3	4	31	3	0	0	16	5	0	2
	12	2	3		33	11	0		15	0	0	
9	17	5	5	0	38	4	0	1	26	3	0	2
	24	4	4		37	7	0		19	5	4	
10	41	5	2	0	39	5	2	0	20	1	0	0
	42	8	0		33	5	1		15	3	0	
11	35	5	0	0	32	7	0	0	14	1	0	9
	39	4	1		32	5	2		21	0	0	
12	36	6	0	1	21	4	0	0	13	2	1	12
	32	2	0		20	3	1		15	3	1	
13	24	2	1	0	27	3	0	0	22	1	0	7
	21	5	1		24	6	0		13	5	0	
14	22	5	0	0	32	1	0	13	16	2	0	1
	17	1	0		28	5	0		12	1	1	
15	16	2	0	2	25	4	2	8	15	3	1	0
	21	3	0		22	5	1		12	4	4	
16	30	6	0	5	38	7	1	0	14	1	0	5
	21	3	1		41	10	1		19	3	0	
17	19	5	0	0	37	4	2	0	18	5	0	2
	21	4	0		34	5	1		16	0	0	
18	24	2	0	0	34	5	1	0	23	2	0	1
	25	2	1		35	6	3		20	0	0	
19	29	3	0	1	31	9	1	1	24	3	1	8
	25	2	0		35	3	3		17	5	0	
20	30	5	1	3	37	14	0	0	16	4	0	0
	13	1	1		37	11	0		19	3	0	
21	30	2	0	13	34	4	0	2	16	1	0	3
	32	3	0		35	3	1		23	1	0	
22	27	2	2	0	31	6	1	0	21	4	0	18

mutS		32	9	1		35	5	2		16	4	1	
	23	21	5	0	0	29	8	0	2	23	6	2	7
		24	0	0		29	5	0		28	4	1	
	24	32	3	0	3	27	5	0	1	21	3	1	17
		24	5	0		27	4	0		24	0	0	
	25	44	6	1	0	34	3	1	0	30	8	1	5
		39	3	0		23	4	0		26	3	0	
	26	24	6	3	0	TNTC	10	1	0	24	0	0	1
		29	1	1		TNTC	8	0		25	4	0	
	27	24	9	2	0	25	1	0	1	22	5	1	2
		22	3	2		28	3	0		25	3	0	
	28	31	2	0	1	28	5	0	0	31	5	1	5
		21	2	1		27	6	0		33	4	0	
	29	18	5	0	56	31	6	1	0	25	4	0	79
		28	4	1		34	6	0		22	3	0	
	30	18	2	0	188	39	7	2	282	29	5	0	54
		23	2	0		38	9	0		15	5	0	
	1	TNTC	11	3	139	17	2	0	115	15	1	0	209
		TNTC	9	7		15	4	0		19	5	1	
	2	22	5	0	92	0	0	0	189	33	3	0	254
		20	6	3		2	0	0		20	2	0	
	3	24	4	0	85	1	0	0	201	13	1	1	99
		27	9	0		0	0	0		20	4	0	
	4	34	8	1	144	14	2	0	118	11	3	0	89
		33	4	3		17	3	0		15	3	0	
	5	24	3	2	153	TNTC	6	1	129	9	1	0	107
		29	3	0		TNTC	9	1		9	2	0	
	6	31	6	0	207	32	2	1	174	16	3	0	199
		34	3	2		22	1	0		12	1	0	
	7	33	3	2	170	14	2	0	136	19	3	1	101
		31	5	0		21	4	0		33	4	0	
	8	25	5	0	152	25	1	0	203	10	1	0	141
		31	8	0		24	7	1		25	4	0	
	9	36	10	0	198	17	2	0	162	12	1	0	138
		39	8	2		17	4	0		19	0	0	
	10	18	1	0	150	16	5	0	185	23	2	0	100
		21	4	0		15	0	0		24	7	0	
	11	26	4	1	88	8	0	0	114	19	1	0	149
		24	2	1		6	0	0		13	5	1	
	12	19	9	0	235	10	2	0	132	17	1	0	138
		17	3	1		17	2	0		22	3	0	
	13	29	6	0	97	0	0	0	119	24	4	0	114

yqiG		22	1	0		0	0	0		23	6	0	
	14	20	2	0	226	0	0	0	101	24	1	0	149
		20	2	1		0	0	0		17	3	0	
	15	14	2	0	193	0	0	0	162	22	3	0	108
		12	7	0		0	0	0		15	2	0	
	16	24	3	1	219	17	0	0	262	27	6	0	104
		27	8	1		17	1	0		27	4	0	
	17	29	5	0	166	14	6	0	185	31	3	0	112
		33	3	0		18	1	0		29	9	0	
	18	16	2	0	139	12	1	0	191	27	8	2	142
		14	0	0		18	3	0		28	5	0	
	19	14	4	0	231	15	0	0	171	17	5	0	122
		18	1	0		11	1	0		21	2	0	
	20	0	0	0	82	8	4	0	166	24	1	0	111
		0	0	0		10	3	0		23	7	0	
	21	30	6	0	122	20	0	0	174	17	2	0	96
		29	8	0		17	4	1		17	2	0	
	22	23	0	0	118	15	4	0	288	22	4	1	92
		28	3	0		19	4	0		28	4	1	
	23	31	4	0	157	19	1	1	193	21	5	0	203
		29	8	0		22	2	1		23	7	0	
	24	26	7	0	241	12	3	0	177	17	5	1	235
		26	4	0		15	1	0		18	3	0	
	25	28	4	0	147	0	0	0	105	19	3	0	233
		20	3	0		0	0	0		7	3	1	
	26	21	2	1	88	0	0	0	199	12	3	0	90
		23	7	2		0	0	0		22	3	1	
	27	19	3	0	123	0	0	0	212	17	3	0	89
		21	6	1		0	0	0		19	5	0	
	28	TNTC	26	3	202	0	0	0	200	25	1	0	116
		TNTC	27	7		0	0	0		21	3	0	
	29	25	7	0	185	0	0	0	106	19	1	1	142
		24	6	3		0	0	0		22	5	1	
	30	28	4	0	151	0	0	0	118	25	1	0	136
		26	9	0		0	0	0		18	6	0	
	1	25	1	0	5	15	0	0	13	19	0	0	0
		26	5	0		20	5	1		19	1	1	
	2	14	2	0	5	16	2	1	2	22	4	0	5
		18	2	1		19	2	0		27	6	1	
	3	11	2	0	2	15	0	0	1	21	1	0	26
		20	3	0		16	2	0		20	1	0	
	4	68	21	4	8	11	1	0	7	17	2	2	0

	65	21	2		12	0	0		16	1	1	
5	48	2	1	2	11	0	0	2	19	2	1	0
	53	4	3		13	1	0		24	3	0	
6	29	3	3	2	25	1	0	5	17	2	0	0
	38	5	0		17	4	0		18	2	0	
7	31	1	0	0	24	3	0	4	23	4	1	3
	21	2	0		24	4	0		28	3	1	
8	18	2	0	10	25	2	0	7	19	3	0	23
	15	4	0		24	3	0		23	3	0	
9	41	7	0	0	13	1	0	3	17	2	1	0
	23	6	0		17	2	0		19	2	0	
10	14	1	0	15	17	0	0	23	11	1	0	0
	19	1	0		22	2	0		9	1	1	
11	14	0	0	8	17	2	0	0	21	2	0	0
	12	2	0		17	3	0		9	2	0	
12	11	0	0	2	21	3	0	27	18	2	0	0
	11	1	0		13	4	0		21	3	0	
13	13	2	0	1	14	2	0	2	18	1	0	0
	15	3	0		27	1	0		21	3	0	
14	13	0	0	0	19	3	1	20	16	2	0	0
	11	2	0		15	3	0		21	2	0	
15	31	2	0	1	12	4	0	13	21	4	0	0
	28	3	0		19	4	0		24	0	0	
16	27	5	1	11	15	0	0	2	26	7	1	0
	40	6	1		22	0	0		31	6	2	
17	15	4	0	13	28	6	1	23	26	2	0	1
	10	2	0		29	3	3		25	2	0	
18	28	2	0	8	30	3	0	3	28	2	0	0
	24	4	0		28	5	0		26	0	0	
19	22	3	0	7	16	4	2	5	29	5	0	1
	23	2	1		28	3	1		33	5	0	
20	23	5	1	1	19	5	2	2	27	4	1	22
	23	2	0		25	3	2		26	3	1	
21	32	3	0	14	33	6	1	6	29	3	1	1
	29	3	0		32	0	0		28	5	0	
22	7	1	0	1	27	1	0	3	8	3	0	8
	16	2	0		29	3	0		6	1	0	
23	25	7	0	5	20	3	0	19	12	1	0	0
	18	2	0		18	0	1		18	3	0	
24	21	0	0	6	18	1	0	63	19	1	1	2
	22	5	0		17	4	0		22	2	0	

25	32	4	0	17	22	3	0	2	20	2	1	0
	20	9	0		27	0	0		22	1	1	
26	34	5	0	5	0	0	0	1	23	1	0	1
	34	3	0		0	0	0		23	4	0	
27	17	3	0	6	25	1	0	13	24	2	0	0
	18	1	0		23	4	0		26	3	2	
28	18	0	0	8	20	1	0	3	23	4	0	0
	22	4	0		16	0	0		27	4	1	
29	11	1	0	7	22	5	0	26	27	4	2	15
	19	3	0		20	3	1		31	6	0	
30	25	2	0	103	23	4	0	245	27	2	0	18
	14	1	0		23	4	0		19	5	0	

*Rep.=Replicate for each Biological Replicate

**TNTC = Too Numerous to Count

Appendix D: Raw Data for Screening Assay

D.1 Raw Data for Screening Assay with 96 Wells

Strain	Technical Replicate	Wells containing viable mutants (out of 96)	Wells containing viable mutants (%)	Technical Replicate	Wells containing viable mutants (out of 96)	Wells containing viable mutants (%)
O111:NM (OLC 0682)	1	75	78.13	1	59	61.46
	2	62	64.6	2	63	65.63
	3	76	79.17	3	63	65.63
	4	55	57.29	4	58	60.42
O157:H7 (OLC 0809)	1	19	19.79	1	12	12.5
	2	13	13.54	2	10	10.42
	3	24	25	3	8	8.33
	4	12	12.5	4	15	15.63
O103:H2 (OLC 0969)	1	31	32.29	1	16	16.67
	2	42	43.75	2	13	13.54
	3	26	27.08	3	14	14.58
	4	10	10.42	4	22	22.92
<i>mutS</i>	1	96	100	1	96	100
	2	95	98.96	2	96	100
	3	93	96.88	3	96	100
	4	96	100	4	96	100
<i>yqiG</i>	1	12	12.5	1	18	18.75
	2	5	5.208	2	20	20.83
	3	15	15.63	3	16	16.67
	4	15	15.63	4	19	19.79

D.2 Raw Data for Screening Assay with 24 Wells.

Strain	Biological Replicate	Wells containing viable mutants (out of 24)	Wells containing viable mutants (%)
O111:NM (OLC 0682)	1	20	83.33
	2	18	75
	3	19	79.17
O157:H7 (OLC 0809)	1	3	12.5
	2	5	20.83
	3	2	8.33
O103:H2 (OLC 0969)	1	2	8.33
	2	4	16.67
	3	5	20.83
<i>mutS</i>	1	24	100
	2	24	100
	3	24	100
<i>yqiG</i>	1	6	25
	2	6	25
	3	29.17	33.33

Appendix E: Raw Data for Screening Assay of 100 VTEC Strains

OLC #	Organism	Serotype	Verotoxin Profile	Growth	No Growth	Total	Percentage Growth (%)
455	E. coli	O111:H11	VT1, eae, hlyA	4	20	24	16.67
457	E. coli	O111:H11	VT1, eae, hlyA	7	17	24	29.17
458	E. coli	O111:H8	VT1, eae, hlyA	3	21	24	12.50
461	E. coli	O111:NM	VT1	7	17	24	29.2
464	E. coli	O26:H11	VT1, eae, hlyA	2	22	24	8.33
465	E. coli	O26:H11	VT1, eae, hlyA	2	22	24	8.33
466	E. coli	O26:H11	VT1, eae, hlyA	5	19	24	20.83
469	E. coli	O157:H7	VT1, VT2, eae, hlyA	7	17	24	29.17
470	E. coli	O157:H7	VT2, eae, hlyA	1	23	24	4.17
471	E. coli	O157:H7	VT2, eae, hlyA	2	22	24	8.33
472	E. coli	O157:H7	VT1, VT2, eae, hlyA	5	19	24	20.83
473	E. coli	O157:H7	VT1, VT2, eae, hlyA	0	24	24	0.00
474	E. coli	O157:H7	VT1, VT2, eae, hlyA	2	22	24	8.33
475	E. coli	O157:H7	VT1, VT2, eae, hlyA	1	23	24	4.17
476	E. coli	O157:H7	eae, hlyA	3	21	24	12.50
477	E. coli	O157:H7	VT2, eae, hlyA	0	24	24	0.00

478	E. coli	O157:H7	VT2, eae, hlyA	0	24	24	0.00
479	E. coli	O157:H7	VT1, VT2, eae, hlyA	0	24	24	0.00
480	E. coli	O157:H7	VT1, VT2, eae, hlyA	0	24	24	0.00
481	E. coli	O157:H7	VT1, VT2, eae, hlyA	4	20	24	16.67
482	E. coli	O157:H7	VT1, VT2, eae, hlyA	2	22	24	8.33
483	E. coli	O157:H7	VT1, VT2, eae, hlyA	1	23	24	4.17
484	E. coli	O157:H7	VT2, eae, hlyA	1	23	24	4.17
609	E. coli	O157:H7	vtx1a	2	22	24	8.33
611	E. coli	O26:H11	VT1, eae, hlyA	0	24	24	0.00
622	E. coli	O78:H11	-	0	24	24	0.00
629	E. coli	O111:H11	VT1, eae, hlyA	1	23	24	4.17
630	E. coli	O111:H11	VT1, eae, hlyA	0	24	24	0.00
631	E. coli	O111:H11	VT1, eae, hlyA	2	22	24	8.33
632	E. coli	O111:NM	VT1, VT2, eae	3	21	24	12.50
633	E. coli	O111:NM	VT1, eae, hlyA	3	21	24	12.50
634	E. coli	O111:NM	VT1, eae, hlyA	3	21	24	12.50
635	E. coli	O26:H11	VT1, eae, hlyA	4	20	24	16.67
636	E. coli	O26:H11	VT1, eae, hlyA	3	21	24	12.50
637	E. coli	O26:H11	VT1, eae, hlyA	5	19	24	20.83
638	E. coli	O26:H11	VT1, eae, hlyA	3	21	24	12.50
639	E. coli	O26:H11	VT1, eae, hlyA	1	23	24	4.17
640	E. coli	O26:H11	VT1, eae, hlyA	3	21	24	12.50
667	E. coli	O119:H25	VT1, eae, hlyA	2	22	24	8.33
669	E. coli	O76:H19	VT1	0	24	24	0.00
675	E. coli	O145:NM	VT1, eae, hlyA	1	23	24	4.17
676	E. coli	O145:NM	VT1, eae, hlyA	0	24	24	0.00
677	E. coli	O121:NM	VT2, eae, hlyA	1	23	24	4.17
678	E. coli	O121:H19	VT2, eae, hlyA	2	22	24	8.33
679	E. coli	O103:H2	VT1, eae, hlyA	5	19	24	20.83

680	E. coli	O103:H25	VT1, eae, hlyA	3	21	24	12.50
683	E. coli	O26:H11	eae	1	23	24	4.17
684	E. coli	O145:NM	eae, hlyA	0	24	24	0
686	E. coli	O113:H21	VT2	1	23	24	4.17
707	E. coli	O165:H25	VT2, eae, hlyA	22	2	24	91.67
708	E. coli	O115:H18	VT1, VT2, hlyA	1	23	24	4.17
709	E. coli	O5:NM	VT1, eae, hlyA	2	22	24	8.33
710	E. coli	O121:H19	VT2, eae, hlyA	3	21	24	12.50
711	E. coli	O103:H2	VT1, eae, hlyA	1	23	24	4.17
712	E. coli	O103:H2	VT1, eae, hlyA	2	22	24	8.33
714	E. coli	O111:NM	VT1, eae	0	24	24	0.00
715	E. coli	O111:NM	VT1, VT2, eae	0	24	24	0.00
716	E. coli	O45:H2	VT1, eae, hlyA	8	16	24	33.33
717	E. coli	O55:H7	VT1, eae	2	22	24	8.33
718	E. coli	O157:H7	VT1, VT2, eae, hlyA	0	24	24	0.00
720	E. coli	O91:H21	VT2, hlyA	0	24	24	0.00
725	E. coli	O26:H11	VT1, eae, hlyA	5	19	24	20.83
727	E. coli	O103:H25	VT1, eae, hlyA	2	22	24	8.33
728	E. coli	O103:H11	VT1, eae, hlyA	0	24	24	0.00
731	E. coli	O26:H11	VT1, VT2, eae, hlyA	1	23	24	4.17
733	E. coli	O85:H1	VT2	3	21	24	12.50
789	E. coli	O121:H19	VT2, eae, hlyA	1	23	24	4.17
791	E. coli	O121:NM	VT2, eae, hlyA	0	24	24	0.00
792	E. coli	O121:NM	VT2, eae, hlyA	1	23	24	4.17
797	E. coli	O157:H7	VT1, VT2, eae, hlyA	4	20	24	16.67
816	E. coli	O104:H7	VT2	4	20	24	16.67
982	E. coli	O121:H7	VT1	3	21	24	12.50
983	E. coli	O145	eae	7	17	24	29.17
984	E. coli	O22:H8	VT1 (vtx1c), VT2 (vtx2b), hlyA	24	0	24	100.00
985	E. coli	O145:H34	VT2(vtx2f), eae	0	24	24	0.00
986	E. coli	O157:H7	VT2 (vtx2a), eae, hlyA	1	23	24	4.17
987	E. coli	O103:H2	VT1 (vtx1a), eae, hlyA	1	23	24	4.17
988	E. coli	O91:H21	VT2 (vtx2d), hlyA	6	18	24	25.00

989	E. coli	O107:H2	VT2 (vtx2b, vtx2c, vtx2d)	0	24	24	0.00
990	E. coli	O139:H1	VT2 (vtx2e)	6	18	24	25.00
991	E. coli	O2:H25	VT2 (vtx2g)	2	22	24	8.33
992	E. coli	O154:H31	VT1 (vtx1d)	2	22	24	8.33
993	E. coli	O146:H21	VT1 (vtx1c), VT2 (vtx2a), hlyA	5	19	24	20.83
994	E. coli	O48:H21	VT1 (vtx1a), VT2 (vtx2a), hlyA	1	23	24	4.17
995	E. coli	O174:H21	VT2 (vtx2b, vtx2c)	0	24	24	0.00
996	E. coli	O157:H7	VT1 (vtx1a), VT2 (vtx2a), eae, hlyA	2	22	24	8.33
997	E. coli	O118:H12	VT2 (vtx2b)	0	24	24	0.00
998	E. coli	O73:H18	VT2 (vtx2d)	1	23	24	4.17
999	E. coli	O2:H25	VT2 (vtx2g)	0	24	24	0.00
1000	E. coli	O8:K85ab:H rough	VT1 (vtx1d), eae	7	17	24	29.17
1001	E. coli	O128ac:[H2]	VT2 (vtx2f), eae	1	23	24	4.17
1002	E. coli	O174:H8	VT1 (vtx1c), VT2 (vtx2b)	1	23	24	4.17
1003	E. coli	O139:K12:H1	VT2 (vtx2e)	5	19	24	20.83
1019	E. coli	O45:H2	VT1, eae, hlyA	3	21	24	12.50
1020	E. coli	O45:H2	VT1, eae, hlyA	0	24	24	0.00
1021	E. coli	O45:H2	VT1, eae, hlyA	0	24	24	0.00
1022	E. coli	O45:H2	VT1, eae, hlyA	7	17	24	29.17
1031	E. coli	O121:H19	VT2, eae, hlyA	3	21	24	12.50
1036	E. coli	O145:H25	VT2, eae, hlyA	2	22	24	8.33
1059	E. coli	O157:H7	VT2 (vtx2c), eae, hlyA	4	20	24	16.67

Appendix F: Quality Metrics for WGS Sequences

Strain	Line	N50	Number of Contigs	Total Length	Average Coverage Depth
O111:NM Passage 50	A	28679	489	5105687	34.54
	B	51137	334	5159235	87.21
	C	3973	2745	6509143	14.64
	D	41930	334	5163222	61.32
	E	75173	248	5168026	79.24
O111:NM Passage 100	A	38258	366	5162871	53.69
	B	40088	376	5164543	54.63
	C	52113	305	5182388	71.93
	D	50837	300	5179779	72.1
	E	51138	281	5174322	69.8
O157:H7 Passage 50	A	56541	308	5226450	44.87
	B	80260	233	5236803	63.91
	C	64525	270	5228571	74.66
	D	116460	247	5230642	96.14
	E	51643	327	5219860	48.17
O157:H7 Passage 100	A	65329	292	5238487	48.68
	B	47370	308	5239381	40.36
	C	93753	206	5247352	41.03
	D	67581	265	5242091	66
	E	36063	402	5194568	35.29
O103:H2 Passage 50	A	47309	241	4950932	59.51
	B	108730	128	4969710	85.2
	C	107933	132	4968828	107.78
	D	53797	198	4961182	41.33
	E	46546	224	4957764	62.91
O103:H2 Passage 100	A	54408	208	4965296	39.75
	B	56976	207	4964241	55.44
	C	54875	193	4965108	57.49
	D	106228	113	4976084	58.16
	E	135875	89	4978602	91.05
O165:H25 Passage 10	A	77007	230	4782569	74.55
O22:H8 Passage 10	A	113643	214	5463443	157.52

References

- Aanen, D. K., & Debets, A. (2019). Mutation-rate plasticity and the germline of unicellular organisms. *Proceedings. Biological sciences*, 286(1902), 20190128.
<https://doi.org/10.1098/rspb.2019.0128>
- Adams, C., Vose, A., Edmond, M. B., & Lyckholm, L. (2017). *Shigella sonnei* and hemolytic uremic syndrome: A case report and literature review. *IDCases*, 8, 6–8.
<https://doi.org/10.1016/j.idcr.2017.02.003>
- Alvarez-Ponce, D., Sabater-Muñoz, B., Toft, C., Ruiz-González, M. X., & Fares, M. A. (2016). Essentiality Is a Strong Determinant of Protein Rates of Evolution during Mutation Accumulation Experiments in Escherichia coli. *Genome biology and evolution*, 8(9), 2914–2927. <https://doi.org/10.1093/gbe/evw205>
- An, D., & Parsek, M. R. (2007). The promise and peril of transcriptional profiling in biofilm communities. *Current opinion in microbiology*, 10(3), 292–296.
<https://doi.org/10.1016/j.mib.2007.05.011>
- Andino, A., & Hanning, I. (2015). Salmonella enterica: survival, colonization, and virulence differences among serovars. *TheScientificWorldJournal*, 2015, 520179.
<https://doi.org/10.1155/2015/520179>
- Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data. Available at: www.bioinformatics.babraham.ac.uk/projects/fastqc.
- Asteris, G., & Sarkar, S. (1996). Bayesian procedures for the estimation of mutation rates from fluctuation experiments. *Genetics*, 142(1), 313–326.

- Ayalew, H., Tsang, P. W., Chu, C., Wang, J., Liu, S., Chen, C., & Ma, X. F. (2019). Comparison of TaqMan, KASP and rhAmp SNP genotyping platforms in hexaploid wheat. *PloS one*, 14(5), e0217222. <https://doi.org/10.1371/journal.pone.0217222>
- Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, 2, 2006.0008. <https://doi.org/10.1038/msb4100050>
- Banjo, M., Iguchi, A., Seto, K., Kikuchi, T., Harada, T., Scheutz, F., *et al.* (2018). *Escherichia coli* H-Genotyping PCR: a Complete and Practical Platform for Molecular H Typing. *Journal of clinical microbiology*, 56(6): e00190-18. <https://doi.org/10.1128/JCM.00190-18>
- Belkaid, Y., & Hand, T. W. (2014). Role of the microbiota in immunity and inflammation. *Cell*, 157(1), 121–141. <https://doi.org/10.1016/j.cell.2014.03.011>
- Bell, B. P., Goldoft, M., Griffin, P. M., Davis, M. A., Gordon, D. C., Tarr, P. I., Bartleson, C. A., Lewis, J. H., Barrett, T. J., & Wells, J. G. (1994). A multistate outbreak of Escherichia coli O157:H7-associated bloody diarrhea and hemolytic uremic syndrome from hamburgers. The Washington experience. *JAMA*, 272(17), 1349–1353.
- Beloin, C., Roux, A., & Ghigo, J. M. (2008). Escherichia coli biofilms. *Current topics in microbiology and immunology*, 322, 249–289. https://doi.org/10.1007/978-3-540-75418-3_12
- Beltz, K., Tsang, D., Wang, J.Z., Rose, S., Bao, Y., Wang, Y., Larkin, K., Rupp, S., Schrepfer, D., Datta, K., Gunderson, K., Sailor, C., Hansen, S., Dobosy, J., Lewis, L., Menezes, A., Walder, J., Behlke, M. and Chen, C.F. (2018) A High-Performing and Cost-Effective

- SNP Genotyping Method Using rhPCR and Universal Reporters. *Advances in Bioscience and Biotechnology*, 9, 497-512. <https://doi.org/10.4236/abb.2018.99034>
- Betts, M. & R. Russell. 2003. Amino Acid Properties and Substitution Consequences. In M. Barnes & I. Gray (Eds.), *Bioinformatics for Geneticists* (pp.289-316). England, UK: John Wiley & Sons.
- Beutin, L., Krause, G., Zimmermann, S., Kaulfuss, S., & Gleier, K. (2004). Characterization of Shiga toxin-producing *Escherichia coli* strains isolated from human patients in Germany over a 3-year period. *Journal of clinical microbiology*, 42(3), 1099–1108. <https://doi.org/10.1128/jcm.42.3.1099-1108.2004>
- Bhatnagar, K., Wong, A. (2019). The mutational landscape of quinolone resistance in *Escherichia coli*. *PLoS ONE* 14(11): e0224650. <https://doi.org/10.1371/journal.pone.0224650>
- Blanco, M., Blanco, J. E., Mora, A., Rey, J., Alonso, J. M., Hermoso, M., Hermoso, J., Alonso, M. P., Dahbi, G., González, E. A., Bernárdez, M. I., & Blanco, J. (2003). Serotypes, virulence genes, and intimin types of Shiga toxin (verotoxin)-producing *Escherichia coli* isolates from healthy sheep in Spain. *Journal of clinical microbiology*, 41(4), 1351–1356. <https://doi.org/10.1128/jcm.41.4.1351-1356.2003>
- Broccanello, C., Chiodi, C., Funk, A., McGrath, J. M., Panella, L., & Stevanato, P. (2018). Comparison of three PCR-based assays for SNP genotyping in plants. *Plant methods*, 14, 28. <https://doi.org/10.1186/s13007-018-0295-6>
- Brown, E., Dessai, U., McGarry, S., & Gerner-Smidt, P. (2019). Use of Whole-Genome Sequencing for Food Safety and Public Health in the United States. *Foodborne pathogens and disease*, 16(7), 441–450. <https://doi.org/10.1089/fpd.2019.2662>

- Bryant, J., Chewapreecha, C., & Bentley, S. D. (2012). Developing insights into the mechanisms of evolution of bacterial pathogens from whole-genome sequences. *Future microbiology*, 7(11), 1283–1296. <https://doi.org/10.2217/fmb.12.108>
- Buchholz, U., Bernard, H., Werber, D., Böhmer, M. M., Remschmidt, C., Wilking, H., Deléré, Y., an der Heiden, M., Adlhoch, C., Dreesman, J., Ehlers, J., Ethelberg, S., Faber, M., Frank, C., Fricke, G., Greiner, M., Höhle, M., Ivarsson, S., Jark, U., Kirchner, M., ... Kühne, M. (2011). German outbreak of *Escherichia coli* O104:H4 associated with sprouts. *The New England journal of medicine*, 365(19), 1763–1770. <https://doi.org/10.1056/NEJMoa1106482>
- Bushnell, B. (2014). BBMap: A Fast, Accurate, Splice-Aware Aligner. Available at: <https://sourceforge.net/projects/bbmap/>
- Canpolat N. (2015). Hemolytic uremic syndrome. *Turk pediatri arsivi*, 50(2), 73–82. <https://doi.org/10.5152/tpa.2015.2297>
- Carballo-Pacheco, M., Nicholson, M. D., Lilja, E. E., Allen, R. J., & Waclaw, B. (2020). Phenotypic delay in the evolution of bacterial antibiotic resistance: Mechanistic models and their implications. *PLoS computational biology*, 16(5), e1007930. <https://doi.org/10.1371/journal.pcbi.1007930>
- Carrillo, C. D., Koziol, A. G., Mathews, A., Goji, N., Lambert, D., Huszcynski, G., Gauthier, M., Amoako, K., & Blais, B. W. (2016). Comparative Evaluation of Genomic and Laboratory Approaches for Determination of Shiga Toxin Subtypes in *Escherichia coli*. *Journal of food protection*, 79(12), 2078–2085. <https://doi.org/10.4315/0362-028X.JFP-16-228>.

- Cavaco, L. M., & Aarestrup, F. M. (2009). Evaluation of quinolones for use in detection of determinants of acquired quinolone resistance, including the new transmissible resistance mechanisms qnrA, qnrB, qnrS, and aac(6')Ib-cr, in *Escherichia coli* and *Salmonella enterica* and determinations of wild-type distributions. *Journal of clinical microbiology*, 47(9), 2751–2758. <https://doi.org/10.1128/JCM.00456-09>
- Centers for Disease Control and Prevention. (1996). Outbreak of *Escherichia coli* O157:H7 infections associated with drinking unpasteurized commercial apple juice--British Columbia, California, Colorado, and Washington, October 1996. *MMWR. Morbidity and mortality weekly report*, 45(44), 975.
- Centers for Disease Control and Prevention (CDC) (2013). Outbreak of *Escherichia coli* O104:H4 infections associated with sprout consumption - Europe and North America, May-July 2011. *MMWR. Morbidity and mortality weekly report*, 62(50), 1029–1031.
- Centers for Disease Control and Prevention. (2016a). Pulsed-field gel electrophoresis (PFGE). Retrieved from: <https://www.cdc.gov/pulsenet/pathogens/pfge.html>
- Centers for Disease Control and Prevention. (2016b). Multiple locus variable-number tandem repeat analysis (MLVA). Retrieved from: <https://www.cdc.gov/pulsenet/pathogens/mlva.html>
- Centers for Disease Control and Prevention. (2020). Serotypes and the Importance of Serotyping *Salmonella*. Retrieved from: <https://www.cdc.gov/salmonella/reportspubs/salmonella-atlas/serotyping-importance.html>
- Chattopadhyay, S., Weissman, S. J., Minin, V. N., Russo, T. A., Dykhuizen, D. E., & Sokurenko, E. V. (2009). High frequency of hotspot mutations in core genes of *Escherichia coli* due to short-term positive selection. *Proceedings of the National Academy of Sciences of the*

United States of America, 106(30), 12412–12417.

<https://doi.org/10.1073/pnas.0906217106>

Clinical and Laboratory Standards Institute.(2012). *Methods for Dilution Antimicrobial Susceptibility Tests for Bacteria That Grow Aerobically; Approved Standard—Ninth Edition*. CLSI document M07-A9 (ISBN 1-56238-783-9 [Print]; ISBN 1-56238-784-7 [Electronic]). Clinical and Laboratory Standards Institute, 950 West Valley Road, Suite 2500, Wayne, Pennsylvania 19087, USA.

Clinical and Laboratory Standards Institute (CLSI) (2015). *Methods for Dilution Antimicrobial Susceptibility Tests for Bacteria That Grow Aerobically; Approved Standard*, 10th edn. CLSI document M07-A10. Wayne, PA: Clinical and Laboratory Standards Institute.

Cody, E. M., & Dixon, B. P. (2019). Hemolytic Uremic Syndrome. *Pediatric clinics of North America*, 66(1), 235–246. <https://doi.org/10.1016/j.pcl.2018.09.011>

Conibear, T. C., Collins, S. L., & Webb, J. S. (2009). Role of mutation in *Pseudomonas aeruginosa* biofilm development. *PloS one*, 4(7), e6289. <https://doi.org/10.1371/journal.pone.0006289>

Coombes, B. K., Gilmour, M. W., & Goodman, C. D. (2011). The evolution of virulence in non-o157 shiga toxin-producing *Escherichia coli*. *Frontiers in microbiology*, 2, 90. <https://doi.org/10.3389/fmicb.2011.00090>

Cooper, A. L., Low, A. J., Koziol, A. G., Thomas, M. C., Leclair, D., Tamber, S., Wong, A., Blais, B. W., & Carrillo, C. D. (2020). Systematic Evaluation of Whole Genome Sequence-Based Predictions of *Salmonella* Serotype and Antimicrobial Resistance. *Frontiers in microbiology*, 11, 549. <https://doi.org/10.3389/fmicb.2020.00549>

- Copelovitch, L., & Kaplan, B. S. (2010). Streptococcus pneumoniae--associated hemolytic uremic syndrome: classification and the emergence of serotype 19A. *Pediatrics*, 125(1), e174–e182. <https://doi.org/10.1542/peds.2007-2017>
- Couturier, M. R., Lee, B., Zelyas, N., & Chui, L. (2011). Shiga-toxigenic Escherichia coli detection in stool samples screened for viral gastroenteritis in Alberta, Canada. *Journal of clinical microbiology*, 49(2), 574–578. <https://doi.org/10.1128/JCM.01693-10>
- Currie, A., Galanis, E., Chacon, P. A., Murray, R., Wilcott, L., Kirkby, P., Honish, L., Franklin, K., Farber, J., Parker, R., Shyng, S., Sharma, D., Tschetter, L., Hoang, L., Chui, L., Pacagnella, A., Wong, J., Pritchard, J., Kerr, A., Taylor, M., ... Investigative Team (2018). Outbreak of Escherichia coli O157:H7 Infections Linked to Aged Raw Milk Gouda Cheese, Canada, 2013. *Journal of food protection*, 81(2), 325–331. <https://doi.org/10.4315/0362-028X.JFP-17-283>
- Currie, A., Honish, L., Cutler, J., Locas, A., Lavoie, M. C., Gaulin, C., Galanis, E., Tschetter, L., Chui, L., Taylor, M., Jamieson, F., Gilmour, M., Ng, C., Mutti, S., Mah, V., Hamel, M., Martinez, A., Buenaventura, E., Hoang, L., Pacagnella, A., ... Team, O. (2019). Outbreak of Escherichia coli O157:H7 Infections Linked to Mechanically Tenderized Beef and the Largest Beef Recall in Canada, 2012. *Journal of food protection*, 82(9), 1532–1538. <https://doi.org/10.4315/0362-028X.JFP-19-005>
- DebRoy, C., Fratamico, P. M., Yan, X., Baranzoni, G., Liu, Y., Needleman, D. S., Tebbs, R., O'Connell, C. D., Allred, A., Swimley, M., Mwangi, M., Kapur, V., Raygoza Garay, J. A., Roberts, E. L., & Katani, R. (2016). Comparison of O-Antigen Gene Clusters of All O-Serogroups of Escherichia coli and Proposal for Adopting a New Nomenclature for O-Typing. *PloS one*, 11(1), e0147434. <https://doi.org/10.1371/journal.pone.0147434>

- Deitsch, K. W., Moxon, E. R., & Wellems, T. E. (1997). Shared themes of antigenic variation and virulence in bacterial, protozoal, and fungal infections. *Microbiology and molecular biology reviews : MMBR*, 61(3), 281–293.
- DePas, W. H., Syed, A. K., Sifuentes, M., Lee, J. S., Warshaw, D., Saggar, V., Csankovszki, G., Boles, B. R., & Chapman, M. R. (2014). Biofilm formation protects *Escherichia coli* against killing by *Caenorhabditis elegans* and *Myxococcus xanthus*. *Applied and environmental microbiology*, 80(22), 7079–7087. <https://doi.org/10.1128/AEM.02464-14>
- Devi, S. G., & Ramya, M. (2015). PCR- RFLP based bacterial diversity analysis of a municipal sewage treatment plant. *Journal of environmental biology*, 36(5), 1113–1118.
- Dobosy, J. R., Rose, S. D., Beltz, K. R., Rupp, S. M., Powers, K. M., Behlke, M. A., & Walder, J. A. (2011). RNase H-dependent PCR (rhPCR): improved specificity and single nucleotide polymorphism detection using blocked cleavable primers. *BMC biotechnology*, 11, 80. <https://doi.org/10.1186/1472-6750-11-80>
- Dragosits, M., & Mattanovich, D. (2013). Adaptive laboratory evolution -- principles and applications for biotechnology. *Microbial cell factories*, 12, 64. <https://doi.org/10.1186/1475-2859-12-64>
- Drake, J. (1991). A constant rate of spontaneous mutation in DNA-based microbes. *Proceedings of the National Academy of Sciences*, 88 (16), 7160-7164. <https://doi.org/10.1073/pnas.88.16.7160>
- Driffield, K., Miller, K., Bostock, J. M., O'Neill, A. J., & Chopra, I. (2008). Increased mutability of *Pseudomonas aeruginosa* in biofilms. *The Journal of antimicrobial chemotherapy*, 61(5), 1053–1056. <https://doi.org/10.1093/jac/dkn044>

- Echols, H., Lu, C., & Burgers, P. M. (1983). Mutator strains of *Escherichia coli*, *mutD* and *dnaQ*, with defective exonucleolytic editing by DNA polymerase III holoenzyme. *Proceedings of the National Academy of Sciences of the United States of America*, 80(8), 2189-92.
- Elena, S. F., & de Visser, J. A. (2003). Environmental stress and the effects of mutation. *Journal of biology*, 2(2), 12. <https://doi.org/10.1186/1475-4924-2-12>
- Elias, W. P., Jr, Czeczulin, J. R., Henderson, I. R., Trabulsi, L. R., & Nataro, J. P. (1999). Organization of biogenesis genes for aggregative adherence fimbria II defines a virulence gene cluster in enteroaggregative *Escherichia coli*. *Journal of bacteriology*, 181(6), 1779–1785.
- Fàbrega, A., Madurga, S., Giralt, E., & Vila, J. (2008). Mechanism of action of and resistance to quinolones. *Microbial biotechnology*, 2(1), 40-61.
- Fakhouri, F., Zuber, J., Frémeaux-Bacchi, V., & Loirat, C. (2017). Haemolytic uraemic syndrome. *Lancet (London, England)*, 390(10095), 681–696. [https://doi.org/10.1016/S0140-6736\(17\)30062-4](https://doi.org/10.1016/S0140-6736(17)30062-4)
- Fan, R., Shao, K., Yang, X., Bai, X., Fu, S., Sun, H., Xu, Y., Wang, H., Li, Q., Hu, B., Zhang, J., & Xiong, Y. (2019). High prevalence of non-O157 Shiga toxin-producing *Escherichia coli* in beef cattle detected by combining four selective agars. *BMC microbiology*, 19(1), 213. <https://doi.org/10.1186/s12866-019-1582-8>
- FAO/WHO STEC EXPERT GROUP (2019). Hazard Identification and Characterization: Criteria for Categorizing Shiga Toxin-Producing *Escherichia coli* on a Risk Basis†. *Journal of food protection*, 82(1), 7–21. <https://doi.org/10.4315/0362-028X.JFP-18-291>
- Ferdous, M., Friedrich, A. W., Grundmann, H., de Boer, R. F., Croughs, P. D., Islam, M. A., Kluytmans-van den Bergh, M. F., Kooistra-Smid, A. M., & Rossen, J. W. (2016).

- Molecular characterization and phylogeny of Shiga toxin-producing *Escherichia coli* isolates obtained from two Dutch regions using whole genome sequencing. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases*, 22(7), 642.e1–642.e6429.
<https://doi.org/10.1016/j.cmi.2016.03.028>
- Foster P. L. (2006). Methods for determining spontaneous mutation rates. *Methods in enzymology*, 409, 195–213. [https://doi.org/10.1016/S0076-6879\(05\)09012-9](https://doi.org/10.1016/S0076-6879(05)09012-9)
- Foster P. L. (2007). Stress-induced mutagenesis in bacteria. *Critical reviews in biochemistry and molecular biology*, 42(5), 373–397. <https://doi.org/10.1080/10409230701648494>
- Foster, P. L., Lee, H., Popodi, E., Townes, J. P., & Tang, H. (2015). Determinants of spontaneous mutation in the bacterium *Escherichia coli* as revealed by whole-genome sequencing. *Proceedings of the National Academy of Sciences of the United States of America*, 112(44), E5990–E5999. <https://doi.org/10.1073/pnas.1512136112>
- Fowler, R., R. Schaaper. 1997. The role of the mutT gene of *Escherichia coli* in maintaining replication fidelity. *FEMS Microbiol Rev*, 21(1):43-54.
- Fratamico, P. M., DebRoy, C., Liu, Y., Needleman, D. S., Baranzoni, G. M., & Feng, P. (2016). Advances in Molecular Serotyping and Subtyping of *Escherichia coli*. *Frontiers in microbiology*, 7, 644. <https://doi.org/10.3389/fmicb.2016.00644>
- Friedberg, E.C., Walker, G.C., Siede, W., Wood, R.D., Schultz, R.A., Ellenberger, T. (2006). DNA repair and mutagenesis. ASM, Washington, DC.
<https://doi.org/10.1128/9781555816704>
- Friedrich, A. W., Bielaszewska, M., Zhang, W. L., Pulz, M., Kuczius, T., Ammon, A., & Karch, H. (2002). *Escherichia coli* harboring Shiga toxin 2 gene variants: frequency and

- association with clinical symptoms. *The Journal of infectious diseases*, 185(1), 74–84.
<https://doi.org/10.1086/338115>
- Galhardo, R. S., Do, R., Yamada, M., Friedberg, E. C., Hastings, P. J., Nohmi, T., & Rosenberg, S. M. (2009). DinB upregulation is the sole role of the SOS response in stress-induced mutagenesis in *Escherichia coli*. *Genetics*, 182(1), 55–68.
<https://doi.org/10.1534/genetics.109.100735>
- Gao, X., Yang, X., Noll, L., Shi, X., Worley, J., Allard, M., Brown, E., Nagaraja, T. G., & Meng, J. (2018). Whole genome shotgun sequencing revealed highly polymorphic genome regions and genes in *Escherichia coli* O157:H7 isolates collected from a single feedlot. *PloS one*, 13(8), e0202775. <https://doi.org/10.1371/journal.pone.0202775>
- García-Alcalde, F., K. Okonechnikov, J. Carbonell, L. M. Cruz, S. Götz, S. Tarazona, J. Dopazo, T. F. Meyer, and A. Conesa. (2012). Qualimap: evaluating next-generation sequencing alignment data. *Bioinforma. Oxf. Engl.* 28:2678–2679.
- Gaulin, C., Ramsay, D., Catford, A., & Bekal, S. (2015). *Escherichia coli* O157:H7 Outbreak Associated with the Consumption of Beef and Veal Tartares in the Province of Quebec, Canada, in 2013. *Foodborne pathogens and disease*, 12(7), 612–618.
<https://doi.org/10.1089/fpd.2014.1919>
- Gerrish P. (2008). A simple formula for obtaining markedly improved mutation rate estimates. *Genetics*, 180(3), 1773–1778. <https://doi.org/10.1534/genetics.108.091777>
- Gill, A. (2018). The decline in Shiga toxin-producing *Escherichia coli* illness in Canada: evidence for a role in changes in beef industry practices. Health Canada, Bureau of Microbial Hazards. Retrieved from: https://cmsa-ascv.ca/documents/Gill_CMSA%20News%20May%202018.pdf

- Gill, A., & Gill, C. O. (2010). Non-O157 verotoxigenic *Escherichia coli* and beef: a Canadian perspective. *Canadian journal of veterinary research = Revue canadienne de recherche veterinaire*, 74(3), 161–169.
- Gillespie, S. H., Voelker, L. L., & Dickens, A. (2002). Evolutionary barriers to quinolone resistance in *Streptococcus pneumoniae*. *Microbial drug resistance (Larchmont, N.Y.)*, 8(2), 79–84. <https://doi.org/10.1089/107662902760190617>
- Gillet-Markowska, A., Louvel, G., & Fischer, G. (2015). bz-rates: A Web Tool to Estimate Mutation Rates from Fluctuation Analysis. *G3 (Bethesda, Md.)*, 5(11), 2323–2327. <https://doi.org/10.1534/g3.115.019836>
- Glickman, B., P. van den Elsen, M. Radman.1978. Induced mutagenesis in dam- mutants of *Escherichia coli*: a role for 6-methyladenine residues in mutation avoidance. *Mol. Gen. Genet.* 163(3): 307-312.
- Goldstein B. P. (2014). Resistance to rifampicin: a review. *The Journal of antibiotics*, 67(9), 625–630. <https://doi.org/10.1038/ja.2014.107>
- Goldstein B. P. (2014). Resistance to rifampicin: a review. *The Journal of antibiotics*, 67(9), 625–630. <https://doi.org/10.1038/ja.2014.107>
- Gon, S., Napolitano, R., Rocha, W., Coulon, S., & Fuchs, R. P. (2011). Increase in dNTP pool size during the DNA damage response plays a key role in spontaneous and induced-mutagenesis in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*, 108(48), 19311–19316. <https://doi.org/10.1073/pnas.1113664108>
- Gurevich, A., V. Saveliev, N. Vyahhi, and G. Tesler. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics*, 29:1072–1075.

- Gutierrez, A., Laureti, L., Crussard, S., Abida, H., Rodríguez-Rojas, A., Blázquez, J., Baharoglu, Z., Mazel, D., Darfeuille, F., Vogel, J., & Matic, I. (2013). β -Lactam antibiotics promote bacterial mutagenesis via an RpoS-mediated reduction in replication fidelity. *Nature communications*, 4, 1610. <https://doi.org/10.1038/ncomms2607>
- Guy, L., Jernberg, C., Arvén Norling, J., Ivarsson, S., Hedenström, I., Melefors, Ö., Liljedahl, U., Engstrand, L., & Andersson, S. G. (2013). Adaptive mutations and replacements of virulence traits in the Escherichia coli O104:H4 outbreak population. *PloS one*, 8(5), e63027. <https://doi.org/10.1371/journal.pone.0063027>
- Gyles C. L. (2007). Shiga toxin-producing Escherichia coli: an overview. *Journal of animal science*, 85(13 Suppl), E45–E62. <https://doi.org/10.2527/jas.2006-508>
- Haiko, J., & Westerlund-Wikström, B. (2013). The role of the bacterial flagellum in adhesion and virulence. *Biology*, 2(4), 1242–1267. <https://doi.org/10.3390/biology2041242>
- Hall, B. M., Ma, C. X., Liang, P., & Singh, K. K. (2009). Fluctuation analysis CalculatOR: a web tool for the determination of mutation rate using Luria-Delbruck fluctuation analysis. *Bioinformatics (Oxford, England)*, 25(12), 1564–1565. <https://doi.org/10.1093/bioinformatics/btp253>
- Hebbelstrup Jensen, B., Olsen, K. E., Struve, C., Krogfelt, K. A., & Petersen, A. M. (2014). Epidemiology and clinical manifestations of enteroaggregative Escherichia coli. *Clinical microbiology reviews*, 27(3), 614–630. <https://doi.org/10.1128/CMR.00112-13>
- Heisig, P. (1996). Genetic evidence for a role of parC mutations in development of high-level fluoroquinolone resistance in Escherichia coli. *Antimicrobial agents and chemotherapy*, 40(4), 879–885.

- Helms, M., Simonsen, J., & Mølbak, K. (2006). Foodborne bacterial infection and hospitalization: a registry-based study. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, 42(4), 498–506.
<https://doi.org/10.1086/499813>
- Herikstad, H., Motarjemi, Y., & Tauxe, R. V. (2002). Salmonella surveillance: a global survey of public health serotyping. *Epidemiology and infection*, 129(1), 1–8.
<https://doi.org/10.1017/s0950268802006842>
- Høiby, N., Bjarnsholt, T., Givskov, M., Molin, S., & Ciofu, O. (2010). Antibiotic resistance of bacterial biofilms. *International journal of antimicrobial agents*, 35(4), 322–332.
<https://doi.org/10.1016/j.ijantimicag.2009.12.011>
- Holmes, A., Allison, L., Ward, M., Dallman, T. J., Clark, R., Fawkes, A., Murphy, L., & Hanson, M. (2015). Utility of Whole-Genome Sequencing of Escherichia coli O157 for Outbreak Detection and Epidemiological Surveillance. *Journal of clinical microbiology*, 53(11), 3565–3573. <https://doi.org/10.1128/JCM.01066-15>
- Hussein H. S. (2007). Prevalence and pathogenicity of Shiga toxin-producing Escherichia coli in beef cattle and their products. *Journal of animal science*, 85(13 Suppl), E63–E72.
<https://doi.org/10.2527/jas.2006-421>
- Hussein, H. S., & Omaye, S. T. (2003). Introduction to the food safety concerns of verotoxin-producing Escherichia coli. *Experimental biology and medicine (Maywood, N.J.)*, 228(4), 331–332. <https://doi.org/10.1177/153537020322800401>
- Hyytiä-Trees, E., Smole, S. C., Fields, P. A., Swaminathan, B., & Ribot, E. M. (2006). Second generation subtyping: a proposed PulseNet protocol for multiple-locus variable-number

- tandem repeat analysis of Shiga toxin-producing *Escherichia coli* O157 (STEC O157). *Foodborne pathogens and disease*, 3(1), 118–131. <https://doi.org/10.1089/fpd.2006.3.118>
- Ishizawa, Y., Ying, B. W., Tsuru, S., and Yomo, T. (2015). Nutrient-dependent growth defects and mutability of mutators in *Escherichia coli*. *Genes Cells* 20, 68–76. doi: 10.1111/gtc.12199
- Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nature protocols*, 10(6), 845–858. <https://doi.org/10.1038/nprot.2015.053>
- Jee, J., Rasouly, A., Shamovsky, I., Akivis, Y., Steinman, S. R., Mishra, B., & Nudler, E. (2016). Rates and mechanisms of bacterial mutagenesis from maximum-depth sequencing. *Nature*, 534(7609), 693–696. <https://doi.org/10.1038/nature18313>
- Joensen, K. G., Scheutz, F., Lund, O., Hasman, H., Kaas, R. S., Nielsen, E. M., & Aarestrup, F. M. (2014). Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic *Escherichia coli*. *Journal of clinical microbiology*, 52(5), 1501–1510. <https://doi.org/10.1128/JCM.03617-13>
- Joensen, K. G., Tetzschner, A. M., Iguchi, A., Aarestrup, F. M., & Scheutz, F. (2015). Rapid and Easy In Silico Serotyping of *Escherichia coli* Isolates by Use of Whole-Genome Sequencing Data. *Journal of clinical microbiology*, 53(8), 2410–2426. <https://doi.org/10.1128/JCM.00008-15>
- Johnson, J. R., Murray, A. C., Gajewski, A., Sullivan, M., Snippes, P., Kuskowski, M. A., & Smith, K. E. (2003). Isolation and molecular characterization of nalidixic acid-resistant extraintestinal pathogenic *Escherichia coli* from retail chicken products. *Antimicrobial*

- agents and chemotherapy*, 47(7), 2161–2168. <https://doi.org/10.1128/aac.47.7.2161-2168.2003>
- Jokiranta T. S. (2017). HUS and atypical HUS. *Blood*, 129(21), 2847–2856. <https://doi.org/10.1182/blood-2016-11-709865>
- Jones, T. F., Ingram, L. A., Cieslak, P. R., Vugia, D. J., Tobin-D'Angelo, M., Hurd, S., Medus, C., Cronquist, A., & Angulo, F. J. (2008). Salmonellosis outcomes differ substantially by serotype. *The Journal of infectious diseases*, 198(1), 109–114. <https://doi.org/10.1086/588823>
- Joseph, A., Cointe, A., Mariani Kurkdjian, P., Rafat, C., & Hertig, A. (2020). Shiga Toxin-Associated Hemolytic Uremic Syndrome: A Narrative Review. *Toxins*, 12(2), 67. <https://doi.org/10.3390/toxins12020067>
- Karch H. (2001). The role of virulence factors in enterohemorrhagic Escherichia coli (EHEC)--associated hemolytic-uremic syndrome. *Seminars in thrombosis and hemostasis*, 27(3), 207–213. <https://doi.org/10.1055/s-2001-15250>
- Kang, M., Kim, K., Choe, D., Cho, S., Kim, S. C., Palsson, B., & Cho, B. K. (2019). Inactivation of a Mismatch-Repair System Diversifies Genotypic Landscape of Escherichia coli During Adaptive Laboratory Evolution. *Frontiers in microbiology*, 10, 1845. <https://doi.org/10.3389/fmicb.2019.01845>
- Karch, H., Friedrich, A. W., Gerber, A., Zimmerhackl, L. B., Schmidt, M. A., & Bielaszewska, M. (2006). New aspects in the pathogenesis of enteropathic hemolytic uremic syndrome. *Seminars in thrombosis and hemostasis*, 32(2), 105–112. <https://doi.org/10.1055/s-2006-939766>

- Karczmarczyk, M., Martins, M., Quinn, T., Leonard, N., & Fanning, S. (2011). Mechanisms of fluoroquinolone resistance in *Escherichia coli* isolates from food-producing animals. *Applied and environmental microbiology*, 77(20), 7113–7120.
<https://doi.org/10.1128/AEM.00600-11>
- Karmali M. A. (2009). Host and pathogen determinants of verocytotoxin-producing *Escherichia coli*-associated hemolytic uremic syndrome. *Kidney international. Supplement*, (112), S4–S7. <https://doi.org/10.1038/ki.2008.608>
- Karpman, D., Loos, S., Tati, R., & Arvidsson, I. (2017). Haemolytic uraemic syndrome. *Journal of internal medicine*, 281(2), 123–148. <https://doi.org/10.1111/joim.12546>
- Karpman, D., & Ståhl, A. L. (2014). Enterohemorrhagic *Escherichia coli* Pathogenesis and the Host Response. *Microbiology spectrum*, 2(5), 10.1128/microbiolspec.EHEC-0009-2013. <https://doi.org/10.1128/microbiolspec.EHEC-0009-2013>
- Kau, A. L., Ahern, P. P., Griffin, N. W., Goodman, A. L., & Gordon, J. I. (2011). Human nutrition, the gut microbiome and the immune system. *Nature*, 474(7351), 327–336.
<https://doi.org/10.1038/nature10213>
- Keithlin, J., Sargeant, J. M., Thomas, M. K., & Fazil, A. (2015). Systematic review and meta-analysis of the proportion of non-typhoidal *Salmonella* cases that develop chronic sequelae. *Epidemiology and infection*, 143(7), 1333–1351.
<https://doi.org/10.1017/S0950268814002829>
- Kepler, T. B., & Oprea, M. (2001). Improved inference of mutation rates: I. An integral representation for the Luria-Delbrück distribution. *Theoretical population biology*, 59(1), 41–48. <https://doi.org/10.1006/tpbi.2000.1498>

- Keys, C., Kemper, S., & Keim, P. (2005). Highly diverse variable number tandem repeat loci in the E. coli O157:H7 and O55:H7 genomes for high-resolution molecular typing. *Journal of applied microbiology*, 98(4), 928–940. <https://doi.org/10.1111/j.1365-2672.2004.02532.x>
- Kinnula, S., Hemminki, K., Kotilainen, H., Ruotsalainen, E., Tarkka, E., Salmenlinna, S., Hallanvuo, S., Leinonen, E., Jukka, O., & Rimhanen-Finne, R. (2018). Outbreak of multiple strains of non-O157 Shiga toxin-producing and enteropathogenic *Escherichia coli* associated with rocket salad, Finland, autumn 2016. *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin*, 23(35), 1700666. <https://doi.org/10.2807/1560-7917.ES.2018.23.35.1700666>
- Kondrashov, F. A., & Kondrashov, A. S. (2010). Measurements of spontaneous rates of mutations in the recent past and the near future. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 365(1544), 1169–1176. <https://doi.org/10.1098/rstb.2009.0286>
- Krašovec, R., Belavkin, R. V., Aston, J. A., Channon, A., Aston, E., Rash, B. M., Kadirvel, M., Forbes, S., & Knight, C. G. (2014). Mutation rate plasticity in rifampicin resistance depends on *Escherichia coli* cell-cell interactions. *Nature communications*, 5, 3742. <https://doi.org/10.1038/ncomms4742>
- Kurusu, Y., Narita, T., Suzuki, M., & Watanabe, T. (2000). Genetic analysis of an incomplete mutS gene from *Pseudomonas putida*. *Journal of bacteriology*, 182(18), 5278–5279.
- Lager, M., Mernelius, S., Löfgren, S., & Söderman, J. (2016). Real-Time PCR Typing of *Escherichia coli* Based on Multiple Single Nucleotide Polymorphisms--a Convenient and Rapid Method. *Clinical laboratory*, 62(3), 349–355.

- Lajhar, S. A., Brownlie, J., & Barlow, R. (2018). Characterization of biofilm-forming capacity and resistance to sanitizers of a range of *E. coli* O26 pathotypes from clinical cases and cattle in Australia. *BMC microbiology*, 18(1), 41. <https://doi.org/10.1186/s12866-018-1182-z>
- Lamba, K., Nelson, J. A., Kimura, A. C., Poe, A., Collins, J., Kao, A. S., Cruz, L., Inami, G., Vaishampayan, J., Garza, A., Chaturvedi, V., & Vugia, D. J. (2016). Shiga Toxin 1-Producing *Shigella sonnei* Infections, California, United States, 2014-2015. *Emerging infectious diseases*, 22(4), 679–686. <https://doi.org/10.3201/eid2204.151825>
- Lambert, D., Carrillo, C. D., Koziol, A. G., Manninger, P., & Blais, B. W. (2015). GeneSippr: a rapid whole-genome approach for the identification and characterization of foodborne pathogens such as priority Shiga toxigenic *Escherichia coli*. *PloS one*, 10(4), e0122928. <https://doi.org/10.1371/journal.pone.0122928>
- Landick, R., Stewart, J., & Lee, D. N. (1990). Amino acid changes in conserved regions of the beta-subunit of *Escherichia coli* RNA polymerase alter transcription pausing and termination. *Genes & development*, 4(9), 1623–1636. <https://doi.org/10.1101/gad.4.9.1623>
- Lang, C., Hiller, M., Konrad, R., Fruth, A., & Flieger, A. (2019). Whole-Genome-Based Public Health Surveillance of Less Common Shiga Toxin-Producing *Escherichia coli* Serovars and Untypeable Strains Identifies Four Novel O Genotypes. *Journal of clinical microbiology*, 57(10), e00768-19. <https://doi.org/10.1128/JCM.00768-19>
- Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., &

- Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics (Oxford, England)*, 23(21), 2947–2948. <https://doi.org/10.1093/bioinformatics/btm404>
- Le, K. K., Whiteside, M. D., Hopkins, J. E., Gannon, V., & Laing, C. R. (2018). Spfy: an integrated graph database for real-time prediction of bacterial phenotypes and downstream comparative analyses. *Database : the journal of biological databases and curation*, 2018, 1–10. <https://doi.org/10.1093/database/bay086>
- Lea, D. E., & Coulson, C. A. (1949). The distribution of the numbers of mutants in bacterial populations. *Journal of genetics*, 49(3), 264–285. <https://doi.org/10.1007/BF02986080>
- LeClerc, J. E., Li, B., Payne, W. L., and Cebula, T. A. (1996). High mutation frequencies among *Escherichia coli* and *Salmonella* pathogens. *Science* 274, 1208–1211. doi: 10.1126/science.274.5290.1208
- Lee, H., Popodi, E., Tang, H., & Foster, P. L. (2012). Rate and molecular spectrum of spontaneous mutations in the bacterium *Escherichia coli* as determined by whole-genome sequencing. *Proceedings of the National Academy of Sciences of the United States of America*, 109(41), E2774–E2783. doi:10.1073/pnas.1210309109
- Legislative Services Branch. (2019, February 04). Consolidated federal laws of Canada, Safe Food for Canadians Regulations. Retrieved from <https://laws-lois.justice.gc.ca/eng/regulations/SOR-2018-108/>
- Li, B., Liu, H., & Wang, W. (2017). Multiplex real-time PCR assay for detection of *Escherichia coli* O157:H7 and screening for non-O157 Shiga toxin-producing *E. coli*. *BMC microbiology*, 17(1), 215. <https://doi.org/10.1186/s12866-017-1123-2>
- Li, J., Hao, H., Dai, M., Zhang, H., Ning, J., Cheng, G., Shabbir, M., Sajid, A., & Yuan, Z. (2019a). Resistance and Virulence Mechanisms of *Escherichia coli* Selected by

- Enrofloxacin in Chicken. *Antimicrobial agents and chemotherapy*, 63(5), e01824-18.
<https://doi.org/10.1128/AAC.01824-18>
- Li, M., Liu, Q., Teng, Y., Ou, L., Xi, Y., Chen, S., & Duan, G. (2019b). The resistance mechanism of *Escherichia coli* induced by ampicillin in laboratory. *Infection and drug resistance*, 12, 2853–2863. <https://doi.org/10.2147/IDR.S221212>
- Li, W., Raoult, D., & Fournier, P. E. (2009). Bacterial strain typing in the genomic era. *FEMS microbiology reviews*, 33(5), 892–916. <https://doi.org/10.1111/j.1574-6976.2009.00182.x>
- Lisboa, L. F., Szelewicki, J., Lin, A., Latonas, S., Li, V., Zhi, S., Parsons, B. D., Berenger, B., Fathima, S., & Chui, L. (2019). Epidemiology of Shiga Toxin-Producing *Escherichia coli* O157 in the Province of Alberta, Canada, 2009-2016. *Toxins*, 11(10), 613.
<https://doi.org/10.3390/toxins11100613>
- López, D., Vlamakis, H., & Kolter, R. (2010). Biofilms. *Cold Spring Harbor perspectives in biology*, 2(7), a000398. <https://doi.org/10.1101/cshperspect.a000398>
- Low, A. J., A. G. Koziol, P. A. Manninger, B. Blais, and C. D. Carrillo. (2019). ConFindr: rapid detection of intraspecies and cross-species contamination in bacterial whole-genome sequence data. *PeerJ*, 7:e6995.
- Lugtenberg, B. (1981). Composition and function of the outer membrane of *Escherichia coli*. *Trends in Biological Sciences*, 6: 262-266. [https://doi.org/10.1016/0968-0004\(81\)90095-5](https://doi.org/10.1016/0968-0004(81)90095-5)
- Lukačišinová, M., Novak, S., & Paixão, T. (2017). Stress-induced mutagenesis: Stress diversity facilitates the persistence of mutator genes. *PLoS computational biology*, 13(7), e1005609. <https://doi.org/10.1371/journal.pcbi.1005609>

- Luria, S. E., & Delbrück, M. (1943). Mutations of Bacteria from Virus Sensitivity to Virus Resistance. *Genetics*, 28(6), 491–511.
- Lynch, M., Ackerman, M. S., Gout, J. F., Long, H., Sung, W., Thomas, W. K., & Foster, P. L. (2016). Genetic drift, selection and the evolution of the mutation rate. *Nature reviews. Genetics*, 17(11), 704–714. <https://doi.org/10.1038/nrg.2016.104>
- Madigan, M., J. Martinko, K. Bender, D. Buckley, and D. Stahl. 2015. *Brock Biology of Microorganisms* (14th ed.). Boston:Pearson.
- Maharjan, R. P., & Ferenci, T. (2018). The impact of growth rate and environmental factors on mutation rates and spectra in Escherichia coli. *Environmental microbiology reports*, 10(6), 626–633. <https://doi.org/10.1111/1758-2229.12661>
- Marchler-Bauer, A. *et al.* 2017. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res.*, 45(D)200-203.
- Martincorena, I., Seshasayee, A. S., & Luscombe, N. M. (2012). Evidence of non-random mutation rates suggests an evolutionary risk management strategy. *Nature*, 485(7396), 95–98. <https://doi.org/10.1038/nature10995>
- Martinez-Medina, M., Naves, P., Blanco, J., Aldeguer, X., Blanco, J. E., Blanco, M., Ponte, C., Soriano, F., Darfeuille-Michaud, A., & Garcia-Gil, L. J. (2009). Biofilm formation as a novel phenotypic feature of adherent-invasive Escherichia coli (AIEC). *BMC microbiology*, 9, 202. <https://doi.org/10.1186/1471-2180-9-202>
- Matic I. (2019). Mutation Rate Heterogeneity Increases Odds of Survival in Unpredictable Environments. *Molecular cell*, 75(3), 421–425. <https://doi.org/10.1016/j.molcel.2019.06.029>

- Maury, M. M., Tsai, Y. H., Charlier, C., Touchon, M., Chenal-Francisque, V., Leclercq, A., Criscuolo, A., Gaultier, C., Roussel, S., Brisabois, A., Disson, O., Rocha, E., Brisse, S., & Lecuit, M. (2016). Uncovering *Listeria monocytogenes* hypervirulence by harnessing its biodiversity. *Nature genetics*, 48(3), 308–313. <https://doi.org/10.1038/ng.3501>
- Mazoyer, A., Drouilhet, R., Despréaux, S., & Ycart, B. (2017). flan: An R Package for Inference on Mutation Models. *R Journal*, 9, 334-351. <https://doi.org/10.32614/RJ-2017-029>.
- McKenzie, G. J., Magner, D. B., Lee, P. L., & Rosenberg, S. M. (2003). The dinB operon and spontaneous mutation in *Escherichia coli*. *Journal of bacteriology*, 185(13), 3972-7.
- Melton-Celsa A. R. (2014). Shiga Toxin (Stx) Classification, Structure, and Function. *Microbiology spectrum*, 2(4), 10.1128/microbiolspec.EHEC-0024-2013. <https://doi.org/10.1128/microbiolspec.EHEC-0024-2013>
- Miller, K., O'Neill, A. J., & Chopra, I. (2004). *Escherichia coli* mutators present an enhanced risk for emergence of antibiotic resistance during urinary tract infections. *Antimicrobial agents and chemotherapy*, 48(1), 23–29. <https://doi.org/10.1128/aac.48.1.23-29.2004>
- Morin, N., Santiago, A. E., Ernst, R. K., Guillot, S. J., & Nataro, J. P. (2013). Characterization of the AggR regulon in enteroaggregative *Escherichia coli*. *Infection and immunity*, 81(1), 122–132. <https://doi.org/10.1128/IAI.00676-12>
- Morton, V. K., Kearney, A., Coleman, S., Viswanathan, M., Chau, K., Orr, A., & Hexemer, A. (2019). Outbreaks of *Salmonella* illness associated with frozen raw breaded chicken products in Canada, 2015-2019. *Epidemiology and infection*, 147, e254. <https://doi.org/10.1017/S0950268819001432>
- Mount D. W. (2008). Using PAM Matrices in Sequence Alignments. *CSH protocols*, 2008, pdb.top38. <https://doi.org/10.1101/pdb.top38>

- Nadon, C., Van Walle, I., Gerner-Smidt, P., Campos, J., Chinen, I., Concepcion-Acevedo, J., Gilpin, B., Smith, A. M., Man Kam, K., Perez, E., Trees, E., Kubota, K., Takkinen, J., Nielsen, E. M., Carleton, H., & FWD-NEXT Expert Panel (2017). PulseNet International: Vision for the implementation of whole genome sequencing (WGS) for global food-borne disease surveillance. *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin*, 22(23), 30544. <https://doi.org/10.2807/1560-7917.ES.2017.22.23.30544>
- Nakano, R., Okamoto, R., Nakano, A., Nagano, N., Abe, M., Tansho-Nagakawa, S., Ubagai, T., Kikuchi-Ueda, T., Koshio, O., Kikuchi, H., & Ono, Y. (2013). Rapid assay for detecting *gyrA* and *parC* mutations associated with fluoroquinolone resistance in Enterobacteriaceae. *Journal of microbiological methods*, 94(3), 213–216. <https://doi.org/10.1016/j.mimet.2013.06.019>
- Nakao, H., & Takeda, T. (2000). Escherichia coli Shiga toxin. *Journal of natural toxins*, 9(3), 299–313.
- National Institute of Health. (2014). Foodborne Illnesses. Retrieved from <https://www.niddk.nih.gov/health-information/digestive-diseases/foodborne-illnesses>
- Nesbitt, A., Ravel, A., Murray, R., McCormick, R., Savelli, C., Finley, R., Parmley, J., Agunos, A., Majowicz, S. E., Gilmour, M., Canadian Integrated Program for Antimicrobial Resistance Surveillance Public Health Partnership, & Canadian Public Health Laboratory Network (2012). Integrated surveillance and potential sources of *Salmonella enteritidis* in human cases in Canada from 2003 to 2009. *Epidemiology and infection*, 140(10), 1757–1772. <https://doi.org/10.1017/S0950268811002548>

- Ninio J. (1996). Gene conversion as a focusing mechanism for correlated mutations: a hypothesis. *Molecular & general genetics : MGG*, 251(5), 503–508.
<https://doi.org/10.1007/BF02173638>
- Noller, A. C., McEllistrem, M. C., Pacheco, A. G., Boxrud, D. J., & Harrison, L. H. (2003). Multilocus variable-number tandem repeat analysis distinguishes outbreak and sporadic *Escherichia coli* O157:H7 isolates. *Journal of clinical microbiology*, 41(12), 5389–5397.
<https://doi.org/10.1128/jcm.41.12.5389-5397.2003>
- Noris, M., & Remuzzi, G. (2009). Atypical hemolytic-uremic syndrome. *The New England journal of medicine*, 361(17), 1676–1687. <https://doi.org/10.1056/NEJMra0902814>
- Notley-McRobb, L., Pinto, R., Seeto, S., & Ferenci, T. (2002). Regulation of mutY and nature of mutator mutations in *Escherichia coli* populations under nutrient limitation. *Journal of bacteriology*, 184(3), 739-45.
- Okhravi, N., Adamson, P., Matheson, M. M., Towler, H. M., & Lightman, S. (2000). PCR-RFLP-mediated detection and speciation of bacterial species causing endophthalmitis. *Investigative ophthalmology & visual science*, 41(6), 1438–1447.
- Okonechnikov, K., A. Conesa, and F. García-Alcalde. (2016). Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinforma. Oxf. Engl.* 32:292–294.
- Oliver, D. H., Thompson, R. E., Griffin, C. A., & Eshleman, J. R. (2000). Use of single nucleotide polymorphisms (SNP) and real-time polymerase chain reaction for bone marrow engraftment analysis. *The Journal of molecular diagnostics : JMD*, 2(4), 202–208. [https://doi.org/10.1016/S1525-1578\(10\)60638-1](https://doi.org/10.1016/S1525-1578(10)60638-1)

- Onda, M., Hanada, K., Kawachi, H., & Ikeda, H. (1999). *Escherichia coli* mutM suppresses illegitimate recombination induced by oxidative stress. *Genetics*, 151(2), 439-46.
- Orskov, F., & Orskov, I. (1992). *Escherichia coli* serotyping and disease in man and animals. *Canadian journal of microbiology*, 38(7), 699–704.
- Petkau, A., Mabon, P., Sieffert, C., Knox, N. C., Cabral, J., Iskander, M., Iskander, M., Weedmark, K., Zaheer, R., Katz, L. S., Nadon, C., Reimer, A., Taboada, E., Beiko, R. G., Hsiao, W., Brinkman, F., Graham, M., & Van Domselaar, G. (2017). SNVPhyl: a single nucleotide variant phylogenomics pipeline for microbial genomic epidemiology. *Microbial genomics*, 3(6), e000116. <https://doi.org/10.1099/mgen.0.000116>
- Pianciola, L., & Rivas, M. (2018). Genotypic Features of Clinical and Bovine *Escherichia coli* O157 Strains Isolated in Countries with Different Associated-Disease Incidences. *Microorganisms*, 6(2), 36. <https://doi.org/10.3390/microorganisms6020036>
- Pizarro-Cerdá, J., & Cossart, P. (2019). Microbe Profile: *Listeria monocytogenes*: a paradigm among intracellular bacterial pathogens. *Microbiology (Reading, England)*, 165(7), 719–721. <https://doi.org/10.1099/mic.0.000800>
- Prager, R., Strutz, U., Fruth, A., & Tschäpe, H. (2003). Subtyping of pathogenic *Escherichia coli* strains using flagellar (H)-antigens: serotyping versus fliC polymorphisms. *International journal of medical microbiology : IJMM*, 292(7-8), 477–486. <https://doi.org/10.1078/1438-4221-00226>
- Public Health Agency of Canada (PHAC). (2010). Canada’s Food-borne Illness Outbreak Response Protocol (FIORP) 2010: To guide a multi-jurisdictional response. Retrieved from: <http://www.phac-aspc.gc.ca/zoono/fiorp-mitioa/index-eng.php>

- Public Health Agency of Canada (PHAC). (2016a). Food-related illnesses, hospitalizations and deaths in Canada. Retrieved from: <https://www.canada.ca/en/public-health/services/publications/food-nutrition/infographic-food-related-illnesses-hospitalizations-deaths-in-canada.html>
- Public Health Agency of Canada (PHAC). (2016b). Salmonellosis (Salmonella). Retrieved from: <https://www.canada.ca/en/public-health/services/diseases/salmonellosis-salmonella.html>
- Public Health Agency of Canada (PHAC). (2016c). Listeriosis (Listeria). Retrieved from: <https://www.canada.ca/en/public-health/services/diseases/listeriosis.html>
- Public Health Agency of Canada (PHAC). (2018). FoodNet Canada annual report 2017. Retrieved from: <https://www.canada.ca/en/public-health/services/surveillance/foodnet-canada/publications/foodnet-canada-annual-report-2017.html>
- Public Health Agency of Canada (PHAC). (2019). FoodNet Canada annual report 2018. Retrieved from: <https://www.canada.ca/en/public-health/services/surveillance/foodnet-canada/publications/foodnet-canada-annual-report-2018.html#info>
- Public Health Agency of Canada (PHAC). 2020. Public Health Notice — Outbreak of Salmonella illnesses linked to raw turkey and raw chicken. Retrieved from: <https://www.canada.ca/en/public-health/services/public-health-notices/2018/outbreak-salmonella-illnesses-raw-turkey-raw-chicken.html#a7>
- Pacheco, A. R., & Sperandio, V. (2012). Shiga toxin in enterohemorrhagic E.coli: regulation and novel anti-virulence strategies. *Frontiers in cellular and infection microbiology*, 2, 81. <https://doi.org/10.3389/fcimb.2012.00081>

- Pope, C. F., O'Sullivan, D. M., McHugh, T. D., & Gillespie, S. H. (2008). A practical guide to measuring mutation rates in antibiotic resistance. *Antimicrobial agents and chemotherapy*, 52(4), 1209–1214. <https://doi.org/10.1128/AAC.01152-07>
- Public Health Agency of Canada (PHAC). (2017). Canada's Foodborne Illness Outbreak Response Protocol (FIORP): A guide to multi-jurisdictional enteric outbreak response. Public Health Agency of Canada, Ottawa, ON. Retrieved from: <https://www.canada.ca/en/public-health/services/publications/health-risks-safety/canadas-foodborne-illness-outbreak-response-protocol-fiorp-guide-multi-jurisdictional-enteric-outbreak-response.html>
- Public Health Agency of Canada (PHAC). (2018). Foodnet Canada Annual Report 2016. Public Health Agency of Canada, Ottawa, ON. Retrieved from: <https://www.canada.ca/en/public-health/services/surveillance/foodnet-canada/publications/foodnet-canada-annual-report-2017.html>
- Public Health Agency of Canada (PHAC). (2016). Yearly food-borne illness estimates for Canada. Retrieved from: <https://www.canada.ca/en/public-health/services/food-borne-illness-canada/yearly-food-borne-illness-estimates-canada.html>
- Ramiro, R.S., Durão, P., Bank, C., Gordo, I. (2020). Low mutational load and high mutation rate variation in gut commensal bacteria. *PLoS Biol* 18(3): e3000617. <https://doi.org/10.1371/journal.pbio.3000617>
- Ribot, E. M., Freeman, M., Hise, K. B., & Gerner-Smidt, P. (2019). PulseNet: Entering the Age of Next-Generation Sequencing. *Foodborne pathogens and disease*, 16(7), 451–456. <https://doi.org/10.1089/fpd.2019.2634>

- Romero, P. A., & Arnold, F. H. (2009). Exploring protein fitness landscapes by directed evolution. *Nature reviews. Molecular cell biology*, 10(12), 866–876.
<https://doi.org/10.1038/nrm2805>
- Rowe, P. C., Orrbine, E., Lior, H., Wells, G. A., Yetisir, E., Clulow, M., & McLaine, P. N. (1998). Risk of hemolytic uremic syndrome after sporadic *Escherichia coli* O157:H7 infection: results of a Canadian collaborative study. Investigators of the Canadian Pediatric Kidney Disease Research Center. *The Journal of pediatrics*, 132(5), 777–782.
[https://doi.org/10.1016/s0022-3476\(98\)70303-8](https://doi.org/10.1016/s0022-3476(98)70303-8)
- Rumore, J., Tschetter, L., Kearney, A., Kandar, R., McCormick, R., Walker, M., Peterson, C. L., Reimer, A., & Nadon, C. (2018). Evaluation of whole-genome sequencing for outbreak detection of Verotoxigenic *Escherichia coli* O157:H7 from the Canadian perspective. *BMC genomics*, 19(1), 870. <https://doi.org/10.1186/s12864-018-5243-3>
- Rusconi, B., Sanjar, F., Koenig, S. S., Mammel, M. K., Tarr, P. I., & Eppinger, M. (2016). Whole Genome Sequencing for Genomics-Guided Investigations of *Escherichia coli* O157:H7 Outbreaks. *Frontiers in microbiology*, 7, 985.
<https://doi.org/10.3389/fmicb.2016.00985>
- Ryder, V. J., Chopra, I., & O'Neill, A. J. (2012). Increased mutability of *Staphylococci* in biofilms as a consequence of oxidative stress. *PloS one*, 7(10), e47695.
<https://doi.org/10.1371/journal.pone.0047695>
- Sáenz, Y., Zarazaga, M., Briñas, L., Ruiz-Larrea, F., & Torres, C. (2003). Mutations in *gyrA* and *parC* genes in nalidixic acid-resistant *Escherichia coli* strains from food products, humans and animals. *The Journal of antimicrobial chemotherapy*, 51(4), 1001–1005.
<https://doi.org/10.1093/jac/dkg168>

- Salvadori, M., & Bertoni, E. (2013). Update on hemolytic uremic syndrome: Diagnostic and therapeutic recommendations. *World journal of nephrology*, 2(3), 56–76.
<https://doi.org/10.5527/wjn.v2.i3.56>
- Scheuermann, R., Tam, S., Burgers, P. M., Lu, C., & Echols, H. (1983). Identification of the epsilon-subunit of Escherichia coli DNA polymerase III holoenzyme as the dnaQ gene product: a fidelity subunit for DNA replication. *Proceedings of the National Academy of Sciences of the United States of America*, 80(23), 7085-9.
- Schiebel, J., Böhm, A., Nitschke, J., Burdukiewicz, M., Weinreich, J., Ali, A., Roggenbuck, D., Rödiger, S., & Schierack, P. (2017). Genotypic and Phenotypic Characteristics Associated with Biofilm Formation by Human Clinical Escherichia coli Isolates of Different Pathotypes. *Applied and environmental microbiology*, 83(24), e01660-17.
<https://doi.org/10.1128/AEM.01660-17>
- Severinov, K., Soushko, M., Goldfarb, A. & Nikiforov, V. (1993). Rifampicin region revisited. New rifampicin-resistant and streptolydigin-resistant mutants in the β subunit of *Escherichia coli* RNA polymerase. *J. Biol. Chem.* 28, 14820–14825.
- Shamloo, E., Hosseini, H., Abdi Moghadam, Z., Halberg Larsen, M., Haslberger, A., & Alebouyeh, M. (2019). Importance of *Listeria monocytogenes* in food safety: a review of its prevalence, detection, and antibiotic resistance. *Iranian journal of veterinary research*, 20(4), 241–254.
- Sharma, V. K., Akavaram, S., Schaut, R. G., & Bayles, D. O. (2019). Comparative genomics reveals structural and functional features specific to the genome of a foodborne Escherichia coli O157:H7. *BMC genomics*, 20(1), 196. <https://doi.org/10.1186/s12864-019-5568-6>

- Shastri B. S. (2009). SNPs: impact on gene function and phenotype. *Methods in molecular biology (Clifton, N.J.)*, 578, 3–22. https://doi.org/10.1007/978-1-60327-411-1_1
- Shee, C., Gibson, J. L., & Rosenberg, S. M. (2012). Two mechanisms produce mutation hotspots at DNA breaks in *Escherichia coli*. *Cell reports*, 2(4), 714–721. <https://doi.org/10.1016/j.celrep.2012.08.033>
- Sheerin, N. S., & Glover, E. (2019). Haemolytic uremic syndrome: diagnosis and management. *F1000Research*, 8, F1000 Faculty Rev-1690. <https://doi.org/10.12688/f1000research.19957.1>
- Shefer, A. M., Koo, D., Werner, S. B., Mintz, E. D., Baron, R., Wells, J. G., Barrett, T. J., Ginsberg, M., Bryant, R., Abbott, S., & Griffin, P. M. (1996). A cluster of *Escherichia coli* O157:H7 infections with the hemolytic-uremic syndrome and death in California. A mandate for improved surveillance. *The Western journal of medicine*, 165(1-2), 15–19.
- Sjölund, M., Bengtsson, S., Bonnedahl, J., Hernandez, J., Olsen, B., & Kahlmeter, G. (2009). Antimicrobial susceptibility in *Escherichia coli* of human and avian origin--a comparison of wild-type distributions. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases*, 15(5), 461–465. <https://doi.org/10.1111/j.1469-0691.2009.02705.x>
- Sneath P. H. (1966). Relations between chemical structure and biological activity in peptides. *Journal of theoretical biology*, 12(2), 157–195. [https://doi.org/10.1016/0022-5193\(66\)90112-3](https://doi.org/10.1016/0022-5193(66)90112-3)
- Souvorov, A., R. Agarwala, and D. J. Lipman. (2018). SKESA: strategic k-mer extension for scrupulous assemblies. *Genome Biol.* 19.

- Spinale, J. M., Ruebner, R. L., Kaplan, B. S., & Copelovitch, L. (2013). Update on *Streptococcus pneumoniae* associated hemolytic uremic syndrome. *Current opinion in pediatrics*, 25(2), 203–208. <https://doi.org/10.1097/MOP.0b013e32835d7f2c>
- Sprouffske, K., Aguilar-Rodríguez, J., Sniegowski, P., & Wagner, A. (2018). High mutation rates limit evolutionary adaptation in *Escherichia coli*. *PLoS genetics*, 14(4), e1007324. <https://doi.org/10.1371/journal.pgen.1007324>
- Strauss, B. S., Roberts, R., Francis, L., & Pouryazdanparast, P. (2000). Role of the *dinB* gene product in spontaneous mutation in *Escherichia coli* with an impaired replicative polymerase. *Journal of bacteriology*, 182(23), 6742–6750. <https://doi.org/10.1128/jb.182.23.6742-6750.2000>
- Sun, L., Alexander, H. K., Bogos, B., Kiviet, D. J., Ackermann, M., & Bonhoeffer, S. (2018). Effective polyploidy causes phenotypic delay and influences bacterial evolvability. *PLoS biology*, 16(2), e2004644. <https://doi.org/10.1371/journal.pbio.2004644>
- Swaminathan, B., Barrett, T. J., Hunter, S. B., Tauxe, R. V., & CDC PulseNet Task Force (2001). PulseNet: the molecular subtyping network for foodborne bacterial disease surveillance, United States. *Emerging infectious diseases*, 7(3), 382–389. <https://doi.org/10.3201/eid0703.010303>
- Swings, T., Van den Bergh, B., Wuyts, S., Oeyen, E., Voordeckers, K., Verstrepen, K. J., Fauvart, M., Verstraeten, N., & Michiels, J. (2017). Adaptive tuning of mutation rates allows fast response to lethal stress in *Escherichia coli*. *eLife*, 6, e22939. <https://doi.org/10.7554/eLife.22939>

- Tang, F., Wang, J., Li, D., Gao, S., Ren, J., Ma, L., Liu, F., Zhuge, X., Yan, G., Lu, Y., & Dai, J. (2019). Comparative genomic analysis of 127 *Escherichia coli* strains isolated from domestic animals with diarrhea in China. *BMC genomics*, 20(1), 212. <https://doi.org/10.1186/s12864-019-5588-2>
- Tang, M., Pham, P., Shen, X., Taylor, J. S., O'Donnell, M., Woodgate, R., & Goodman, M. F. (2000). Roles of *E. coli* DNA polymerases IV and V in lesion-targeted and untargeted SOS mutagenesis. *Nature*, 404(6781), 1014–1018. <https://doi.org/10.1038/35010020>
- Tataryn, J., Morton, V., Cutler, J., McDonald, L., Whitfield, Y., Billard, B., Gad, R. R., & Hexemer, A. (2014). Outbreak of *E. coli* O157:H7 associated with lettuce served at fast food chains in the Maritimes and Ontario, Canada, Dec 2012. *Canada communicable disease report = Releve des maladies transmissibles au Canada*, 40(Suppl 1), 2–9. <https://doi.org/10.14745/ccdr.v40is1a01>
- Taylor, M. (2008). Enterohaemorrhagic *Escherichia coli* and *Shigella dysenteriae* type 1-induced haemolytic uraemic syndrome. *Pediatric nephrology (Berlin, Germany)*, 23(9), 1425–1431. <https://doi.org/10.1007/s00467-008-0820-3>
- Tenaillon, O., Skurnik, D., Picard, B., & Denamur, E. (2010). The population genetics of commensal *Escherichia coli*. *Nature reviews. Microbiology*, 8(3), 207–217. <https://doi.org/10.1038/nrmicro2298>
- Ternhag, A., Törner, A., Svensson, A., Ekdahl, K., & Giesecke, J. (2008). Short- and long-term effects of bacterial gastrointestinal infections. *Emerging infectious diseases*, 14(1), 143–148. <https://doi.org/10.3201/eid1401.070524>

- Thomas, M. K., Murray, R., Nesbitt, A., & Pollari, F. (2017). The Incidence of Acute Gastrointestinal Illness in Canada, Foodbook Survey 2014-2015. *The Canadian journal of infectious diseases & medical microbiology = Journal canadien des maladies infectieuses et de la microbiologie medicale*, 2017, 5956148.
<https://doi.org/10.1155/2017/5956148>
- Timme, R. E., Rand, H., Sanchez Leon, M., Hoffmann, M., Strain, E., Allard, M., Roberson, D., & Baugher, J. D. (2018). GenomeTrakr proficiency testing for foodborne pathogen surveillance: an exercise from 2015. *Microbial genomics*, 4(7), e000185.
<https://doi.org/10.1099/mgen.0.000185>
- Timme, R. E., Sanchez Leon, M., & Allard, M. W. (2019). Utilizing the Public GenomeTrakr Database for Foodborne Pathogen Traceback. *Methods in molecular biology (Clifton, N.J.)*, 1918, 201–212. https://doi.org/10.1007/978-1-4939-9000-9_17
- Totten, S., Medaglia, A., & McDermott, S. (2019). Updates to the Canadian Notifiable Disease Surveillance System and its interactive website. *Canada communicable disease report = Releve des maladies transmissibles au Canada*, 45(10), 4257–4261.
<https://doi.org/10.14745/ccdr.v45i10a02>
- Tsujimoto, H., Gotoh, N., & Nishino, T. (1999). Diffusion of macrolide antibiotics through the outer membrane of *Moraxella catarrhalis*. *Journal of infection and chemotherapy : official journal of the Japan Society of Chemotherapy*, 5(4), 196–200.
<https://doi.org/10.1007/s101560050034>
- Turrientes, M. C., Baquero, F., Levin, B. R., Martínez, J. L., Ripoll, A., González-Alba, J. M., Tobes, R., Manrique, M., Baquero, M. R., Rodríguez-Domínguez, M. J., Cantón, R., &

- Galán, J. C. (2013). Normal mutation rate variants arise in a Mutator (Mut S) *Escherichia coli* population. *PloS one*, 8(9), e72963. <https://doi.org/10.1371/journal.pone.0072963>
- Uphoff, S., Lord, N. D., Okumus, B., Potvin-Trottier, L., Sherratt, D. J., & Paulsson, J. (2016). Stochastic activation of a DNA damage response causes cell-to-cell mutation rate variation. *Science (New York, N.Y.)*, 351(6277), 1094–1097. <https://doi.org/10.1126/science.aac9786>
- Valilis, E., Ramsey, A., Sidiq, S., & DuPont, H. L. (2018). Non-O157 Shiga toxin-producing *Escherichia coli*-A poorly appreciated enteric pathogen: Systematic review. *International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases*, 76, 82–87. <https://doi.org/10.1016/j.ijid.2018.09.002>
- Vanderkooi, O. G., Kellner, J. D., Wade, A. W., Jadavji, T., Midgley, J. P., Louie, T., & Tyrell, G. J. (2003). Invasive *Streptococcus pneumoniae* infection causing hemolytic uremic syndrome in children: Two recent cases. *The Canadian journal of infectious diseases = Journal canadien des maladies infectieuses*, 14(6), 339–343. <https://doi.org/10.1155/2003/219027>
- Wagner J., Gruz P., Kim S.-R., Yamada M., Matsui K., Fuchs R.P.P., Nohmi T. 1999. The *dinB* gene encodes a novel *E. coli* DNA polymerase, DNA pol IV, involved in mutagenesis. *Mol Cell.*, 4(2):281-286.
- Wang, L., Rothmund, D., Curd, H., & Reeves, P. R. (2003). Species-wide variation in the *Escherichia coli* flagellin (H-antigen) gene. *Journal of bacteriology*, 185(9), 2936–2943. <https://doi.org/10.1128/jb.185.9.2936-2943.2003>
- Ward, T. J., Gorski, L., Borucki, M. K., Mandrell, R. E., Hutchins, J., & Pupedis, K. (2004). Intraspecific phylogeny and lineage group identification based on the *prfA* virulence gene

- cluster of *Listeria monocytogenes*. *Journal of bacteriology*, 186(15), 4994–5002.
<https://doi.org/10.1128/JB.186.15.4994-5002.2004>
- Wasilenko, J. L., Fratamico, P. M., Sommers, C., DeMarco, D. R., Varkey, S., Rhoden, K., & Tice, G. (2014). Detection of Shiga toxin-producing *Escherichia coli* (STEC) O157:H7, O26, O45, O103, O111, O121, and O145, and *Salmonella* in retail raw ground beef using the DuPont™ BAX® system. *Frontiers in cellular and infection microbiology*, 4, 81.
<https://doi.org/10.3389/fcimb.2014.00081>
- Weintraub, L., Ahluwalia, M., Dogra, S., Uehlinger, J., Skversky, A., & Vasovic, L. (2014). Management of streptococcal pneumoniae-induced hemolytic uremic syndrome: a case report. *Clinical nephrology. Case studies*, 2, 9–17. <https://doi.org/10.5414/CNCS107887>
- Wendel, A. M., Johnson, D. H., Sharapov, U., Grant, J., Archer, J. R., Monson, T., Koschmann, C., & Davis, J. P. (2009). Multistate outbreak of *Escherichia coli* O157:H7 infection associated with consumption of packaged spinach, August-September 2006: the Wisconsin investigation. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, 48(8), 1079–1086.
<https://doi.org/10.1086/597399>
- Whitfield C. (2006). Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli*. *Annual review of biochemistry*, 75, 39–68.
<https://doi.org/10.1146/annurev.biochem.75.103004.142545>
- Wielgoss, S., Barrick, J. E., Tenaillon, O., Cruveiller, S., Chane-Woon-Ming, B., Médigue, C., Lenski, R. E., & Schneider, D. (2011). Mutation Rate Inferred From Synonymous Substitutions in a Long-Term Evolution Experiment With *Escherichia coli*. *G3 (Bethesda, Md.)*, 1(3), 183–186. <https://doi.org/10.1534/g3.111.000406>

- Willems, R. J., Top, J., Smith, D. J., Roper, D. I., North, S. E., & Woodford, N. (2003). Mutations in the DNA mismatch repair proteins MutS and MutL of oxazolidinone-resistant or -susceptible *Enterococcus faecium*. *Antimicrobial agents and chemotherapy*, 47(10), 3061-6.
- Wong, A., Rodrigue, N., & Kassen, R. (2012). Genomics of adaptation during experimental evolution of the opportunistic pathogen *Pseudomonas aeruginosa*. *PLoS genetics*, 8(9), e1002928. <https://doi.org/10.1371/journal.pgen.1002928>
- Woodward, D. L., Clark, C. G., Caldeira, R. A., Ahmed, R., & Rodgers, F. G. (2002). Verotoxigenic *Escherichia coli* (VTEC): a major public health threat in Canada. *The Canadian journal of infectious diseases = Journal canadien des maladies infectieuses*, 13(5), 321–330. <https://doi.org/10.1155/2002/383840>
- Xu, M., Zhou, Y. N., Goldstein, B. P. & Jin., D. J. (2005). Cross-resistance of *Escherichia coli* RNA polymerases conferring rifampin resistance to different antibiotics. *Antimicrob. Agents Chemother.* 187, 2783–2792.
- Yang, S., Xi, D., Jing, F., Kong, D., Wu, J., Feng, L., Cao, B., & Wang, L. (2018). Genetic diversity of K-antigen gene clusters of *Escherichia coli* and their molecular typing using a suspension array. *Canadian journal of microbiology*, 64(4), 231–241. <https://doi.org/10.1139/cjm-2017-0620>
- Yoshida, H., Bogaki, M., Nakamura, M., Yamanaka, L. M., & Nakamura, S. (1991). Quinolone resistance-determining region in the DNA gyrase *gyrB* gene of *Escherichia coli*. *Antimicrobial agents and chemotherapy*, 35(8), 1647–1650. <https://doi.org/10.1128/aac.35.8.1647>

- Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., Aarestrup, F. M., & Larsen, M. V. (2012). Identification of acquired antimicrobial resistance genes. *The Journal of antimicrobial chemotherapy*, 67(11), 2640–2644.
<https://doi.org/10.1093/jac/dks261>
- Zheng Q. (2002). Statistical and algorithmic methods for fluctuation analysis with SALVADOR as an implementation. *Mathematical biosciences*, 176(2), 237–252.
[https://doi.org/10.1016/s0025-5564\(02\)00087-1](https://doi.org/10.1016/s0025-5564(02)00087-1)
- Zheng Q. (2005). New algorithms for Luria-Delbrück fluctuation analysis. *Mathematical biosciences*, 196(2), 198–214. <https://doi.org/10.1016/j.mbs.2005.03.011>
- Zheng Q. (2017). rSalvador: An R Package for the Fluctuation Experiment. *G3 (Bethesda, Md.)*, 7(12), 3849–3856. <https://doi.org/10.1534/g3.117.300120>
- Zhu, L., Cai, Z., Zhang, Y., and Li, Y. (2014). Engineering stress tolerance of Escherichia coli by stress-induced mutagenesis (SIM)-based adaptive evolution. *Biotechnol. J.* 9, 120–127.
doi: 10.1002/biot.201300277