USING WHOLE GENOME SEQUENCING TO TRACK COLIBACILLOSIS IN SASKATCHEWAN BROILER FLOCKS

A Thesis Submitted to the
College of Graduate and Postdoctoral Studies
In Partial Fulfillment of the Requirements
For the Degree of Master of Science
In the Department of Biochemistry, Microbiology, and Immunology
University of Saskatchewan
Saskatoon

By

MADELINE CHRISTINE VALERIE McCARTHY

PERMISSION TO USE

In presenting this thesis/dissertation in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis/dissertation in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis/dissertation work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis/dissertation or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis/dissertation.

DISCLAIMER

This document was exclusively created to meet the thesis and/or exhibition requirements for the degree of Master of Science at the University of Saskatchewan. Reference in this thesis/dissertation to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not constitute or imply its endorsement, recommendation, or favoring by the University of Saskatchewan. The views and opinions of the author expressed herein do not state or reflect those of the University of Saskatchewan and shall not be used for advertising or product endorsement purposes.

Requests for permission to copy or to make other uses of materials in this thesis/dissertation in whole or part should be addressed to:

Head of the Department of Biochemistry, Microbiology, and Immunology
GA20, Health Sciences, 107 Wiggins Rd
University of Saskatchewan
Saskatoon, Saskatchewan S7N 5E5 Canada

OR

Dean
College of Graduate and Postdoctoral Studies
University of Saskatchewan
116 Thorvaldson Building, 110 Science Place
Saskatoon, Saskatchewan S7N 5C9 Canada

ABSTRACT

Colibacillosis is a systemic infection caused by *Escherichia coli* resulting in significant morbidity and mortality in broiler flocks worldwide. Little is known about the group of *E. coli* that cause colibacillosis, collectively termed avian pathogenic *E. coli* (APEC). My MSc research focused on determining how APEC differ from resident *E. coli* that live in the chicken gut but do not cause disease. I hypothesized that systemic and cecal *E. coli* are genetically distinct, and *E. coli* that cause colibacillosis are virulent outbreak strains. My objectives were to isolate *E. coli* from Saskatchewan broilers, sequence their genomes using Nanopore and Illumina technology, and screen them for virulence, antimicrobial resistance, and disinfectant resistance. I developed a pipeline to isolate and sequence *E. coli* from Saskatchewan colibacillosis outbreaks, selecting isolates based on outbreak, disease status, and biofilm profiles. I sequenced 96 *E. coli* isolates, consisting of 58 from diseased broilers with confirmed colibacillosis (systemic *E. coli*), and 38 from the cecal contents of healthy broilers in the same flocks (cecal *E. coli*). Our initial experiments were optimized for whole genome assembly and excluded DNA fragments under 500bp; therefore, we likely missed plasmids present in *E. coli* isolates. I tested six plasmid kits and two sequencing protocols to develop a methodology to capture missed plasmids in avian *E. coli* isolates and successfully identified new plasmids in both types of isolates. Systemic *E. coli* were more drug-resistant than cecal *E. coli* against a panel of 27 antimicrobial agents and possessed significantly more plasmids than cecal *E. coli*. plasmids contained multiple virulence and antimicrobial resistance genes that may contribute to disease. Since biofilms can provide protection from antibiotics and disinfectants, I quantified biofilm formation in three different medias. Systemic isolates were significantly more likely to form biofilms in rich media, but there was no correlation between biofilm formation and antimicrobial resistance. My characterization led us to conclude that systemic and cecal *E. coli* represent two different populations of strains. This will need to be confirmed with the analysis of more isolates. Characterization of avian pathogenic *E. coli* will help us understand how these isolates cause disease.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. Aaron White, for his mentorship and guidance during my time in his laboratory.

I would like to thank my committee members-Drs. Jo-Anne Dillon, Janet Hill, and Wei Xiao-for their time and feedback throughout this process.

I would like to thank my colleagues at VIDO and the faculty and staff of the Department of Biochemistry, Microbiology, and Immunology for supporting me during my graduate studies.

Thank you to Tennille, Tyra, and Nikki of the Poultry Extension Team for answering my many questions and helping me collect samples. This project would not exist without you.

Thank you to Lindsay Balezantis for being a wonderful friend and mentee.

To Kristen and Sean, my pandemic bubble, thank you both for being incredible friends and for helping me through such a rough year.

Finally, to all my family, thank you for your unending love and support.

DEDICATION

To my father, Paul.
For never giving up on me, even when I doubted myself.
It wasn't easy, but we made it.

To my grandfather, John.
For encouraging me and giving me so much advice.
Not a day goes by that I don't think of you.

To Scott.
For being my best friend, and for making me smile when nothing else did.
I couldn't have done it without you.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 2

CHAPTER 3

CHAPTER 4

APPENDIX B

LIST OF FIGURES

CHAPTER 2

CHAPTER 3

LIST OF ABBREVIATIONS

1.1 AFEC	Avian fecal *Escherichia coli*
1.2 APEC	Avian pathogenic *Escherichia coli*
1.3 CFU	Colony forming units
1.4 ColV	Colicin V
1.5 DNA	Deoxyribonucleic acid
1.6 ExPEC	Extraintestinal pathogenic *Escherichia coli*
1.7 GWAS	Genome-wide association study
1.8 HGT	Horizontal gene transfer
1.9 IPEC	Intestinal pathogenic *Escherichia coli*
1.10 LEE	Locus of enterocyte effacement
1.11 MALDI-TOF	Matrix-assisted laser desorption ionization time of flight
1.12 NMEC	Neonatal meningitis *Escherichia coli*
1.13 PCR	Polymerase chain reaction
1.14 SNP	Single nucleotide polymorphism
1.15 UPEC	Urinary pathogenic *Escherichia coli*
1.16 UTI	Urinary tract infection
2.1 APEC	Avian Pathogenic *Escherichia coli*
2.2 NGS	Next generation sequencing
2.3 PBS	Phosphate buffered saline
2.4 PDS	Prairie Diagnostic Services
2.5 PEX	Poultry Extension Team
2.6 TSB	Tryptic soy broth
2.7 WGS	Whole genome sequencing
3.1 AMR	Antimicrobial resistance
3.2 APEC	Avian pathogenic *Escherichia coli*
3.3 DNA	Deoxyribonucleic acid
3.4 LB	Luria bertoni
3.5 LSK	Ligation Sequencing Kit
3.6 ONT	Oxford Nanopore Technologies
3.7 RBK	Rapid Barcoding Kit

3.8 WGS	Whole genome sequencing
4.1 AMR	Antimicrobial resistance
4.2 APEC	Avian pathogenic *Escherichia coli*
4.3 BHI	Brain heart infusion
4.4 DDAC	Didecyldimethylammonium chloride
4.5 ESBL	Extended-spectrum β-lactamases
4.6 MBC	Minimum bactericidal concentration
4.7 MBEC	Minimum biofilm eradication concentration
4.8 MH	Mueller Hinton
4.9 MIC	Minimum inhibitory concentration
4.10 MLST	Multi-locus sequence typing
4.11 QAC	Quaternary ammonium compounds
4.12 TSA	Tryptic soy agar
5.1 AMR	Antimicrobial resistance
5.2 APEC	Avian Pathogenic *E. coli*
5.3 BHI	Brain heart infusion
5.4 DNA	Deoxynucleic acid
5.5 LB	Luria bertoni
5.6 LSK	Ligation sequencing kit
5.7 MALDI-TOF	Matrix-assisted laser desorption ionization time of flight
5.8 PEX	Poultry Extension Team
5.9 RBK	Rapid barcoding kit
5.10 TSB	Tryptic soy broth

CHAPTER 1: LITERATURE REVIEW & INTRODUCTION

1.1 Colibacillosis

1.1.1 Clinical Disease in Broiler Chickens

Colibacillosis is an umbrella term for numerous local and systemic diseases in poultry caused by the bacterium *Escherichia coli* (Guabiraba and Schouler, 2015; Newman *et al.,* 2021). *E. coli* infections can manifest locally as numerous reproductive tract and yolk sack infections or infections of the umbilical stump in chicks, or systemically as septicemia (Newman *et al.,* 2021). Outbreaks of disease are more likely to be acute infections, which are characterized by the presence of multiple lesions on the internal organs of birds that have succumbed to sepsis (Lutful Kabir, 2010; Newman *et al.,* 2021). These lesions are commonly found on the liver, spleen, heart, and air sacs, but may be found on other internal organs as well (Schouler *et al.,* 2012; Guabiraba and Schouler, 2015; Newman *et al.,* 2021). The route of infection is still unclear, but the primary hypothesis is that septic colibacillosis begins as a local respiratory infection. Broilers in a crowded poultry barn are likely breathing in dust that has been contaminated with fecal material. Fecal dust in poultry barns may carry up to 10^6 colony-forming units (CFU) of *E. coli* per gram of dust (Lutful Kabir, 2010). This fecal dust transports viable *E. coli* into the trachea. Upon colonizing the trachea, the bacteria can migrate into the respiratory tract, where pathogenic strains can penetrate through the epithelial layers and disseminate systemically via the bloodstream (Newman *et al.,* 2021). Examples of colibacillosis being passed vertically from hen to chick have also been observed (Lutful Kabir, 2010; Kim *et al.*, 2020; Delannoy *et al.*, 2021).

Colibacillosis can be caused by primary and secondary infections (Guabiraba and Schouler, 2015; Newman *et al.,* 2021). Birds already infected with respiratory viruses such as Newcastle disease virus and infectious bronchitis virus are particularly susceptible to colibacillosis (Vandekerchove *et al.,* 2004; Johnson *et al.,* 2008; Guabiraba and Schouler, 2015). In addition, deteriorating housing conditions in the broiler barn may negatively affect the birds. High levels of ammonia and dust in a broiler barn can impact the immune system and cause physical distress to the respiratory tract of broilers, increasing their susceptibility to colibacillosis (Swelum et al., 2021). However, examples of colibacillosis outbreaks in otherwise healthy birds exist, suggesting that some *E. coli* strains act as primary pathogens (Vandekerchove et al., 2004).

1.1.2 Economic Impacts of Disease

Colibacillosis is an economically expensive disease for producers (Vandekerchove *et al*., 2004; Lutful Kabir, 2010; Guabiraba and Schouler, 2015). It is one of the most impactful diseases affecting poultry, both in Canada and worldwide (Guabiraba and Schouler, 2015). Mortality associated with colibacillosis ranges from 5-20% of a flock (Newman *et al*., 2021). Additional costs to the producer are associated with decreased egg production and hatching rates due to illness and vertical transmission of disease (Lutful Kabir, 2010; Guabiraba and Schouler, 2015). In the case of broiler chickens, additional carcasses may be condemned at the slaughterhouse if colibacillosis lesions are found in other birds, and the remaining meat may be given a lower grade and price (Guabiraba and Schouler, 2015).

1.1.3 Treatment

Colibacillosis is a difficult disease to prevent, due to the natural presence of *E. coli* on the skin and feathers of broilers, as well as in the intestinal tract (Sargeant et al., 2019). Proper biosecurity practices on farms are essential to try and mitigate outbreaks caused by unsanitary conditions. Rodents and insects may be potential carriers of pathogenic *E. coli* strains (Nielsen *et al*., 2004). *E. coli* may also persist in the barn bedding or water systems of a large broiler farm. The addition of chlorine to drinking water systems in poultry barns helps mitigate colibacillosis outbreaks in broiler flocks (Dhillon and Jack, 1996).

Historically, colibacillosis has been controlled with prophylactic and therapeutic antibiotic treatment (Kazemnia et al., 2014). Antimicrobial use in food production animals has been linked to the spread of antimicrobial resistance to humans via multiple routes, including the bacterial contamination of food products (Aidara-Kane et al., 2018). Concerns of heavily drug-resistant bacterial infections in humans has resulted in recommendations from the World Health Organization to heavily restrict the use of antimicrobials in food animals (Aidara-Kane et al., 2018). In many first-world countries, antimicrobial drugs with high importance to human medicine are no longer used in livestock, and most other drugs require a veterinary prescription for use (Landers *et al*., 2012). Globally, the majority of *E. coli* isolates recovered from broilers with colibacillosis are resistant to at least one antimicrobial agent, with many being multi-drug resistant (Kaemnia et al., 2014; Joshi et al., 2012; Halfaoui et al., 2017).

Prophylactic treatment of colibacillosis with probiotics has been attempted with limited success. In theory, microbes present in probiotic substances can directly inhibit intestinal pathogens by outcompeting them, blocking access to the epithelial layer, or stimulating the host immune response to enhance protection (Lee et al., 2010). Lee et al. determined that three strains of *Bacillus subtilis*-based probiotics were able to inhibit the growth of avian pathogenic *E. coli in vitro* (2010). While probiotic therapy shows promise as a potential alternative for antibiotic growth promotion, there is limited support for the use of probiotic therapy to prevent or treat colibacillosis (Amerah et al., 2013; Redweik et al., 2020).

Bacteriophage therapy is another potential treatment option for colibacillosis (Sorour et al., 2020). Broiler chickens treated with an oral dose of a *Salmonella enterica* serovar Enteritidis phage were able to fight bacterial infection faster than placebo, while broiler chicks treated two hours post-infection with an *E. coli* lytic bacteriophage had milder symptoms and recovered faster overall (Fiorentin et al., 2005; Lau et al., 2010). Bacteriophage therapy is an appealing candidate in light of globally increasing antimicrobial resistance. Phage therapy continues to show promise in treating colibacillosis outbreaks, however due to the wide diversity of *E. coli* strains that may cause disease, results may be variable (Sorour et al., 2020).

Autogenous vaccines have been proposed as another prophylactic candidate for colibacillosis treatment. Numerous efforts to develop an autogenous vaccine for colibacillosis have been attempted over the past decades. Poulvac is a currently licensed vaccine given to broilers and chicks to prevent septicemia caused by *E. coli* (Li et al., 2017). This vaccine contains a Δ*aroA*-mutant strain of *E. coli* which affects amino acid synthesis (Felgner et al., 2016). Poulvac has largely been successful at preventing or reducing the severity of colibacillosis in broiler flocks, but it does not provide long-term protection and does not cover all strains of *E. coli* that can cause disease (Li et al., 2017). Once again, the diversity of *E. coli* strains responsible for colibacillosis, and the presence of non-pathogenic *E. coli* strains present in and on broilers, makes development of a universal vaccine candidate challenging.

1.2 Avian Pathogenic *E. coli*

1.2.1 Diversity and Pathogenesis

Escherichia coli is a Gram-negative bacillus that is a common commensal member of the intestines of warm-blooded animals, including poultry (Newman et al., 2021). Some *E. coli* can escape the intestinal niche of poultry to cause extraintestinal disease. These strains, deemed Avian pathogenic *E. coli (*APEC), are difficult to characterize due to the great diversity of strains recovered from diseased broilers (Schouler *et al.,* 2012). Shortly after birth, chicks are colonized with a wide variety of *E. coli* strains likely acquired from the local barn environment (Kemmett *et al.*, 2013). Unlike other extraintestinal pathogenic *E. coli* (ExPEC) lineages which cause pandemic disease in humans, there are no clear lineages of APEC that are dominant locally or globally (Riley, 2014; Cordoni et al., 2016). The gold standard method of APEC identification is the chick lethality assay (Skyberg *et al.,* 2007). However, because this method is slow, many methods have been developed to attempt to predict strain pathogenesis. Serogrouping, a technique using antiserum specific to the O-antigens present on the surface of Gram-negative bacteria, has been a routine tool for the identification of APEC for decades (DebRoy et al., 2018; Cordoni et al., 2016). In most countries, six serogroups (O1, O2, O5, O8, O18, and O78) are responsible for over 50% of colibacillosis infections in broilers (Dziva and Stevens, 2008). However, classification of APEC strains by serogroup is a poor strategy. One study which attempted to develop a diagnostic assay to identify APEC by serotype found that this method failed to identify over 40% of pathogenic strains (10). *E. coli* with these serogroups are also found in the intestines of healthy broilers, suggesting that the intestinal tract serves as a potential reservoir for APEC strains (Dziva and Stevens, 2008). Multilocus sequence typing of APEC strains confirms a diverse global population (Maluta *et al.,* 2019). Several sequence types are associated with potentially zoonotic *E. coli*, including those from ST95, ST117, and ST131 (Maluta *et al.,* 2019). These are all global ExPEC lineages associated with extraintestinal disease in both poultry and humans (Totsika *et al.,* 2011; Ronco *et al.,* 2017; Forde *et al.*, 2018). *E. coli* in these lineages frequently possess resistance plasmids and are associated with the spread of antimicrobial resistance genes in hospitals and farm animals (Mamani *et al.*, 2019).

E. coli can be classified into eight groups called phylogroups based on the presence or absence of specific DNA fragments (Clermont *et al.,* 200). These groups are derived from a

common ancestry and provide some insight into the lifestyle and habitat of *E. coli* strains (Clermont *et al.,* 2000; 2018). Isolates recovered from broilers with confirmed colibacillosis are distributed across all eight *E. coli* phylogenetic lineages, though they are predominantly identified in phylogroups B2 and D, while commensal avian isolates are typically categorized as phylogroup A and B1 (Bin Kim *et al.,* 2019). Strains were originally classified by phylogroup using a triplex PCR screening method to identify the presence or absence of three genes-*chuA* (an outer membrane protein involved in the uptake of heme molecules), *yjaA* (a protein involved in response to low cellular pH), and a DNA fragment identified as a lipase esterase gene-that loosely characterize *E. coli* strains (Clermont *et al.,* 2000; Clermont *et al.,* 2013). Phylogroup A contains *E. coli* strains from a wide variety of habitats and host organisms (Duriez *et al.,* 2001). Group A is made up of attenuated, lab strains such as *E. coli* K12, as well as true commensal strains that do not cause disease in the host (Abram *et al.,* 2020). Likewise, phylogroup B1 predominantly contains non-pathogenic *E. coli* frequently isolated from environmental sources (Duriez *et al.,* 2001; Gordon and Cowling, 2003). *E. coli* in phylogroups A and B1 typically possess few, if any, virulence factors (Abram *et al.,* 2020). Phylogroup B2, and to a lesser extent phylogroup D, typically contains pathogenic extraintestinal *E. coli* (Clermont *et al.,* 2000). Group B2 strains are frequently isolated from farm animals, and in humans they are associated with urinary tract infections (Hutton *et al.,* 2017). The original screening method has been modified and expanded to include phylogroups C, E, F, and G, based on the high rate at which *E. coli* gain and lose genes (Clermont *et al.,* 2013). Phylogroup C is related to group B1 but harbors *E. coli* strains with a greater number of virulence genes (Moissenet *et al.,* 2010). Group C strains are more likely to cause extraintestinal disease than B1 strains (Clermont *et al.,* 2011). Phylogroup E strains typically have the largest genomes of all phylogroups, due largely to the presence of extracellular DNA and mobile genetic elements, such as plasmids and pathogenicity islands (Clermont *et al.,* 2021). This phylogroup contains lineages from human, animal, and environmental sources (Clermont *et al.,* 2021). Phylogroup F contains pathogenic lineages often associated with high virulence and zoonotic potential (Zhuge *et al.,* 2020). Finally, phylogroup G has recently been recognized as a pathogenic lineage related to groups B2 and F, which contains highly virulent extraintestinal strains that are often drug-resistant (Clermont *et al.,* 2019). The phylogroups have evolved largely due to horizontal gene transfer and niche specialization (Mellata *et al.,* 2010).

There are no clear defining features of APEC (Johnson *et al*., 2008; Schouler *et al*., 2012; Cordoni *et al*., 2016; Newman *et al*., 2021). There is significant overlap in sero- and phylogroup among pathogenic and commensal *E. coli* isolated from broilers (Cordoni *et al*., 2016; Cunha *et al*., 2017; Halfaoui *et al*., 2017; DebRoy *et al*., 2018). This overlap makes it difficult to clearly categorize and define an *E. coli* strain as APEC without performing chicken lethality experiments (Kim et al., 2020). Historically, the presence of certain virulence genes has been used to identify APEC strains (Mohamed et al., 2017). Genes for numerous adhesins, iron acquisition systems, auto-transport systems, and anti-host defense are routinely found in pathogenic isolates (Anato *et al*., 2009; Roquet *et al*., 2009; Li *et al*., 2010). However, no single virulence gene or group of genes has been universally identified in APEC isolates, and many of these virulence genes are also present in commensal avian *E. coli* genomes (Touchon *et al*., 2009; Schouler *et al*., 2012; Newman *et al*., 2021).

The *E. coli* genome is incredibly diverse. Typical genome size for an *E. coli* isolate ranges from 4.6 to 5.8Mb, representing between 4600 to 5800 protein encoding genes (Dziva and Stevens, 2008; Touchon et al., 2009). Roughly 2,200 of those genes are conserved among all *E. coli* isolates in the core genome, which accounts for less than half of the total gene content of many members of the species (Touchon et al., 2009). *E. coli* are known to acquire extracellular DNA through horizontal gene transfer (HGT), and this acquisition of extracellular DNA plays a role in the development of pathogenicity (Johnson *et al*., 2008a; Touchon et al., 2009). Regions of the *E. coli* genome known as 'hotspots' have high recombination rates and play a role in the diversity seen in many pathogenic isolates (Touchon *et al*., 2009; Kemmett *et al*., 2013). The origin of the *stx* genes responsible for severe intestinal disease and hemolytic uremic syndrome caused by enterotoxigenic *E. coli* are phage-mediated genes which have subsequently been incorporated into the *E. coli* chromosome (Gyles, 2007). Likewise, a virulence plasmid found in many *E. coli* strains that cause urinary tract infections contains a pathogenicity island that plays a role in the early stages of cellular invasion during infection in humans (Cusumano et al., 2020).

1.2.2 APEC Plasmids

Many virulence genes are found on virulence plasmids, such as the colicin V (ColV) plasmid associated with APEC strains (Johnson *et al*., 2008b). Johnson *et al*. identified ColV plasmids in over 80% of pathogenic strains but only 25% of commensal avian *E. coli* (Johnson *et al*., 2006).

These large APEC-associated virulence plasmids often co-exist in the bacterial cell with additional R plasmids that contain numerous antimicrobial resistance genes (Johnson *et al*., 2012; Schouler *et al*., 2012). Possession of multiple large virulence plasmids has become a defining characteristic of APEC (Mellata *et al*., 2010). APEC strains cured of virulence plasmids are severely attenuated in chicken models and have reduced serum resistance compared to wild-type strains (Mellata *et al*., 2010; Mellata, 2013). The large plasmids identified in APEC strains likely contribute to the diversity of APEC as a whole: a wide variety of virulence factors present on the plasmids allows *E. coli* with different genetic backgrounds to adapt to unpredictable environmental conditions. Some genes, such as those for iron uptake and serum survival, are essential for colonizing the chicken host (Johnson *et al*., 2008; Mellata *et al*., 2010; Schouler *et al*., 2012).

The presence of a large plasmid is not enough to characterize a strain as APEC (Cunha *et al*., 2012; Johnson *et al*., 2012; Mageiros *et al*., 2021). Numerous studies determined that large plasmids are ubiquitous in avian *E. coli* isolates (Mellata *et al*., 2010; Johnson *et al*., 2012; *Kim* et al., 2020; Mageiros *et al*., 2021). Highly virulent APEC strains typically possess five conserved virulence genes-*iroN, ompT, hlyF, iss*, and *iutA*-on a single plasmid, while less virulent strains may possess a mixture of plasmid and chromosomal virulence genes (Schouler *et al*., 2012; Ikuta *et al*., 2014; Schouler *et al*., 2014). These virulence genes are often spread across multiple large plasmids in APEC isolates, and several of these plasmids have been implicated in the zoonotic potential of certain avian *E. coli* (Mellata *et al*., 2010).

The most well-known APEC plasmid is pAPEC-1, a 103 kb plasmid which was isolated from an O78:K80:H9 *E. coli* isolate recovered from the liver of a turkey (Mellata *et al*., 2009; Mellata *et al*., 2010). This plasmid contains numerous siderophores which enhance an *E. coli* isolate's ability to grow in iron-restricted media (Mellata *et al*., 2010). pAPEC-1 also plays a role in serum survival and complement resistance, which are commonly identified traits in APEC strains (Mellata *et al*., 2010). Among several large plasmids associated with APEC, pAPEC-1 had the greatest impact on lethality in day old chicks, though when combined the plasmids had a synergistic effect on lethality (Mellata *et al*., 2010). APEC plasmids affect strain growth in different media (Mellata *et al*., 2010). In minimal media, the presence of plasmids pAPEC-1 and pAPEC-2 were detrimental to fitness compared to growth in LB broth (Mellata *et al*., 2010). Plasmid pAPEC-2 is another large (90 kb) plasmid frequently found in APEC isolates (Skyberg *et al*., 2006; Mellata *et al*., 2010).

Unlike pAPEC-1, it does not contain the typical virulence factors seen in pathogenic isolates (Mellata *et al*., 2012). Rather, this self-replicating plasmid carries several antimicrobial resistance genes, and several genes present in human *E. coli* isolates (Mellata *et al*., 2012). The homology to plasmids found in other *E. coli* strains suggests that this may be a widely circulating plasmid not specific to avian infections (Mellata *et al*., 2012).

The presence of multiple APEC plasmids in a bacterial cell can have a synergistic effect on the physiology of the bacterium. Mellata *et al*. demonstrated that the combined effect of three APEC plasmids promoted bacterial biofilm formation at host body temperatures (2012). Bacterial biofilms are commonly induced at lower temperatures (Brombacher *et al*., 2003).

The presence of APEC plasmids may play a significant role in the transition from commensal to pathogenic isolate (Mellata *et al*., 2010). Curing an APEC strain of all large plasmids completely attenuated that strain in a chicken embryo lethality assay model of colibacillosis (Mellata *et al*., 2010). Conversely, transferring the APEC plasmid pAPEC-2 to a commensal *E. coli* isolate significantly enhanced the lethality of the isolate and allowed it to grow and colonize the mouse urinary tract (Skyberg *et al*., 2006). A test for the presence of multiple plasmid-encoded virulence factors has been used to putatively identify APEC isolates, but the lethality assay remains the gold standard (Tivendale *et al*., 2009). Numerous genes are associated with APEC, but not exclusive to APEC strains. However, APEC plasmids such as pAPEC-1 and pAPEC-2 have multiple highly conserved regions which encode genes associated with APEC strains and are consistent in isolates from around the globe (Tivendale *et al*., 2009).

1.2.3 Antimicrobial Resistance

Antimicrobial resistance is a growing global threat, and the historic widespread use of antibiotic agents for prophylaxis and enhanced growth in the food animal industry contributed significantly to this problem (Davies and Wales, 2019). *E. coli* play an important role in the spread of antimicrobial resistant plasmids due to their genetic plasticity and promiscuity. Multi-drug resistance is common among *E. coli* isolated from avian sources, irrespective of host disease status (Johnson *et al*., 2006). However, correlations between antimicrobial resistance and specific APEC virulence factors have been observed in multiple studies (Johnson *et al*., 2006; Johnson *et al*., 2012; Davis *et al*., 2018). APEC genes responsible for enhanced adherence (*afa, iha*) and iron uptake (*iutA)* during infection are significantly correlated with ampicillin resistance, while both *afa* and

iutA are also correlated with trimethoprim-sulfamexazole resistance (Servin, 2005;Johnson *et al*., 2012; Landragf *et al*., 2012; Leveille *et al*., 2020). Previous studies identified an association between APEC virulence factors found on ColV plasmids (*cvaC, iss, iutA, traT*) and resistance to gentamycin, tetracycline, and certain cephalosporin antibiotics (Johnson *et al*., 2012). Drug-resistant avian *E. coli* may be introduced into poultry food products at various points of production (Davis *et al*., 2018). *E. coli* are routinely isolated from poultry meat in commercial stores (Jonson *et al*., 2012; Davis *et al*., 2018). Some studies cultured *E. coli* from over 90% of chicken meat samples; of those, roughly 25% of the isolates were APEC (Lyhs *et al*., 2012).

1.3 *E. coli* Biofilms

1.3.1 Biofilms as an Alternative Reservoir to the Chicken Gut

The majority of research into APEC has focused on characterizing pathogenic *E. coli* strains (Johnson *et al*., 2008; Ewers *et al*., 2009). Many questions remain about the reservoir for these bacteria. The current and dominant hypothesis is that broilers acquire colibacillosis infections from contaminated fecal dust that is seeded with APEC isolates transiting through the chicken gut (Lutful Kabir, 2010). Comparisons of *E. coli* isolated from the feces of healthy chickens (avian fecal *E. coli* or AFEC) suggest that up to 15% of *E. coli* isolates in the feces may be APEC (Ewers *et al*., 2009). Virulence genes associated with traits such as serum and complement resistance are common in both APEC and AFEC isolates (Ewers *et al*., 2009). Broiler chickens often come from a limited number of hatcheries, where they are bred to grow quickly and efficiently (Kemmett *et al*., 2013). Thus, vertical transmission of *E. coli* from hen to chick may lead to colonization with *E. coli* strains with pathogenic potential early in life (Kemmet *et al*., 2013). Detection of APEC strains in young broilers typically declines with age, and after 3 weeks there is a significant decline in APEC identified in chicks from hatcheries (Kemmet *et al*., 2013). This suggests that APEC may be transient in the chicken gut.

1.3.2 Biofilm Formation and Environmental Persistence

Bacteria can persist on surfaces in aggregated, static communities instead of as singular planktonic cells (Vogeleer *et al*., 2014). These communities, or biofilms, are complex matrices composed of proteins, extracellular DNA, and many polysaccharides (Vogeleer *et al*., 2014).

Bacteria within a biofilm are protected from external pressures such as desiccation, lack of nutrients, antimicrobials, and disinfectants (Brombacher *et al*., 2003; Abdallah *et al*., 2014; MacKenzie *et al*., 2017). Curli and cellulose are major components of the extracellular matrix of *E. coli* and *Salmonella* biofilms (Barnhart and Chapman, 2006). Curli fimbriae are amyloid proteins that help promote the formation of aggregates during early biofilm formation (Saldana *et al*., 2009). In APEC isolates curli play a role in pathogenesis: expression of curli enhanced colonization of the liver, spleen, and cecum in day-old chicks (Ragione *et al*., 1999). High levels of curli expression also correlate to an increased level of internalization of *E. coli* and *Salmonella* strains in eukaryotic cells (Gophna *et al*., 2001). Cellulose is a polymer of glucose and one of the main building blocks of the biofilm matrix (Serra *et al*., 2013). Production of cellulose in a bacterial biofilm significantly increases resistance to desiccation (Gualdi *et* al., 2008). Curli and cellulose play a synergistic role in the structural formation of environmental biofilms in Enterobacteriaceae (Saldana *et al*., 2009). However, while the overexpression of curli has been shown to promote adhesion to host cells during infection, this effect can be mediated by increased cellulose production (Gualdi *et al*., 2008).

Bacteria that become adapted to their host form may develop mutations in biofilm genes as they become disadvantageous for survival in the host environment (MacKenzie *et al*., 2017). It is advantageous for bacteria passing transiently through a host to form biofilms to survive in unpredictable conditions (MacKenzie *et al*., 2017). A broiler barn provides numerous surfaces upon which bacteria can form biofilms. *E. coli* biofilms have been found in and on the drinking water systems in barns (Heinemann *et al*., 2020). They can also form biofilms in the dirt and bedding of a broiler barn, where the organic matter present may provide additional protection from disinfectants during barn cleaning (Abdallah *et al*., 2014). Dust from commercial barns is commonly identified as an environmental reservoir of *E. coli* (Schulz *et al*., 2016). Viable *E. coli* have been isolated from barn dust samples up to 20 years after collection (Schulz *et al*., 2016). The ability of APEC isolates to form environmental biofilms is highly variable, though most isolates form at least a weak biofilm *in vitro* (Skyberg *et al*., 2007; Nielsen *et al*., 2018; Rodrigues *et al*., 2018). Due to the variation in determining the APEC pathotype, and the uncertainty surrounding the circumstances of colibacillosis infection, it is difficult to evaluate what role biofilm formation may play in persistence and infection.

1.3.3 Biofilm Formation and Disinfectant Resistance

Broiler barns are thoroughly disinfected in between flocks of birds. Common commercial disinfectants include active ingredients such as quaternary ammonium compounds, reactive oxygen species, and potassium peroxymonosulfate (Payne *et al.,* 2005). The concentration and surface contact time of the disinfectant largely determines its ability to kill microorganisms in broiler barns (Course *et al.,* 2021). Disinfection is essential at multiple stages of the food production line, particularly in the barn and in food production plants (Sun *et al.,* 2019). *E. coli* isolated from retail meats frequently possess disinfectant resistance genes that are often associated with antimicrobial resistance genes (Zhang *et al.,* 2016). Biofilm formation promotes the attachment of *E. coli* cells to surfaces such as drinking water and feed pens in barns, and stainless steel in packaging plants (Silagyi *et al.,* 2009). Biofilms can shield bacteria close to the centre from lethal concentrations of disinfectants, allowing them to persist and potentially infect new broilers placed in barns (Marin *et al.,* 2009). *E. coli* can form biofilms on a wide variety of abiotic and biotic surfaces, including dust and litter particles on barn floors (Sliagyi *et al.,* 2009). Organic matter is known to interfere with the efficacy of many disinfectants, potentially resulting in exposure to sub-lethal concentrations of disinfectants (Ruano *et al.,* 2001).

1.3.4 Virulence Factors that Enhance Biofilm Formation

Biofilm formation plays an important role in both environmental persistence and infection. (Vogeleer *et al.,* 2014). During infection, biofilms provide increased protection from bactericidal factors like the host immune response and antibiotics (Flament-Simon *et al.*, 2019). Many APEC isolates are capable of forming biofilms, and many possess virulence genes that enhance biofilm formation (Germon *et al.*, 2005; Wang *et al.*, 2016). Adherence to host epithelial cells is often a preliminary step during infection (Wang *et al.*, 2016). APEC strains normally carry multiple genes for adhesins such as *papC* and several fimbriae genes (Wang *et al.*, 2016). *E. coli* isolates that possess *papC* and *papG* adhesin alleles are significantly more likely to form strong biofilms (Naves *et al.*, 2007). Previous studies have shown that similar to urinary pathogenic *E. coli*, type I pili in APEC are involved in the early stages of colibacillosis infection, while P-fimbriae play a role later in infection (Germon *et al.*, 2005). The group of *Yad* fimbriae genes play numerous roles in biofilm formation, both at and below physiological host body temperature (Larsonneur *et al.*, 2016; Verma *et al.*, 2016). Enhanced ability to form biofilms at temperatures below host body temperature suggests that these APEC strains may persist in the environment (Verma *et al.*, 2016).

Other virulence factors that play a role in biofilm formation include invasins, toxins, quorum sensing molecules, and secretion systems (Kathayat *et al*., 2021). IbeA is a protein which plays a role in the invasion of epithelial cells, particularly in the brain (Germon *et* al., 2005). A mutant strain lacking the *ibeA* gene produced significantly less biofilm, though the mechanism behind this is not clear (Wang *et al*., 2011). *yicS*, a virulence factor encoding a protein involved in invasion and motility, enhances wild-type biofilm formation in APEC isolates (Verma *et al*., 2018). Biofilm formation in a Δ*yicS* mutant *E. coli* was severely diminished, suggesting that YicS may play a role in transporting proteins involved in biofilm formation (Verma *et al*., 2018). A survey analyzing early biofilm formation in multiple *E. coli* lineages determined that the toxin hlyA was significantly associated with strong biofilm formation (Flament-Simon *et al*., 2019).

1.4 Zoonotic Potential of APEC Isolates

1.4.1 Similarity to Other ExPEC

E. coli can be broadly classified in two groups: intestinal pathogenic *E. coli* (IPEC), which contain many foodborne diarrheagenic isolates, and extraintestinal pathogenic *E. coli* (ExPEC) (Mellata, 2013). *E. coli* classified as ExPEC cause disease outside of the gastrointestinal tract (Mellata, 2013). In humans, ExPEC are responsible for many cases of sepsis, urinary tract infections, and neonatal meningitis (Kathayat *et al*., 2021; Mellata, 2013). The recovery of APEC isolates from food production facilities and retail meat has raised concerns about potential spread from food animals to humans (Liu *et al*., 2018; Mitchell *et al*., 2015). APEC isolates sharing sequence types with global lineages of urinary pathogenic *E. coli* (UPEC) and *E. coli* that cause neonatal meningitis (NMEC) have been identified in multiple countries, raising questions about their ability to cause extraintestinal disease in humans (Liu *et al*., 2018; Tivendale *et al*., 2010).

Human ExPEC strains typically share a few defining virulence genes, but are overall genetically diverse (Mellata, 2013). Many ExPEC isolates are ultimately diagnosed based on their isolation source (Lloyd *et al*., 2007). NMEC are identified by the presence of six genes: capsular antigen (*kpsII, K1),* sialic acid synthesis (*neuC*), iron transport (*iucC, sitA*), and vacuolating autotransporter toxin (*vat*) (Wijetunge *et al*., 2015; Vann *et al*., 2004; Lorenzo and Neilands, 1986; Sabri *et al*., 2006; Diaz *et al*., 2020). UPEC isolates are capable of growth in human urine, and

typically possess genes for fimbrial adhesion (*pap, dra, afa, foc*), hemolysin (*hlyA*), serin protease autotransporter (*pic*), and iron transport/acquisition (*ibeABC, iuc, iroN, chu, sitABCD*) (Sarowska *et al*., 2019).

1.4.2 Shared Genetic and Virulence Traits

APEC share many genetic characteristics with other ExPEC isolates. A study comparing single nucleotide polymorphisms (SNPs) in the genomes of 323 APEC and human ExPEC ST95 isolates revealed that APEC are a not a distinct lineage, and significant genetic overlap was observed between APEC, UPEC, and NMEC strains (Jorgensen *et al*., 2019). Meat products have been flagged as a potential source of UPEC, particularly from the global ST131 *E. coli* lineage (Liu *et al*., 2018). This study analyzed the genomes of ST131-H22 *E. coli* recovered from commercial poultry products and compared them to human UPEC strains. Once again, phylogenetic analysis and *in silico* multilocus sequence typing showed that numerous isolates recovered from meat were closely related to human pathogenic isolates (Liu *et al*., 2018). This study also identified the presence of APEC plasmids in human UPEC isolates (Liu *et al*., 2018). Of the 25 human isolates that were closely related to poultry isolates, 23 possessed APEC ColV plasmids (pAPEC-1 and pAPEC-2). These plasmids are a distinguishing feature of avian pathogenic isolates, and the genes encoded on such plasmids are often used to identify APEC strains (Johnson *et al*., 2006). The presence of APEC plasmids in human *E. coli* strains suggests that they may have a zoonotic origin.

Other plasmids identified in human ExPEC isolates share virulence clusters that are homologous with APEC plasmids (Peigne *et al*., 2009). These clusters typically contain genes encoded on pathogenicity islands for iron uptake and outer membrane transporters (Peigne *et al*., 2009). Related extended-spectrum beta-lactamase plasmids belonging to the IncI replication type have been identified in multiple avian and human *E. coli* (Accogli *et al*., 2013). The plasmid-derived gene for increased serum survival (*iss*) plays a significant role in extraintestinal pathogenesis (Johnson *et al*., 2008b). An insertion sequence element containing this gene was prevalent in both APEC and NMEC isolates in a study of almost 500 ExPEC strains (Johnson *et al*., 2008b). APEC and NMEC often share large, ExPEC associated plasmid regions. Notably, possession of pathogenicity-island linked genes (*hlyF, etsAB*, and *sitA)* is common among these isolates (Johnson *et al*., 2008). Among APEC, UPEC, and NMEC isolates, two genes (*iutA* and

ompT) are common (Johnson *et al*., 2008b). These genes, for iron transport and outer membrane transport, respectively, are found in many ExPEC isolates (Kathayat *et al*., 2021*)*.

Due to the natural promiscuity of *E. coli*, and the prevalence of horizontal gene transfer in nature, many APEC isolates possess genes characteristic of other ExPEC (Mitchell *et al*., 2015; Meena *et al*., 2021). The *pap* operon genes, which are a defining feature of UPEC, were identified in 67% of APEC isolates characterized in a study by Najafi *et al*. (2019). Because many defining ExPEC virulence genes are located within pathogenicity islands on plasmids, other *E. coli* may gain new genes which enhance their ability to infect other mammals (Sarowska *et al*., 2019). Recently, hybrid IPEC/ExPEC strains have been identified in human urinary tract infection (UTI) patients. These strains, isolated from patients with UTIs, possess genes associated with the locus of enterocyte effacement (LEE) (Valiatti *et al*., 2020). The LEE is a pathogenicity island characteristic of enteropathogenic *E. coli* (McDaniel *et al*., 1995). The hybrid isolates were able to secrete effector proteins encoded by the LEE, but they produced weak attaching/effacing lesions on eukaryotic cells (Valiatti *et al*., 2020). These strains are likely ExPEC in origin but have acquired the LEE pathogenicity island via HGT. IPEC genes have been identified in APEC isolates, raising additional concerns if these isolates were to infect humans (Lindstedt *et al*., 2018).

1.5 Challenges in Defining the APEC Pathotype

1.5.1 Early Classification of APEC Isolates Using Serotyping and PCR

Producers and researchers have been struggling to identify and manage APEC in poultry for decades. One of the biggest challenges in the management of colibacillosis is defining the APEC pathotype. APEC are currently defined based on the presence of multiple virulence genes and isolation from a broiler with confirmed colibacillosis (Kemmett *et al*., 2013). Fecal isolates from healthy chickens often possess virulence genes found in APEC, creating confusion about specific requirements for infectivity (Collingwood *et al*., 2014). There are no clear traits which define the APEC pathotype. To produce better and more targeted treatments, disease manifestation and geographic location need to be taken into consideration.

Before genomic technologies became the norm, attempts to define the APEC pathotype were limited to serotyping and virulence gene screening (Koga *et al*., 2014; Joensen *et al*., 2015). The

great diversity of *E. coli* isolates capable of infecting poultry belong to many serotypes, and none stand out as a dominant type (Joensen *et al*., 2015). In one study, approximately 56% of APEC isolates and 22% of avian fecal (AFEC) isolates fell into one of six serogroups: O1, O2, O5, O8, O18, and O78) (Schouler *et al*., 2012). Due to the overlap of some fecal isolates, and the fact that as many as 30% of APEC isolates are not typeable, serotyping is an unreliable method for the identification of APEC strains (Schouler *et al*. 2012; Maluta *et al*., 2014; Awad *et al*., 2020).

Polymerase chain reaction (PCR) screening is highly accurate but limited solely to known genes (Hu *et al*., 1999). Previous research into APEC and colibacillosis relied exclusively on PCR to determine the presence and absence of virulence genes to try and define a pathotype (Ewers *et al*., 2004; Rodriguez-Siek *et al*., 2005; Jonson *et al*., 2012; Subedi *et al*., 2018). PCR panels range from a small group of frequently identified genes to massive panels surveying a wide range of ExPEC and IPEC genes. Targeted PCR panels have been proposed as a method of identifying APEC isolates (Johnson *et al*., 2008b; Schouler *et al*., 2012; Ikuta *et al*., 2014). These PCR panels screen for the presence of several virulence genes frequently associated with APEC strains and typically found on APEC virulence plasmids (Schouler *et al*., 2012). A multiplex PCR from 2003 screened for *iucC*, a gene encoding an iron uptake protein, *iss*, a gene for increased serum survival, *tsh,* the gene encoding temperature sensitive hemagglutinin, and *cvi*, a colicin gene (Skyberg *et al*., 2003). In this small study, 8/10 strains isolated from diseased chickens possessed 75% of the genes (Skyberg *et al*., 2003). Other common APEC genes targeted by multiplex PCR include the aerobactin iron receptor gene *iutA*, hemolysin gene *hlyF*, the iron uptake receptor *iroN*, and the outer membrane transport protein gene *ompT* (Johnson *et al*., 2008b; Ikuta *et al*., 2014). Using a multiplex PCR with the *iutA, iroN, hlyF, ompT,* and *iss* genes, Johnson *et al*. were able to correctly predict 85% of APEC isolates in a study of almost 1,000 *E. coli* strains based on possession of at least 4/5 genes (2008b). However, up to 40% of the 200 fecal avian *E. coli* tested in this study possessed one or more of the same virulence genes. To date, no multiplex PCR has been designed that can identify 100% of APEC isolates.

Neither serotyping nor screening for virulence genes via PCR are reliable methods for identifying APEC isolates (Johnson *et al*., 2008b; Schouler *et al*., 2012). To determine which isolates are truly pathogenic, both methods need to be followed by the gold standard lethality assays using embryos and day-old chicks (Johnson *et al*., 2008b; Awad *et al*., 2020). Recent comparisons

of serotyping and lethality assays demonstrated that serotyping does not predict virulence of an APEC isolate, but screening for virulence genes can identify highly virulent isolates (Awad *et al.,* 2020).

1.5.2 APEC Source and Disease Manifestation

Further diagnostic challenges lay in defining the source of APEC isolates in each study. APEC is an umbrella term given to *E. coli* isolated from poultry with confirmed *E. coli* infections (Mehat *et al*., 2021). Until recently, there was some debate about whether APEC were primary or secondary pathogens in poultry (Mellata, 2013; Collingwood *et al*., 2014). While some strains appear to be opportunistic pathogens and struggle to cause disease in otherwise healthy broilers, other APEC strains are indeed highly virulent primary pathogens (Collingwood *et al*., 2014). Numerous factors including the site of infection, the immune response of the bird, and the virulence repertoire of the *E. coli* isolate determines the manifestation of disease (Mehat *et al*., 2021). Extraintestinal infections can manifest as respiratory infections, heart, liver, or air sac infections, as well as colibacillosis (septicemia) (Sadeyen *et al.,* 2014; Mehat *et al*., 2021). In layer hens, *E. coli* infections of the reproductive tract are common and may transmit APEC to eggs and newborn chicks (Van Goor *et al*., 2020). Many colibacillosis studies do not differentiate between clinical disease manifestations, rather they group together APEC isolates that cause various extraintestinal disease (Mehat *et al*., 2021). Including multiple manifestations of disease under the umbrella term 'colibacillosis' is likely hindering the definition of an APEC pathotype that causes severe systemic disease in poultry.

In order to increase the number of strains for higher statistical power, many studies combine and analyze large pools of APEC and AFEC strains from a wide geographic area (Skyberg *et al.,* 2007; Nielsen *et al*., 2018; Mageiros *et al*., 2021). While these studies agree that APEC isolates typically possess a greater number of virulence genes and virulence plasmids than AFEC isolates, methods to successfully diagnosis of APEC isolates still fall short (Skyberg *et al.,* 2007; Nielsen *et al*., 2018; Mageiros *et al*., 2021). Virulence profiles differ between broiler flocks, with almost 70% of farms possessing strains with unique virulence gene patterns (Kemmett *et al*., 2013).

1.5.3 Genomic Surveillance as a Tool for Future Outbreaks

Most early research into colibacillosis was biased by PCR screening and the search for a single, defined virulotype amongst APEC isolates (Johnson *et al.,* 2006; Schouler *et al.,* 2012;

Mageiros *et al*., 2021). Genomic analysis is a powerful tool that could identify regions of interest in APEC genomes. In the largest genome-wide study done to-date, Mageiros *et al*. created a machine learning model to try and predict the virulence of an APEC isolate using whole genome data from over 500 *E. coli* isolates (2021). The model included genes present in all *E. coli* lineages and ranked them by estimated importance in infection (Mageiros *et al*., 2021). A model trained with data from all *E. coli* phylogroups was able to reach a predictive accuracy of almost 77%. A second model was designed to predict the disease status of the bird (healthy vs. diseased) and trained using data from three phylogroups-A, B1, and B2 (Mageiros *et al*., 2021). This model achieved an accuracy of 73-75%. Finally, they attempted to develop a set of 10 APEC-defining genes using whole genome data. Their predictive set of genes performed no better than previous multiplex PCR screening sets: the panel had a predictive accuracy of 73.5% (Mageiros *et al*., 2021).

The idea that a specific set of genes is required for virulence has focused the direction of colibacillosis research for decades (Skyberg *et al*., 2003; Johnson *et al*., 2008b; Schouler *et al*., 2014). In the Mageiros *et al*. study they used a genome-wide association study (GWAS) to try and identify differences in the genomes of presumed pathogenic *E. coli* and *E. coli* isolated from healthy broilers (2021). This study included 309 APEC strains and 234 strains isolated from the feces of healthy broilers. Out of over 15,000 genes analyzed they identified 65 core genes and 78 accessory genes associated with pathogenicity (Mageiros *et al*., 2021). One possible takeaway from this study is that the concept of a single APEC pathotype is overly simplistic (Collingwood *et al*., 2014). A 2020 study investigating enterohemorrhagic *E. coli* infections in a mouse model of disease determined that there were multiple combinations of effector genes sufficient to cause disease (Ruano-Gallego *et al*., 2020). Different effector combinations induced different immune responses and caused slightly different manifestations of disease (Ruano-Gallego *et al*., 2020). In a parallel with APEC strains, many of these effector genes are present on plasmids (Johnson *et al*., 2008a; Ruano-Gallego *et al*., 2020). Rather than a single set of genes required for infection, it is likely that there are many combinations of genes sufficient to cause infection. Strains with a greater number of virulence genes are more likely to be deadly primary pathogens capable of causing systemic infection (Johnson *et al*., 2008a; Collingwood *et al*., 2012; Mehat *et al*., 2020).

There are two major weaknesses in the study by Mageiros *et al*.: the avian sources and geographic origins of the study strains analyzed (2021). To gather enough isolates for the study the

authors used a combination of collected strains and strains from previously published studies (Mageiros *et al*., 2021). Strains collected by the authors were recovered from commercial broilers, commercial turkeys, wildfowl, and gulls. Strains were considered APEC based on either recovery from various extraintestinal organs or the symptoms of the animal prior to autopsy (Mageiros *et al*., 2021). Strains from previously published collections also varied by poultry type and isolation source (Mageiros *et al*., 2021). A discussed in section 1.5.2., the disease manifestation of the bird may provide important information about the pathogenicity of the APEC strain. Including APEC strains representing multiple manifestations of disease sampled from different kinds of poultry introduces additional variables into the study. Different virulence profiles may cause different manifestations of disease (Ruano-Gallego *et al*., 2020). Strains analyzed in this study also came from a broad geographic area. One set of isolates represents broilers from Nordic farms, while the others consisted of global APEC genomes uploaded to EnteroBase and GenBank (Ronco *et al*., 2017; Jorgensen *et al*; 2019; Najafi *et al*., 2019; Mageiros *et al*., 2021). As previously discussed, between farms there are typically unique virulence patterns identified in APEC isolates (Kemmett *et al*., 2013). Previous studies have failed to identify global APEC lineages (Riley, 2014; Cunha *et al*., 2017; Manges *et al*., 2019). Analyzing APEC strains with such a broad brush may fail to identify more subtle similarities between local outbreaks, or even country-specific strains.

1.6 Understanding APEC Isolates Found on Saskatchewan Broiler Farms

The aim of this thesis was to investigate the differences in *E. coli* that cause colibacillosis in broilers and *E. coli* that reside in the intestinal tract but do not cause disease. To that end, we designed a methodology to collect, sequence, and characterize APEC and commensal strains sourced from colibacillosis outbreaks on farms within two hours of Saskatoon, Saskatchewan. We hypothesized that there are genetic differences between systemic and cecal *E. coli*, and systemic *E. coli* are virulent outbreak strains rather than opportunistic pathogens. Our objectives were as follows: 1) to isolate *E. coli* from diseased and healthy broilers from the same flocks, allowing us to match systemic and cecal isolates by farm for future genomic analysis, 2) screen all isolates for biofilm formation, antimicrobial resistance, and disinfectant resistance and, 3) select 96 *E. soli* strains for whole genome sequencing using Nanopore and Illumina technology to generate highly accurate genomes with the correct gene order. We compared draft genome assemblies to identify

potential differences in commensal and pathogenic *E. coli* present on Saskatchewan broiler farms. Chapter 2 of this thesis introduces preliminary whole genome sequencing results which identified potential genetic differences in systemic and cecal isolates. In chapter 3 I present a methodology to capture and identify avian *E. coli* plasmids missed using our initial whole genome sequencing protocol. Finally, in chapter 4 I analyzed biofilm formation, antimicrobial resistance, and disinfectant trends in systemic and cecal isolates to look for phenotypic differences in the two groups.

1.6.1 Collection and Whole Genome Sequencing of Saskatchewan *E. coli*

To understand how APEC strains cause systemic disease in broiler chickens we designed a study which sampled broiler chickens from a small geographic area using a strict definition of disease. To qualify for our study, colibacillosis septicemia was confirmed via both veterinary necropsies to identify the presence of characteristic lesions on the heart, liver, and spleen typical of systemic disease, and matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry to confirm sepsis was caused by *E. coli*. This contrasts with previously published papers that include *E. coli* isolated from chickens with other forms of disease (yolk sac, air sac, and skin infections), from instruments of meat processing and packing facilities, and from retail chicken meat (Skyberg *et al.,* 2007; Nielsen *et al*., 2018; Mageiros *et al.,* 2021). After a confirmed colibacillosis diagnosis, producers submitted healthy, asymptomatic broilers from the same flock to serve as a comparison group. We isolated *E. coli* from the cecal contents of healthy broilers to avoid contamination with APEC strains from the external barn environment. We have developed a protocol for whole genome sequencing of APEC and AFEC isolates using a modified Nanopore sequencing protocol (Lohman *et al*., 2016). I confirmed that this methodology produced sequencing reads of sufficient length and quantity for assembling *E. coli* genomes by assembling 96 preliminary draft genomes. Using those draft genomes we identified potential differences in plasmid content, phylogeny, and genes of interest in systemic and cecal isolates.

1.6.2 Plasmid Content

Many whole genome sequencing protocols include a size selection step to optimize long read sequencing. Because our initial whole genome sequencing protocol excluded DNA fragments smaller than 500bp in length, we may have failed to identify some plasmids in our initial genome sequencing runs. We developed a plasmid capture methodology with a dedicated plasmid

extraction step to test whether our whole genome sequencing method was sufficient for identifying plasmids in avian *E. coli* isolates.

We tested combinations of six popular commercial plasmid extraction kits and two Nanopore library preparation kits to determine the best method for sequencing *E. coli* plasmids. Three commercial plasmid kits produced average read lengths over 3kb, which produced the best *de novo* assemblies after sequencing. We hypothesized that some plasmids would become nicked during the plasmid extraction process, allowing for a standard ligation-based library prep. The alternative, a transposon-barcoding system, would fragment circularized DNA before sequencing. Ultimately, a combination of a plasmid extraction kit that produces average read lengths of over 3kb plus the ligation sequencing preparation produced the best assemblies. We were able to identify an additional three plasmids in a systemic isolate and one additional plasmid in a cecal isolate.

1.6.3 Biofilm, Antimicrobial Resistance, and Disinfectant Resistance

We screened 96 avian *E. coli* for antimicrobial resistance and their ability to form a biofilm. Multi-drug resistant *E. coli* are a significant concern to producers trying to manage disease outbreaks among their flocks. Bacteria that persist in the inner layers of a biofilm are protected from high concentrations of antimicrobials and disinfectants and may acquire resistance through exposure to sublethal concentrations (Fitzgerald, 2019; Santos-Lopez *et al.*, 2019). We hypothesized that APEC persist in the poultry barn environment in bacterial biofilms before infecting broilers. Thus, antimicrobial and disinfectant resistance, as well as biofilm formation may be linked in pathogenic isolates.

Antimicrobial resistance is widespread among *E. coli* isolated from Saskatchewan broiler farms, but we were unable to correlate resistance with biofilm formation. *E. coli* isolated from diseased broilers (systemic *E. coli*) are more likely to be drug-resistant than *E. coli* isolated from the cecal contents of healthy broilers (cecal *E. coli*). The majority of both systemic and cecal *E. coli* produce the biofilm components curli and cellulose. Out of 96 *E. coli*, only 2 failed to produce either cellulose or curli. Both were cecal strains. Systemic *E. coli* formed significantly stronger biofilms in rich media compared to cecal isolates. However, overall, there was no significant difference in total biofilm formation between systemic and cecal isolates. The strongest biofilm producing strains possessed plasmid-mediated virulence genes that enhance biofilm formation. Commercial concentrations of disinfectants were sufficient to kill planktonic and biofilm cells,

however bacteria in a biofilm were resistant to higher concentrations of disinfectants compared to planktonic cells.

1.7 References

Abdallah, M., Benoliel, C., Drider, D., Dhulster, P., Chihib, N. (2014). Biofilm formation and persistence on abiotic surfaces in the context of food and medical environments. *Archives of Microbiol*. 196, 453-472. doi: 10.1007/s00203-014-0983-1.

Abram, K., Uadondo, Z., Bleker, C., Wanchai, V., Wassenaar, T.M., Robeson, M.S., Ussery, D.W. (2020). What can we learn from over 100,000 *Escherichia coli* genomes? *BioRxiv Preprint*. doi: 10.1101/708131.

Accogli, M., Fortini, D., Giufre, M., Graziani, C., Dolejska, M,. Carattoli, A., Cerquetti, M. (2013). IncI 1 plasmids associated with the spread of CMY-2, CTX-M-I, and SHV-12 in *Escherichia coli* of animal and human origin. *Clin. Microbiol. and Infect*. 19: E238-240. doi: 10.1111/1469-0691.12128.

Aidara-Kane, A., Angulo, F.J., Conly, J.M., Minato, Y., Silbergeld, E.K., McEwen, S.A., *et al*. (2018). World Health Organization (WHO) guidelines on use of medically important antimicrobials in food-producing animals. *Antimicrob. Resist Infect. Control*. 17(7). doi: 10.1187/s13756-017-0294-9.

Amerah, A.M., van Rensburg, C.J., Plumstead, P.W., Kromm, C., Dunham, S. (2012). Effect of feeding diets containing a probiotic or antibiotic on broiler performance, intestinal mucosa-associated avian pathogenic *E. coli* and litter water-soluble phosphorus. *J. Appl. Animal Nutr*. 1(E7). doi: 10.1017/jan.2013.4.

Antao, E., Ewers, C., Gurlebeck, D., Preisinger, R., Homeier, T., Li, G., Wieler, L.H. (2009). Signature-tagged mutagenesis in a chicken infection model leads to the identification of a novel avian pathogenic *Escherichia coli* fimbrial adhesin. *PLoS One* 4(11): 37796. doi: 10.1371/journal.pone.0007796.

Awad, A.M., El-Shall, N.A., Khalil, D.S., El-Hack, M.E., Swelum, A.A., Mahmoud, A.H., Ebaid, H., *et al*. (2020). Incidence, pathotyping, and antibiotic susceptibility of avian pathogenic *Escherichia coli* among diseased broiler chicks. *Pathogens*. 9, 114. doi: 10.3390/pathogens9020114.

Barnhart, M.M., Chapman, M.R. (2006). Curli biogenesis and function. *Annu. Rev. Microbiol*. 60: 131-47. doi: 10.1146/annurev.micro.60.080805.142106.

Bin Kim, Y., Yoon, M.Y., Ha, J.S., Seo, K.W., Noh, E.B., Son, S.H., Lee, Y.J. (2019). Molecular characterization of avian pathogenic *Escherichia coli* from broiler chickens with colibacillosis. *Poultry Sci*. 99(2). Doi: 10.1016/j.psj.2019.10.047.

Brombacher, E., Dorel, C., Zehnder, A.J.B., Landini, P. (2003). The curli biosynthesis regulator CsgD coordinates the expression of both positive and negative determinants for biofilm formation in *Escherichia coli*. *Microbiol*. *Soc*. 149(10). doi: 10.1099/mic.0.26306-0.

Clermont, O., Bonacorsi, S., Bingen, E. (2000). Rapid and simple determination of the *Escherichia coli* phylogenetic group. *App*. *Environ*. *Microbiol*. 66(10): 4555-8. doi: 10.1128/AEM.66.10.4555-4558.2000.

Clermont, O., Olier, M., Hoede, C., Diancourt, L., Brisse, S., Kerodean, M., Glodt, J., *et al*. (2011). Animal and human pathogenic *Escherichia coli* strains share common genetic backgrounds. *Infect*. *Gen*. *Evol*. 11(3): 654-662. doi: 10.1016/j.meegid.2011.02.005.

Clermont, O., Christenson, J.K., Denamur, E., Gordon, D.M. (2012). The Clermont *Escherichia coli* phylo-typing method revisited: improvements of specificity and detection of new phylo-groups. *Environ*. *Microbiol*. *Rep*. 5(1): 58-65. doi: 10.1111/1758-2229.12019.

Clermont, O., Christensen, J.K., Denamur, E., Gordon, D.M. (2013). The Clermont *Escherichia coli* phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. *Environ*. *Microbiol*. *Rep*. 5(1): 58-65.

Clermont, O., Gordon, D., Denamur, E. (2015). Guide to the various phylogenetic classification schemes for *Escherichia coli* and the correspondence among schemes. *Microbiol*. 161: 980-988. doi: 10.1099/mic.0.000063.

Clermont, O., Dixit, O.V.A., Vangchhia, B., Condamine, B., Dion, S., Bridier-Nahmias, A., *et al*. (2019). Characterization and rapid identification of phylogroup G in *Escherichia coli*, a lineage with high virulence and antibiotic resistance potential. *Environ*. *Microbiol*. 21(8): 3107-3117. doi: 10.1111/1462-2920.14713.

Clermont, O., Condamine, B., Dion, S., Gordon, D.M., Denamur, E. (2021). The E phylogroup of *Escherichia coli* is highly diverse and mimics the whole *E. coli* species population structure. *Environ*. *Microbiol*. doi: 10.111/1462-2920.15742.

Collingwood, C., Kemmett, K., Williams, N., Wigley, P. (2014). Is the concept of avian pathogenic *Escherichia coli* as a single pathotype fundamentally flawed? *Front*. *Vet*. *Sci*. 1(5). doi: 10.3389/fvets.2014.00005.

Cordoni, G., Woodward, M.J., Wu, H., Alanazi, M., Wallis, T., La Ragione, R.M. (2016). Comparative genomics of European avian pathogenic *E. coli* (APEC). *BMC Genomics*. 17, 960. doi: 10.1186/s12864-016-3289-7.

Course, C.E., Boerlin, P., Slavic, D., Vaillancourt, J., Guerin, M.T. (2021). Factors associated with *Salmonella enterica* and *Escherichia coli* during downtime in commercial broiler chicken barns in Ontario. *Poul. Sci*. 100(5): 101065. doi: 10.1016/j.psj.2021.101065.

Cunha, M.P.V., Saidenberg, A.B., Moreno, A.M., Ferreira, A.J.P., Vieira, M.A.M., Gomes, T.A.T., Knobl, T. (2017). Pandemic extra-intestinal pathogenic *Escherichia coli* (ExPEC) clonal group O6-B2-ST73 as a cause of avian colibacillosis in Brazil. *PLoS One* 12(6): e0178970. doi: 10.1371/journal.pone.0178970.

Cusumano, C.K., Hung, C.S., Chen, S.L., Hultgren, S.J. (2010). Virulence plasmid harbored by uropathogenic *Escheria coli* functions in acute stages of pathogenesis. *Infect. Immun*. 78(4): 1457-1467. doi: 10.1128/IAI.01260-09.

Cusumano, C.K., Hung, C.S., Chen, S.L., Hultgren, S.J. (2020). Virulence plasmid harboured by uropathogenic *Escherichia coli* functions in acute stages of pathogenesis. *Infect and Immun*. 78(4). doi: 10.1128/IAI.01260-09.

Davies, R., Wales, A. (2019). Antimicrobial resistance on farms: a review including biosecurity and the potential role of disinfectants in resistance selection. *Comp. Reviews in Food Safety*. doi: 10.1111/1541-4337.12438.

Davis, G.S., Waits, K., Nordstrom, L., Grande, H., Weaver, B., Papp, K., Horwinski, J., Koch, B., Hungate, B.A., Liu, C.M., Price, L.B. (2018). Antibiotic-resistant *Escherichia coli* from retail poultry meat with different antibiotic use claims. *BMC Microbiol*. 18:174. doi: 10.1186/s12866-018-1322-5.

DebRoy, C., Fratamico, P.M., Roberts, E. (2018). Molecular serogrouping of *Escherichia coli*. *Animal Health Rev. Res*. 19(1): 1-16. doi: 10.1017/s1466252317000093.

Delannoy, S., Schouler, C., Souillard, R., Yousfi, L., Le Devendec, L., Lucas, C., *et al*. (2021). Diversity of *Escherichia coli* strains isolated from day-old broilers, their environment, and colibacillosis lesions in 80 flocks in France. *Vet. Microbiol*. 252: 108923. doi: 10.1016/j.vetmic.2020.108923.

Dhillon, A.S., Jack, O.K. (1996). Two outbreaks of colibacillosis in commercial caged layers. *Avian Dis*. 40: 742-746. doi: 10.2307/1592290.

Diaz, J.M., Dozois, C.M., Avelar-Gonzalez, F.J., Hernandez-Cuellar, E., Pokharel, P., de Santiago, A.S., Guerrero-Barrera, A.L. (2020). The vacuolating autotransporter toxin (Vat) of *Escherichia coli* causes cell cytoskeleton changes and produces non-lysosomal vacuole formation in bladder epithelial cells. *Front. Cell. Infect. Microbiol*. doi: 10.3389/fcimb.2020.00299.

Duriez, P., Clermont, O., Bonacorsi, S., Bingen, E., Chanventre, A., Elion, J., *et al*. (2001). Commensal *Escherichia coli* isolates are phylogenetically distributed among geographically distinct human populations. *Microbiol. Soc*. 147(6). doi: 10.1099/00221287-147-6-1671.

Dziva, F., Stevens, M.P. (2008). Colibacillosis in poultry: unravelling the molecular basis of virulence of avian pathogenic *Escherichia coli* in their natural hosts. *Avian Pathol*. 37(4): 355-366. doi: 10.1080/03079450802216652.

Ewers, C., Antao, E., Diehl, I., Philipp, H., Wieler, L.H. (2009). Intestine and environment of the chicken as reservoirs for extraintestinal pathogenic *Escherichia coli* strains with zoonotic potential. *Appl. Environ. Microbiol*. 75(1): 184-192. doi: 10.1128/AEM.01324-08.

Felgner, S., Frahm, M., Kocijancic, D., Rohde, M., Eckweiler, D., Bielecka, A., Bueno, E. et al. (2016). *aroA*-Deficient *Salmonella enterica* serovar Typhimurium is more than a metabolically attenuated mutant. *mBio* 7(5): 301220-16. doi: 10.1128/mBio.01220-16.

Filho, T.F., Favaro, C., Ingberman, M., Beirao, B.C.B., Inoue, A., Gomes, L., Caron, L.F. (2013). Effect of spray *Escherichia coli* vaccine on the immunity of poultry. *Avian Dis*. 57(3): 671-676. doi: 10.1637/10456-112612-ResNote.1.

Flament-Simon, S., Duprilot, M., Mayer, N., Garcia, V., Alonso, M.P., Blanco, J., Nicolas-Chanoine, M. (2019). Association between kinetics of early biofilm formation and clonal lineage in *Escherichia coli*. *Front Microbiol*. 10: 1183. doi: 10.3389/fmicb.2019.01183.

Fiorentin, L., Vieria, M.D., Barioni, W. (2005). Use of lytic bacteriophages to reduce Salmonella Enteritidis in experimentally contaminated chicken cuts. *Poultry Sci*. 7(4). doi: 10.1590/S1516-635X2005000400010.

Fitzgerald, D.M. (2019). Bacterial evolution: the road to resistance. *eLife*. 8: 352092. doi: 10.7554/eLife.52092.

Forde, B.M., Zowawi, H.M., Harris, P.N.A., Roberts, L., Ibrahim, E., Shaikh, N., *et al*. (2018). Discovery of *mcr-1*-mediated colistin resistance in a highly virulent *Escherichia coli* lineage. *mSphere* 3:e00486-18. doi: 10.1128/mSphere.00486-18.

Germon, P., Chen, Y., He, L., Blanco, J.E., Bree, A., Schouler, C., Huang, S., Moulin-Schouler, M. (2005). *ibeA*, a virulence factor of avian pathogenic *Escherichia coli*. *Microbiol*. 151: 1179-1186. doi: 10.1099/mic.0.27809-0.

Gophna, U., Barlev, M., Seijffers, R., Oelschlager, T.A., Hacker, J., Ron, E.Z. (2001). Curli fibers mediate internalization of *Escherichia coli* by eukaryotic cells. *Infect. Immun*. 69(4): 2659-65. doi: 10.1128/IAI.69.4.2659-2665.2001.

Gordon, D.M., Cowling, A. (2003). The distribution and genetic structure of *Escherichia coli* in Australian vertebrates: host and geographic effects. *Microbiol. Soc*. 149(12). doi: 10.1099/mic.0.26486-0.

Guabiraba, R., Schouler, C. (2015). Avian Colibacillosis: still many black holes. *FEMS Micro. Lett*. 362(15): fnv118. doi: 10.1093/femsle/fnv118.

Gualdi, L., Tagliabue, L., Bertagnoli, S., Ierano, T., De Castro, C., Landini, P. (2008). Cellulose modulates biofilm formation by counteracting curli-mediated colonization of solid surfaces in *Escherichia coli*. *Microbiol*. 154: 2017-2024. doi: 10.1099/mic.0.2008/018093-0.

Gyles, C.L. (2007). Shiga toxin-producing *Escherichia coli*: An overview. *J. Animal Sci*. 85: E45-62. doi: 10.2527/jas.2006-508.

Halfaoui, Z., Menoueri, N.M., Bendali, L.M. (2017). Serogrouping and antibiotic resistance of *Escherichia coli* isolated from broiler chickens with colibacillosis in center of Algeria. *Vet. World*. 10(7): 830-835. doi: 10.14202/vetworld.2017.830-835.

Heinemann, C., Leubner, C.D., Savin, M., Sib, E., Schmithausen, R.M., Steinhoff-Wagner, J. (2020). Research note: tracing pathways of entry and persistence of facultative pathogenic and antibiotic-resistant bacteria in a commercial broiler farm with substantial health problems. *Poultry Sci*. 99(11): 5481-5486. doi: 10.1016/j.psj.2020.08.050.

Hu, Y., Zhang, Q., Meitzler, J.C. (1999). Rapid and sensitive detection of *Escherichia coli* O157:H7 in bovine faeces by a multiplex PCR. *J. Appl. Microbiol*. 87(6): 867-876.

Hutton, T.A., Innes, G.K., Harel, J., Garneau, P., Cucchiara, A., Schifferli, D.M., *et al*. (2017). Phylogroup and virulence genes association with clinical characteristics of *Escherichia coli* urinary tract infections from dogs and cats. *J. Vet. Diagnost. Invest*. 30(1): 64-70. doi: 10.1177/1040638717729395.

Ikuta, N., Sobral, F., Lehman, F.K.M., da Silveira, P., de Carli, S., Casanova, Y.S., *et al*. (2014). Taqman real-time PCR assays for rapid detection of avian pathogenic *Escherichia coli* isolates. *Avian Dis*. 58(4): 628-631. doi: 10.1637/10871-052414.

Joensen, K.G., Tetzchner, A.M.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. *J. Clin. Microbiol*. 53(8): 2410-2426. doi: 10.1128/JCM.00008-15.

Johnson, T.J., Siek, K.E., Johnson, S.J., Nolan, L.K. (2006). DNA sequence of a ColV plasmid and prevalence of selected plasmid-encoded virulence genes among avian *Escherichia coli* strains. *J. Bacteriol*. 188(2): 745-58. doi: 10.1128/JB/188/2/745-758.2006.

Johnson, T.J., Wannemuehler, Y., Johnson, S.J., Stell, A.L., Doetkott, C., Johnson, J.R., Kim, K.S., Spanjaard, L., Nolan, L.K. (2008a). Comparison of extraintestinal pathogenic *Escherichia coli*

strains from human and avian sources reveals a mixed subset representing potential zoonotic pathogens. *Appl. Environ. Microbiol.* 74(22): 7043-7050.

Johnson, T.J., Wannemuehler, Y., Doetkott, C., Johnson, S.J., Rosenberger, S.C., Nolan, L.K. (2008b). Identification of minimal predictors of avian pathogenic *Escherichia coli* virulence for use as a rapid diagnostic tool. *J. Clin. Microbiol.* 46(12): 3987-96. doi: 10.1128/JCM.00816-08.

Johnson, T.J., Logue, C.M., Johnson, J.R., Koskowski, M.A., Sherwood, J.S., Barnes, H.J., DebRoy, C., Wannemuehler, Y.M., Obata-Yasuoka, M., Spanjaard, L., Nolan, L.K. (2012). Associations between multidrug resistance, plasmid content, and virulence potential among extraintestinal pathogenic and commensal *Escherichia coli* from humans and poultry. *Food. Pathog. And Dis.* 9(1). doi: 10.1089/fpd.2011/0961.

Jorgensen, S., Stegger, M., Kudirkiene, E., Lilje, B., Poulsen, L.L., Ronco, T., Dos Santos, T.P., Kiil, K., Bisgaard, M., *et al.* (2019). Diversity and population overlap between avian and human *Escherichia coli* belonging to sequence type 95. *mSphere.* 4(1): 00333-18. doi: 10.1128/mSphere.00333-18.

Joshi, S., Singh, R., Singh, S.P. (2012). Antibiotic resistance profile of *Escherichia coli* isolates from Colibacillosis in and around Pantnagar, India. *Vet. World.* 5(7): 405-408. doi: 10.5455/vetworld.2012.405-408.

Kathayat, D., Lokesh, D., Ranjit, S., Rajashekara, G. (2021). Avian pathogenic *Escherichia coli* (APEC): an overview of virulence and pathogenesis factors, zoonotic potential, and control strategies. *Pathog.* 10(4): 467. doi: 10.3390/pathogens10040467.

Kazemnia, A., Ahmadi, M., and Dilmaghani, M. (2014). Antibiotic resistance pattern of different *Escherichia coli* phylogenetic groups isolated from human urinary tract infection and avian colibacillosis. *Iran Biomed. J.* 18(4): 219-224. doi: 10.6091/ibj.1394.2014.

Kim, Y.B., Yoon, M.Y., Ha, J.S., Seo, K.W., Noh, E.B., Son, S.H., Lee, Y.J. (2020). Molecular characterization of avian pathogenic *Escherichia coli* from broiler chickens with colibacillosis. *Poultry Sci.* 99(2): 1088-1095. doi: 10.1016/j.psj.2019.10.047.

Kemmett, K., Humphrey, T., Rushton, S., Close, A., Wigley, P., Williams, N.J. (2013). A longitudinal study simultaneously exploring the carriage of APEC virulence associated genes and the molecular epidemiology of faecal and systemic *E. coli* in commercial broiler chickens. *PLoS ONE.* 8(6): e66749. doi: 10.1371/journal.pone.0067749.

Koga, V.L., Tomazetto, G., Cyoia, P.S., Neves, M.S., Vidotto, M.C., Nakazato, G., *et al.* (2014). Molecular screening of virulence genes in extraintestinal pathogenic *Escherichia coli* isolated from human blood culture in Brazil. *Blood Stream Infect.* 465054. doi: 10.1155/2014/465054.

Koutsianos, D., Gantelet, H., Franzo, G., Lecoupeur, M., Thibault, E., Cecchinato, M., Koutoulis, K.C. (2020). An assessment of the level of protection against colibacillosis conferred by several autogenous and/or commercial vaccination programs in conventional pullets upon experimental challenge. *Vet. Sci*. 7,80. doi: 10.3390/vetsci7030080.

Landers, T.F., Cohen, B., Wittum, T.E., Larson, E.L. (2012). A review of antibiotic use in food animals: perspective, policy, and potential. *Public Health Rep*. 127(1): 4-22. doi: 10.1177/003335491212700103.

Landgraf, T.N., Berlese, A., Fernandes, F.F., Milanezi, M.L., Martinez, R., Panunto-Castelo, A. (2012). The ferric aerobactin receptor IutA, a protein isolated on agarose column, is not essential for uropathogenic *Escherichia coli* infection. *Rev. Lat. Am. Enfermagem*. 20(2): 340-5. doi: 10.1590/s0104-11692012000200017.

Larsonneur, F., Martin, F.A., Mallet, A., Martinez-Gil, M., Semetey, V., Ghigo, J., Beloin, C. (2016). Functional analysis of *Escherichia coli* Yad fimbriae reveals their potential role in environmental persistence. *Environ. Microbiol*. 18(12): 5228-5248. doi: 10.1111/1462-2920.13559.

Lau, G.L., Sieo, C.C., Tan, W.S., Hair-Bejo, M., Jalila, A., Ho, Y.W. (2010). Efficacy of a bacteriophage isolated from chickens as a therapeutic agent for colibacillosis in broiler chickens. *Poultry Sci*. 89(12). doi: 10.3382/ps.2010-00904.

Lee, K., Lillehoj, H.S., Siragusa, G.R. (2010). Direct-fed microbials and their impact on the intestinal microflora and immune system of chickens. *Japan Poultry Sci. Assoc*. doi: 10.2141/jpsa.009096.

Li, G., Feng, Y., Kariyawasam, S., Tivendale, K.A., Wannemuehler, Y., Zhou, F., Logue, C.M., Miller, C.L., Nolan, L.K. (2010). AatA is a novel autotransporter and virulence factor of avian pathogenic *Escherichia coli*. *Infect. Immun*. 78(3): 898-906. doi: 10.1128/IAI.00513-09.

Li, L., Thofner, I., Christensen, J.P., Ronco, T., Pedersen, K., Olsen, R.H. (2017). Evaluation of the efficacy of an autogenous *Escherichia coli* vaccine in broiler breeders. *Avian Pathol*. 46:3, 300-308. doi: 10.1080/03079457.2016.1267857.

Lindstedt, B., Finton, M.D., Porcellato, D., Brandal, L.T. (2018). High frequency of hybrid *Escherichia coli* strains with combined intestinal pathogenic *Escherichia coli* (IPEC) and extraintestinal pathogenic *Escherichia coli* (ExPEC) virulence factors isolated from human fecal samples. *BMC Infect. Dis*. 18(544). doi: 10.1186/s12879-3449-2.

Liu, C.M., Stegger, M., Aziz, M., Johnson, T.J., Waits, K., Nordstrom, L., Gauld, L., *et al*. (2018). *Escherichia coli* ST131-H22 as a foodborne uropathogen. *mBio*. 9(4): 300470-18. doi: 10.1128/mBio.00470-18.

Leveille, S., Caza, M., Johnson, J.R., Clabots, C., Sabri, M., Dozois, C.M. (2020). Iha from an *Escherichia coli* urinary tract infection outbreak clonal group A strain is expressed in vivo in the mouse urinary tract and functions as a catecholate siderophore receptor. *J. Clin. Microbiol.* 74(6). doi: 10.1128/IAI.00107-06.

Lloyd, A.L., Rasko, D.A., Mobley, H.L.T. (2007). Defining genomic islands and uropathogen-specific genes in uropathogenic *Escherichia coli*. *J. Bacteriol.* 189(9). doi: 10.1128/JB.10744-06.

Lorenzo, V., Neilands, J.B. (1986). Characterization of iucA and iucC genes of the aerobactin system of plasmid ColV-K30 in *Escherichia coli*. *J. Bacteriol.* 167(1): 350-355. doi: 10.1128/jb.167.1.350-355.1986.

Lutful Kabir, S.M. (2010). Avian colibacillosis and salmonellosis: a closer look at epidemiology, pathogenesis, diagnosis, control and public health concerns. *Int. J. Environ. Res. Public Health* 7, 89-114. doi: 10.3390/ijerph7010089.

Lyhs, U., Ikonen, I., Pohjanvirta, T., Raninen, K., Perko-Makela, P., Pelkonen, S. (2012). Extraintestinal pathogenic *Escherichia coli* in poultry meat products on the Finnish retail marker. *Acta Vet. Scand* 54(64). doi: 10.1186/1751-0147-54-64.

MacKenzie, K.D., Palmer, M.B., Koster, W.L., White, A.P. (2017). Examining the link between biofilm formation and the ability of pathogenic *Salmonella* strains to colonize multiple host species. *Front. Vet. Sci.* doi: 10.3389/fvets.2017.00138.

Mageiros, L., Meric, G., Bayliss, S.C., Pensar, J., Pascoe, B., Mourkas, E., Calland, J.K., *et al.* (2021). Genome evolution and the emergence of pathogenicity in avian *Escherichia coli*. *Nature Comm.* 12:765. doi: 10.1038/s41467-021-20988-w.

Maluta, R.P., Logue, C.M., Casas, M.R.T., Meng, T., Guastalli, E.A.L., Rojas, T.C.G., Montelli, A.C., *et al.* (2014). Overlapped sequence types (STs) and serogroups of avian pathogenic (APEC) and human extra-intestinal pathogenic (ExPEC) *Escherichia coli* isolated in Brazil. *PLoS ONE.* 9(8): e105016. doi: 10.1371/journal.pone.0105016.

Mamani, R., Flament-Simon, S.C., Garcia, V., Mora, A., Alonso, M.P., Lopez, C., Garcia-Menino, I. *et al.* (2019). Sequence types, clonotypes, serotypes, and virotypes of extended-spectrum B-lactamase-producing *Escherichia coli* causing bacteremia in a Spanish hospital over a 12-year period (2000-2011). *Front. Microbiol.* 10:5130. doi: 10.3389/fmicb.2019.10530.

Manges, A.R., Min Geum, H., Guo, A., Edens, T.J., Fibke, C.D., Pitout, J.D.D. (2019). Global extraintestinal pathogenic *Escherichia coli* (ExPEC) lineages. *Clin. Microbiol. Rev.* 32(3). doi: 10.1128/CMR.00135-18.

Marin, C., Hernandiz, A., Lainez, M. (2009). Biofilm development capacity of *Salmonella* strains isolated in poultry risk factors and their resistance against disinfectants. *Poul. Sci*. 88(2): 424-431. doi: 10.3382/ps.2008-00241.

McDaniel, T.K., Jarvis, K.G., Donnenberg, M.S., Kaper, J.B. (1995). A genetic locus of enterocyte effacement conserved among diverse enterobacterial pathogens. *PNAS*. 92(5): 1664-1668. doi: 10.1073/pnas.92.5.1664.

Meena, P.R., Yadav, P., Hemlata, H., Tejavath, K.K., Singh, A.P. (2021). Poultry-origin extraintesintal *Escherichia coli* strains carrying the traits associated with urinary tract infection, sepsis, meninigitis, and avian colibacillosis in India. *J. Appl. Microbiol*. 130(6): 2087-2101. doi: 10.1111/jam.14905.

Mehat, J.W., van Vliet, A.H.M., Ragione, R.M. (2021). The avian pathogenic *Escherichia coli* (APEC) pathotype is comprised of multiple distinct, independent genotypes. *Avian Pathol*. 1(15). doi: 10.1080/03079457.2021.1915960.

Mellata, M., Ameiss, K., Mo, H., Curtiss, R. (2010). Characterization of the contribution to virulence of three large plasmids of avian pathogenic *Escherichia coli* $_{X}$7122 (O78:K80:H9). *Infect Immun*. 78(4): 1528-41. doi: 10.1128/IAI.00981-09.

Mellata, M. (2013). Human and avian extraintestinal pathogenic *Escherichia coli*: infections, zoonotic risks, and antibiotic resistance trends. *Foodborn Pathog. Dis*. 10(11): 916-932. doi: 10.1089/fpd.2013.1533

Mitchell, N.M., Johnson, J.R., Johnston, B., Curtiss, R., Mellata, M. (2015). Zoonotic potential of *Escherichia coli* isolates from retail chicken meat products and eggs. *Appl. Environ. Microbiol*. 81(3): 1177-87. doi: 10.1128/AEM.03524-14.

Mohamed, L., Ge, Z., Yuehua, L., Yubin, G., Rachid, K., Mustapha, O., Junewi, W., Karine, O. (2017). Virulence traits of avian pathogenic (APEC) and fecal (AFEC) *E. coli* isolated from broiler chickens in Algeria. *Tropical Animal Health and Prod*. 50, 547-553. doi: 10.1007/s11250-017-1467-5.

Moissenet, D., Salauze, B., Clermont, O., Bingen, E., Arlet, G., Denamur, E., *et al*. (2010). Meningitis caused by *Escherichia coli* producing TEM-52 extended-spectrum B-lactamase within an extensive outbreak in a neonatal ward: epidemiological investigation and characterization of the strain. *J. Clin. Microbiol*. 48(7): 2459-2463. doi: 10.1128/JCM.00529-10.

Najafi, S., Rahimi, M., Nikousefat, Z. (2019). Extraintestinal pathogenic *Escherichia coli* from human and avian origin: detection of the most common virulence-encoding genes. *Vet Res. Forum*. 10(1): 43-49. doi: 10.30466/vrf.2019.34307.

Naves, P., del Prado, G., Huelves, L., Gracia, M., Ruiz, V., Blanco, J., Dahbi, G., Blanco, M., del Carmen Ponte, M., Soriano, F. (2008). Correlation between virulence factors and *in vitro* biofilm formation by *Escherichia coli* strains. *Microb. Pathog.* 45: 86-91.

Newman, D.M., Barbieri, N.L., de Oliviera, A.L., Willis, D., Nolan, L.K., Logue, C.M. (2021). Characterizing avian pathogenic *Escherichia coli* (APEC) from colibacillosis cases, 2018. *PeerJ* 9: 311025. doi: 10.7717/peerj.11025.

Nicholson, B.A., West, A.C., Mangiamele, P., Barbieri, N., Wannemuehler, Y., Nolan, L.K., Logue, C.M., Li, G. (2016). Genetic characterization of ExPEC-like virulence plasmids among a subset of NMEC. *PLos One*.

Nielsen, E.M., Skov, M.N., Madsen, J.J., Lodal, J., Jespersen, J.B., Baggesen, D.L. (2004). Verocytotoxin-producing *Escherichia coli* in wild birds and rodents in close proximity to farms. *Appl. Environ. Microbiol.* 70(11). doi: 10.1128/AEM.70.11.6944-6947.2004.

Nielsen, D.W., Klimavicz, J.S., Cavender, T., Wannemuehler, Y., Barbieri, N.L., Nolan, L.K., Logue, C.M. (2018). The impact of media, phylogenetic classification, and *E. coli* pathotypes on biofilm formation in extraintestinal and commensal *E. coli* from humans and animals. *Front. Microbiol.* 9(902). doi: 10.3389/fmicb.2018.00902.

Payne, J.B., Kroger, E.C., Watkins, S.E. (2005). Evaluation of disinfectant efficacy when applied to the floor of poultry grow-out facilities. *J. Appl. Poul. Res.* 14(2): 322-329. doi: 10.1093/japr/14.2.322.

Peigne, C., Bidet, P., Mahjoub-Messai, F., Plainvert, C., Barbe, V., Medigue, C., Frapy, E., Nassif, X., Denamur, E., Bingen, E., Bonacorsi, S. (2009). The plasmid of *Escherichia coli* strain S88 (O45:K1:H7) that causes neonatal meningitis is closely related to avian pathogenic *E. coli* plasmids and is associated with high-level bacteremia in a neonatal rat meningitis model. *Infect. Immun.* 77(6): 2272-84. doi: 10.1128/IAI.01333-08.

Ragione, R.M., Collighan, R.J., Woodward, M.J. (1999). Non-curliation of *Escherichia coli* O78:K80 isolates associated with IS1 insertion in csgB and reduced persistence in poultry infection. *FEMS Microbiol. Lett.* 175(2): 247-53. doi: 10.1111/j.1574-6968.1999.tb13627.x.

Riley, L.W. (2014). Pandemic lineages of extraintestinal pathogenic *Escherichia coli*. *Clin. Microbiol. Infect.* 20(5): 380-90. doi: 10.1111/1469-0691.12646.

Redweik, G.A., Stromberg, Z.R., Van Goor, A., Mellata, M. (2020). Protection against avian pathogenic *Escherichia coli* and *Salmonella* Kentucky exhibited in chickens given both probiotics and live *Salmonella* vaccine. *Poultry Sci.* 99(2): 752-762. doi: 10.1016/j.psj.2019.10.038.

Rodriguez-Siek, K.E., Giddings, C.W., Doetkott, C., Johnson, T.J., Nolan, L.K. (2005). Characterizing the APEC pathotype. *Vet Res.* 36(2): 241-56. doi: 10.1051/vetres:2004057.

Rodrigues, S.V., Laviniki, V., Borges, K.A., Furian, T.Q., Moraes, H.L.S., Nascimento, V.P., Salle, C.T.P. (2019). Biofilm formation by avian pathogenic *Escherichia coli* is not related to in vivo pathogenicity. *Curr. Microbiol*. 76(2): 194-199. doi: 10.1007/s00284-018-1608-8.

Ronco, T., Stegger, M., Olsen, R.H., Sekse, C., Nordstoga, A.B., Pohjanvirta, T., Lilje, B., Lyhs, U., Andersen, P.S., Pedersen, K. (2017). Spread of avian pathogenic *Escherichia coli* ST117 O78:H4 in Nordic broiler production. *BMC Genomics*. 18(13). doi: 10.1186/s12864-016-3415-6.

Rouquet, G., Porcheron, G., Barra, C., Reperant, M., Chanteloup, N.K., Schouler, C., Gilot, P. (2009). A metabolic operon in extraintestinal pathogenic *Escherichia coli* promotes fitness under stressful conditions and invasion of eukaryotic cells. *J. Bacteriol*. 191(13): 4427-40. doi: 10.1128/JB.00103-09.

Ruano, M., El-Attrache, J., Villegas, P. (2001). Efficacy comparisons of disinfectants used by the commercial poultry industry. *Avian Dis*. 45(4): 972-977.

Ruano-Gallego, D., Sanchez-Garrido, J., Kozik, Z., Nunez-Berrueco, E., Cepeda-Molero, M., Mullineaux-Sanders, C., Clark, J.N., *et al*. (2020). Type III secretion system effectors form robust and flexible intracellular virulence networks. *Science*. 371(1122). doi: 10.1126/science.abc9531.

Sabri, M., Leveille, S., Dozois, C.M. (2006). A SitABCD homologue from an avian pathogenic *Escherichia coli* strain mediates transport of iron and manganese and resistance to hydrogen peroxide. *Microbiol*. 152: 745-758. doi: 10.1099/mic.0.28682-0.

Saldana, Z., Xicohetencatl-Cortes, J., Avelino, F., Phillips, A.D., Kaper, J.B., Puente, J.L., Giron, J.A. (2009). Synergistic role of curli and cellulose in cell adherence and biofilm formation of attaching and effacing *Escherichia coli* and identification of Fis as a negative regulator of curli. *Environ. Microbiol*. 11(4): 992-1006. doi: 10.1111/j.1462-2920.2008.01842.x.

Santos-Lopez, A., Marshall, C.W., Scribner, M.R., Snyder, D.J., Cooper, V.S. (2019). Evolutionary pathways to antibiotic resistance are dependent upon environmental structure and bacterial lifestyle. *eLife*. 8: 347612. doi: 10.7554/eLife.47612.

Sargeant, J.M., Bergevin, M.D., Churchill, K., Dawkins, K., Deb, B., Dunn, J., Logue, C.M., et al. (2019). The efficacy of antibiotics to control colibacillosis in broiler poultry: a systematic review. *Animal Health Res. Reviews*. 20: 263-273. doi: 10.1017/s1466252319000264.

Sarowska, J., Futoma-Koloch, B., Jama-Kmiecik, A., Frej-Madrzak, M., Ksiazczyk, M., *et al*. (2019). Virulence factors, prevalence and potential transmission of extraintesitnla pathogenic *Escherichia coli* isolated from different sources: recent reports. *Gut Pathog*. 11(10). doi: 10.1186/s13099-109-0290-0.

Schouler, C., Schaeffer, B., Bree, A., Mora, A,. Dahbi, G., Biet, F., Oswald, E., Mainil, J., Blanco, J., Moulin-Schouler, M. (2012). Diagnostic strategy for identifying avian pathogenic *Escherichia*

coli based on four patterns of virulence genes. *J. Clin. Microbiol.* 50(5): 1673-78. doi: 10.1128/JCM.05057-11.

Schulz, J., Ruddat, I., Hartung, J., Hamscher, G., Kemper, N., Ewers, C. (2016). Antimicrobial-resistant *Escherichia coli* survived in dust samples for more than 20 years. *Front. Microbiol.* 10. doi: 10.3389/fmicb.2016.00866.

Serra, D.O., Richter, A.M., Hengge, R. (2013). Cellulose as an architectural element in spatially structured *Escherichia coli* biofilms. *J. Bacteriol.* 195(24). doi: 10.1128/JB.00946-13.

Servin, A.L. (2005). Pathogenesis of Afa/Dr diffusely adhering *Escherichia coli*. *Clin. Microbiol. Rev.* 18(2): 264-92. doi: 10.1128/CMR.18.2.264-292.2005.

Silagyi, K., Kim, S., Lo, Y.M., Wei, C. (2009). Production of biofilm and quorum sensing by *Escherichia coli* O157:H7 and its transfer from contact surfaces to meat, poultry, ready-to-eat deli, and produce products. *Food Microb.* 26(5): 514-519. doi: 10.1016/j.fm.2009.03.004.

Skyberg, J.A., Horne, S.M., Giddings, C.W., Wooley, R.E., Gibbs, P.S., Nolan, L.K. (2003). Characterizing avian *Escherichia coli* isolates with multiplex polymerase chain reaction. *Avian Dis.* 47(4): 1441-1447.

Skyberg, J., Siek, K.E., Doetkoff, C., Nolan, L.K. (2007). Biofilm formation by avian *Eschericha coli* in relation to media, source and phylogeny. *J. Appl. Microbiol.* 102(2): 548-54. doi: 10.1111/j.1365-2672.2006.03076.x

Sorour, H.K., Gaber, A.F., Hosny, R.A. (2020). Evaluation of the efficiency of using *Salmonella* Kentucky and *Escherichia coli* O119 bacteriophages in the treatment and prevention of salmonellosis and colibacillosis in broiler chickens. *Lett. Appl. Microbiol.* 13347. doi: 10.1111/lam.13347.

Subedi, M., Luitel, H., Devkota, B., Bhattarai, R.K., Phuyal, S., Panthi, P., Shreshta, A., Chaudhary, D.K. (2018). Antibiotic resistance pattern and virulence gene content in avian pathogenic *Escherichia coli* (APEC) from broiler chickens in Chitwan, Nepal. *BMC Vet Res.* 14: 113. doi: 10.1186/s12917-018-1442-z.

Sun, Y., Guo, D., Hu, X., Shi, C., Zhang, C., Peng, X., Yang, H., Xia, X. (2019). Disinfectant resistance profiles and biofilm formation capacity of *Escherichia coli* isolated from retail chicken. *Microb. Drug Res.* 25(5). doi: 10.1089/mdr.2018.0175.

Swelum, A.A., El-Saadony, M.T., El-Hack, M.E., Ghanima, M.M.A., Shukry, M., alhotan, R.A., Hussein, E.O.S., *et al.* (2021). Ammonia emissions in poultry houses and microbial nitrification as a promising reduction strategy. *Sci. Total Environ.* 781: 146978. doi: 10.1016/j.scitotenv.2021.146978.

Tivendale, K.A., Allen, J.A., Browning, G.F. (2009). Plasmid-borne virulence-associated genes have a conserved organization in virulent strains of Avian Pathogenic *Escherichia coli*. *J. Clin. Microbiol*. 47(8): 2513-2519. doi: 10.1128/JCM.00391-09.

Tivendale, K.A., Logue, C.M., Kariyawasam, S., Jordan, D., Hussein, A., Li, G., Wannemuehler, Y., Nolan, L.K. (2010). Avian-pathogenic *Escherichia coli* strains are similar to neonatal meningitis *E. coli* strains and are able to cause meningitis in the rat model of human disease. *Infect. Immun*. 78(8): 3412-3419. doi: 10.1128/IAI.00347-10.

Totsika, M., Beatson, S.A., Sarkar, S., Phan, M., Petty, N.K., Bachmann, N., Szubert, M., *et al*. (2011). Insights into a multidrug resistant *Escherichia coli* pathogen of the globally disseminated ST131 lineage: genome analysis and virulence mechanisms. *PLoS ONE* 6(10): e26578. doi: 10.1371/journal.pone.0026578.

Touchon, M., Hoede, C., Tenaillon, O., Barbe, V., Baeriswyl, S., *et al*. (2009). Organized genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. *PLoS Genetics*. 5(1): 31000344. doi: 10.1371/journal.pgen.1000344.

Valiatti, T.B., Santos, F.F., Santos, A.C.M., Nascimento, J.A.S., Silva, R.M., Carvalho, E., Sinigaglia, R., Gomes, T.A.T. (2020). Genetic and virulence characteristics of a hybrid atypical enteropathic and uropathic *Escherichia coli* (aEPEC/UPEC) strain. *Front. Cell. Infect. Microbiol*. doi: 10.3389/fcimb.2020.00492.

Vandekerchove, D., De Herdt, P., Laevens, H., Pasmans, F. (2004). Risk factors associated with colibacillosis outbreaks in caged layer flocks. *Avian Pathol*. 33(3). doi: 10.1080/0307945042000220679.

Van Goor, A., Redweik, G.A.J., Stromberg, Z.R., Treadwell, C.G., Xin, J., Mellata, M. (2020). Microbiome and biological blood marker changes in hens at different laying stages in conventional and cage free housing. *Poultry Sci*. 99(5): 2362-2374. doi: 10.1016/j.psj.2020.01.011.

Vann, W.F., Daines, D.A., Murkin, A.S., Tanner, M.E., Chaffin, D.O., Rubens, C.E., *et al*. (2004). The NeuC protein of *Escherichia coli* K1 is a UDP N-Acetylglucosamine 2-Epimerase. *J. Bacteriol*. 186(3): 706-712. doi: 10.1128/JB.186.3.706-712.2004.

Verma, R., Rojas, T.C.G., Maluta, R.P., Leite, J.L., da Silva, L.P.M., Nakazato, G., da Silveira, W.D. (2016). Fimbria-encoding gene *yadC* has a pleiotropic effect on several biological characteristics and plays a role in avian pathogenic *Escherichia coli* pathogenicity. *Infect. Immun*. 84(1): 187-193. doi: 10.1128/IAI.01138-15.

Vogeleer, P., Tremblay, Y.D.N., Mafu, A.A., Jacques, M., Harel, J. (2014). Life on the outside: role of biofilms in environmental persistence of shiga-toxin producing *Escherichia coli*. *Front. Microbiol*. 5: 317. doi: 10.3389/fmicb.2014.00317.

Wang, S., Niu, C., Shi, Z., Xia, Y., Yagoob, M., Dai, J., Lu, C. (2011). Effects of *ibeA* deletion on virulence and biofilm formation of avian pathogenic *Escherichia coli*. *Infect. and Immun*. 79(1). doi: 10.1128/IAI.00821-10.

Wang, Y., Yi, L., Wang, Y., Wang, Y., Cai, Y., Zhao, W., Ding, C. (2016). Isolation, phylogenetic group, drug resistance, biofilm formation, and adherence genes of *Escherichia coli* from poultry in central China. *Poultry Sci*. 95(12): 2895-2901. doi: 10.3382/ps/pew252.

Wijetunge, D.S.S., Gongati, S., DebRoy, C., Kim, K.S., Couraud, P.O., Romero, I.A., Weksler, B., Kariyawasam, S. (2015). Characterizing the pathotype of neonatal meningitis causing *Escherichia coli* (NMEC). *BMC Microbiol*. 15, 211. doi: 10.1186/s12866-015-0547-9.

Zhang, A., He, X., Meng, Y., Guo, L., Long, M., Yu, H., Li, B., *et al*. (2016). Antibiotic and disinfectant resistance of *Escherichia coli* isolated from retail meats in Sichuan, China. *Microb. Drug Resist*. 22(1): 80-87. doi: 10.1089/mdr.2015.0061.

Zhuge, X., Zhou, Z., Jiang, M., Wang, Z., Sun, Y., Tang, F. *et al*. (2020). Chicken-source *Escherichia coli* within phylogroup F shares virulence genotypes and is closely related to extraintestinal pathogenic *E. coli* causing human infections. *Transbound. Emerg. Dis*. 68(2): 880-895. doi: 10.1111/tbed.13755.

CHAPTER 2. DEVELOPING AND TROUBLESHOOTING A METHODOLOGY FOR WHOLE GENOME SEQUENCING AND PRELIMINARY CHARACTERIZATION OF SYSTEMIC AND CECAL *E. COLI* ISOLATES FROM SASKATCHEWAN BROILER FARMS

2.1 Introduction

The use of whole genome sequencing (WGS) to track and understand bacterial species significantly increased as the cost of sequencing technologies became more affordable (Quainoo *et al.,* 2017). WGS captures all the genetic material within a bacterial cell, including extrachromosomal elements such as plasmids (Leopold *et al.,* 2014). The resolution captured by WGS is an invaluable tool in tracking outbreaks of bacterial disease (Gilchrist *et al.,* 2015). Analysis of sequencing data does not require higher level biosecurity, meaning that fewer workers are at risk from pathogens. Compared to multi-locus sequence typing, which identifies allelic variations in known housekeeping genes to characterize bacteria based on their overall allelic profile, tracking single nucleotide polymorphisms in an outbreak population can provide a much clearer picture of transmission (Maiden *et al.,* 2013).

Next generation sequencing (NGS) technologies used for WGS primarily include sequencers developed by Illumina, Oxford Nanopore, and/or Pacific Biosciences (Goodwin *et al.,* 2016; McCombie *et al.,* 2019). Illumina technology uses a massively parallel sequencing by synthesis approach to determine the sequence of DNA fragments of up to 500bp in length (Goodwin *et al.,* 2016). Nucleotides with attached fluorophores are added sequentially, and the sequence is determined based on the colour of the fluorophore attached to the base. In contrast, Oxford Nanopore and Pacific Biosciences technologies can sequences much longer DNA fragments. Both Pacific Biosciences and Oxford Nanopore instruments can accommodate DNA ranging from under 1kb to the length of an entire bacterial chromosome (Goodwin *et al.,* 2016; McCombie *et al.,* 2019). Pacific Biosciences instruments have a fixed polymerase at the bottom of a sequencing well. Hairpin adapters attached to either end of a DNA fragment create a circular fragment that can be ratcheted through the polymerase multiple times in order to obtain an accurate sequence (Wagner *et al.,* 2016). Bases are determined by the incorporation of fluorophore-labelled

nucleotides (Goodwin *et al.,* 2016). Oxford Nanopore flow cells contain biological protein pores fixed in a membrane across which an electrical current is run. Sequencing adapters attached to DNA fragments ratchet the DNA through the pores 5 nucleotides at a time (Kono and Arakawa, 2019). Each base makes a characteristic disruption in the DNA signal, and these electrical squiggles are converted to basecalls by Nanopore software (Kono and Arakawa, 2019).

Illumina sequencing is one of the most reliable and accurate technologies for sequencing available today (McCombie *et al.,* 2019). Because reads are sequenced in a massively parallel fashion, basecalling accuracy can reach up to 99.99% (Goodwin *et al.,* 2016). However, in the absence of a reference genome, *de novo* assembly using only Illumina reads will produce a fragmented genome. This is because Illumina reads are often significantly shorter than repeat regions found within bacterial genomes, and *de novo* assemblers cannot reconcile those repeats (Sonh and Nam, 2016). Pacific Biosciences and Oxford Nanopore technologies can overcome these assembly problems by producing significantly longer reads, however they suffer from a lower accuracy (Sonh and Nam, 2018; McCombie *et al.,* 2019). Combining short- and long-read technologies together can produce highly accurate and complete genomes (Utturkar *et al.,* 2014).

We developed a method to isolate *E. coli* from avian tissue samples, extract genomic DNA, and sequence those isolates using Illumina and Oxford Nanopore technologies. We use a spin-column based DNA extraction, which are more likely to fragment DNA than other methods such as DNA precipitation but produces DNA fragments up to 60kb in length (Quick and Lohman, 2019). We determined that fragments produced using our genomic extraction kit are sufficiently long enough to assemble a circular draft chromosome in our *E. coli* isolates. We also identified some preliminary differences in *E. coli* isolated from broilers with confirmed colibacillosis (systemic *E. coli*) and isolates recovered from the cecal contents of healthy birds (cecal *E. coli*). Ultimately, this method will be used to characterize and track outbreaks of colibacillosis on Saskatchwan broiler farms to help producers manage disease.

2. 2 Isolating and Characterizing Avian *E. coli*

2.2.1 Isolating *E. coli* from Diseased and Healthy Broilers

Broiler carcasses were collected for this study in collaboration with Dr. Tyra Dickson, Poultry Veterinarian, Tennille Knezacek, Professional Research Associate, and Nikki Storbakken, Animal Health Technician, of the Poultry Extension Team (PEX) at the Western College of Veterinary Medicine, University of Saskatchewan. We took advantage of an existing system in which producers submit diseased broilers to PEX for necropsy. After a preliminary diagnosis of colibacillosis is made, typically based on the presence of internal lesions on the organs, portions of the tissues are sent to PDS for bacterial culturing (Prairie Diagnostic Services; http://pdsinc.ca/). PDS confirms the colibacillosis diagnosis via matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry. MALDI-TOF MS separates ionized particles from an unknown bacterial sample according to a mass-to-charge ratio. A measurement of the particle ratios and the flight time of these particles from one end of a tube to a detector at the other end of the tube determines the spectra of the sample, and this can be compared to the spectra produced by samples of a known origin (Wunschel *et al*., 2005). MALDI-TOF MS is frequently used to identify bacteria down to the species level and is comparable to sequencing in accuracy (Chui *et al*., 2015; Sauget *et al*., 2017).

For our study, PEX stored a portion of the heart, liver, and spleen of diseased broilers at 4°C until confirmation of *E. coli* diagnosis was received from PDS. Equal sized pieces of each organ were placed in a 2 mL SafeLock Eppendorf tube with 1mL of phosphate buffered saline (PBS) and a 5-mm steel bead (Qiagen #69989) and homogenized for 5 min at 30 Hz using a mixer mill (Retsch; MM400). 10 µL aliquots of homogenized samples were added to 90 uL of PBS and subsequently serially diluted in PBS; dilutions of 10^{-3}, 10^{-4}, and 10^{-5} were plated on MacConkey Agar. Colonies were grown for 18-24 hours at 37°C. Upon visual inspection, 3-6 suspected *E. coli* colonies were streaked onto fresh MacConkey agar and incubated for an additional 18 hours at 37°C. Methods for biochemical confirmation of *E. coli* isolates and determining biofilm formation can be found in appendix A.

2.2.2 Genomic DNA Extraction

Genomic DNA was extracted from each *E. coli* isolate using the GenElute Bacterial Genomic DNA Extraction Kit (Sigma), which produces high quality genomic DNA fragments up to 60 kb in length (Lohman *et al*., 2016). Isolates were streaked on LB agar from a 50% glycerol freezer stock and grown at 37°C for 18 hours. A single colony was inoculated into 5 mL of LB

broth and grown at 37°C with shaking for 18 hours. DNA was isolated according to the kit manufacturers protocol. Purity of the DNA was assessed using a NanoDrop spectrophotometer (Thermo Fisher) and quantified using a Qubit Fluorometer (Thermo Fisher) (Desjardins and Conklin, 2010; Mardis and McCombie, 2017).

2.2.3 Nanopore Sequencing

All *E. coli* isolates were sequenced on a Nanopore MinION according to the protocol developed by Nick Lohman, Matt Loose, and Mick Watson for Porecamp Vancouver 2018 (http://porecamp.github.io). Because the genomic DNA extraction step uses spin column purification, we expect that the genomic DNA will shear when it is physically forced through the column during centrifugation. Small fragments produced during DNA extraction (less than 500bp in length) must be removed prior to library preparation because the amount of DNA ligase added to each step of the Porecamp protocol is calculated based on the expected number of DNA molecules available for sequencing (Lohman *et al*., 2018). An equal mass of high molecular weight DNA will have fewer "ends" available for ligation than a sample with many short reads (**Fig 2.1**). Failure to remove these short reads can dramatically affect the yield of a Nanopore run, because only a portion of the total DNA fragments in a sample will be sequenced and shorter reads yield less data than large reads (Lopez *et al*., 2019). The Porecamp protocol is optimized for reads greater than 8kb (http://porecamp.github.io).

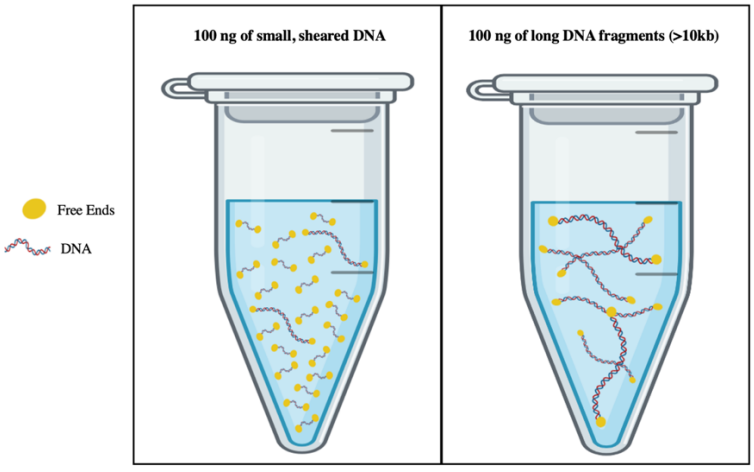

Figure 2.1. The fragment length of a Nanopore sequencing library contributes to the overall quality of a sequencing run. Two samples with identical masses of DNA but different fragment lengths will not have the same number of ends available for ligation (yellow circles).

To remove small fragments from the genomic DNA eluate, we added 0.4x magnetic beads to each sample (NucleoMag NGS beads, Macherey-Nagel) according to the protocol described in the kit (Nagar and Schwessinger, 2018; Maghini *et al*., 2020). DNA fragments larger than 500 bp were saved and washed in fresh 80% ethanol and eluted in 1 mM Tris buffer, pH 8.0. DNA was stored at -20°C for up to one week before sequencing. Any DNA not used for Nanopore sequencing was stored at -80°F for future Illumina sequencing.

Up to 200 fmol of purified and size-selected DNA (in a maximum of 24 µL volume) was added to a lo-bind Eppendorf tube. dH_20 was added to 30 µL. Nicks and gaps in the DNA were repaired using FFPE DNA Repair Mix (NEBNext #M6630L) and polyA-tails were added to DNA fragments using the Ultra II End Repair/dA-Tailing Module (NEBNext #E7645). Nanopore barcodes (NBD-104) were ligated to DNA fragments using Blunt/TA Master Mix (NEB). Multiplexed samples were pooled into a single Eppendorf tube and the reaction was adjusted to 1 M NaCl. DNA repair enzymes and excess DNA barcodes were removed using a 0.2X magnetic bead clean-up. Beads were washed twice with 80% EtOH and the DNA was eluted in 10 mM Tris, pH 8.0 according to the protocol. Sequencing adapters (AMII; Oxford Nanopore Technologies EXP-NBD104) were ligated to the ends of DNA fragments using Quick T4 DNA ligase (NEB) before a final bead clean-up (1:1 ratio of beads to sample) was performed to remove T4 DNA ligase and excess sequencing adapters from the final purified DNA library. Beads were washed by incubating two times in long chain fragment buffer (LFB) and resuspending by flicking. DNA was eluted in 1mM Tris Buffer, pH 8.0. The final library was prepared and loaded onto the sequencer as described in the SQK-LSK109 protocol. Sequencing was performed using R.9.4.1 MinION flow cells. Sequencing runs were terminated after approximately 24 hours.

2.2.4 Illumina Sequencing

96 strains previously sequenced on the MinION were selected for Illumina Sequencing. Purified and size selected DNA previously used for MinION sequencing was used for Illumina library preparation. Libraries were prepared according to the NEBNext Ultra II DNA Library Prep Kit protocol (NEB #E7103). Approximately 100 ng of genomic DNA was sheared by incubating the DNA with the NEBNext Ultra II FS Enzyme Mix (NEB #E7805L) in a

thermocycler with the following conditions: 12 minutes at 37°C, 30 minutes at 65°C, hold at 4°C. NEBNext Adaptors for Illumina were ligated to the fragmented DNA using the NEBNext Ultra II Ligation Mix (NEB #E7595L) and incubated at 20°C for 15 minutes. Next, 25 μL of magnetic beads (NucleoMag NGS beads, Macherey-Nagel #MN-744970.50) was added to the DNA to purify the DNA and remove excess enzyme, followed by a size selection step with 10uL of magnetic beads. Finally, DNA was enriched via PCR using NEBNext Multiplex Oligos for Illumina (NEB #E6609) using the following thermocycler conditions: denature for 30 seconds at 98°C, 4 cycles of denaturing at 98°C for 10 seconds followed by annealing and extension at 65°C for 75 seconds, and a final extension at 65°C for five minutes. Sequencing was performed by Novogene Corporation Inc. (Sacramento, CA) using an Illumina HiSeq 400. Samples were distributed equally across 3 lanes (32 samples/lane). Illumina-specific sequencing adapters were removed by Novogene after sequencing. The quality of the reads was checked via FastQC, and barcodes were removed using Trimmomatic (Andrews, 2010; Bolger *et al*., 2014). A summary of the sequencing data quality and reads generated from this run can be found in appendix B (**table B1**).

2.2.5 Trimming and Filtering of Nanopore Reads

A visual schematic outlining data clean-up and assembly can be found in **Figure 2.2**. Nanopore reads were base-called and debarcoded using Oxford Nanopore Technologies software programs MinKNOW (version 2.0) and Guppy (version 2.1.1). Read quality and lengths for each sequencing run were determined and visualized using the NanoPack software (version 1.35.4) developed by De Coster *et al*. (**Fig 2.3A and B**) (2018). Longer read lengths are beneficial for *de novo* sequence assembly, and the absence of a large number of small fragments indicates a higher quality sequencing run. Additional read trimming was performed using Porechop to remove any barcodes missed by Guppy (Wick, 2018a). Reads were filtered for quality using Nanofilt to ensure that only reads with a Q score of 10 or greater were included in draft assemblies (-q 10) (De Coster *et al*., 2018). During basecalling, a score is assigned to each base which gives the confidence that that basecall is correct (Smith *et al*., 2008). A higher quality score indicates a lower chance of basecalling errors and a more accurate sequence. Sequences with a quality score of 10 are predicted to be 90% accurate (Smith *et al*., 2008). Unlike Illumina sequencing, which can generate reads with Q scores over 20 (99% accuracy), basecalling via Nanopore sequencing relies on a program's

ability to correctly determine the raw electrical signal of three to five bases within a pore (Rang *et al*., 2018). This means that a basecalling program must be able to correctly distinguish the set of five bases from over 2,000 possible combinations of bases (Rang *et al*., 2018). In addition, homopolymers, or repeat stretches of a single base, are difficult for Nanopore technologies to distinguish.

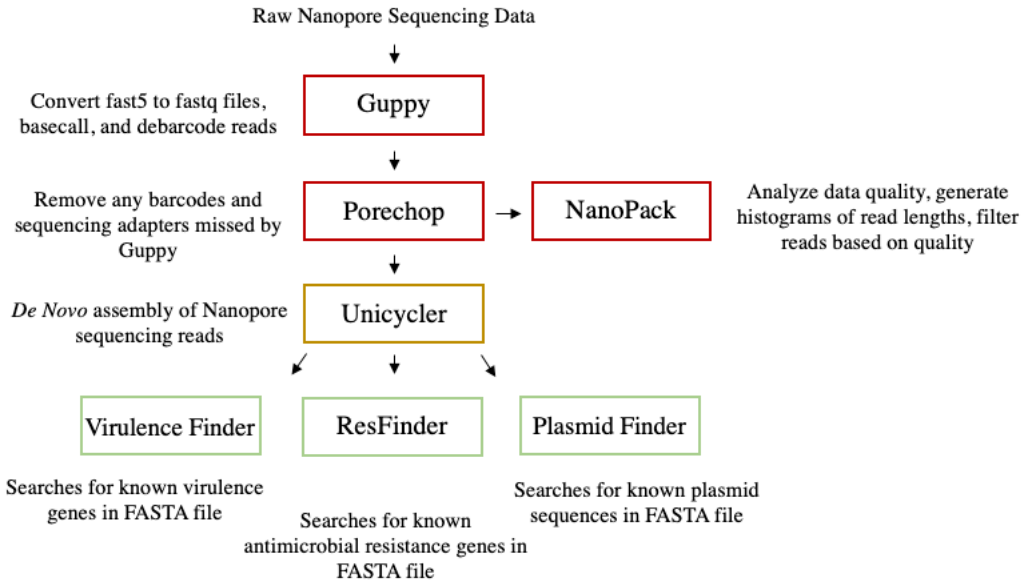

Figure 2.2. Schematic of genome assembly and sequencing data analysis pipeline. All programs were run in an Anaconda environment (Anaconda Software Distribution, 2016).

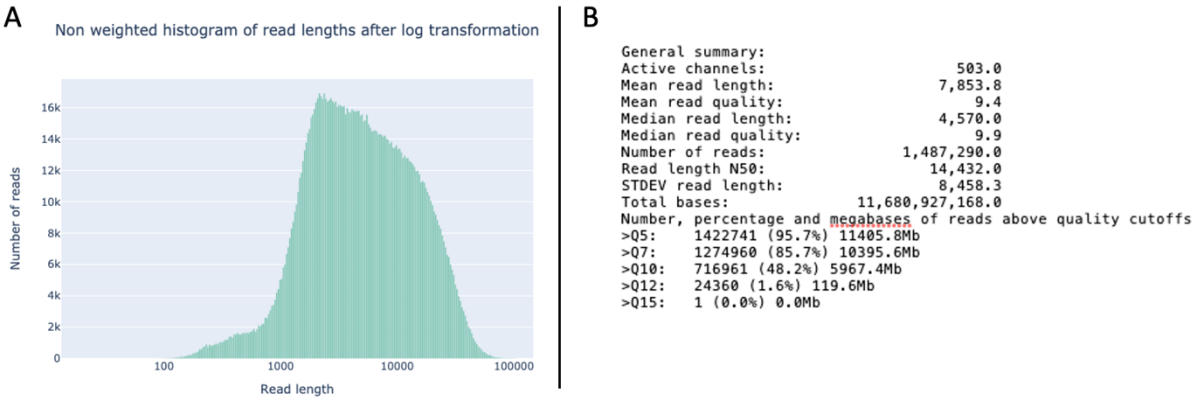

Figure 2.3. A. Example output from NanoPack software showing an unweighted histogram of read lengths producing during a Nanopore sequencing run from September 24, 2019, and **B.** a summary report of the same sequencing run generated using NanoStat (De Coster *et al*., 2018).

A high quality Nanopore sequencing run should produce a many reads of a sufficient mean read length to be able to assemble a single, circular chromosomal contig after assembly. While this varies depending on the isolate, in general, an average read length of around 7kb is generally sufficient for *de novo* assembly (Sohn and Nam, 2018).

2.3 Preliminary WGS Results and Discussion

2.3.1 WGS Data Summary and Quality

A summary of the sequencing data quality and reads generated from the Illumina sequencing run can be found in appendix B (**Table B1**). These results demonstrate that the DNA extraction protocol and library preparation were successful. Illumina data will be used in future hybrid genome assemblies but was not further analysed in this thesis.

As of October 2021, a total of 496 *E. coli* have been collected from 16 Saskatchewan colibacillosis outbreaks. Systemic isolates have been recovered from 16 outbreaks and cecal isolates from healthy birds of the same flocks have been recovered from 14/16 outbreaks. A set of 96 strains were selected for both Nanopore and Illumina sequencing based on outbreak, disease status, cellulose and curli production, and colony morphology. All 96 have been sequenced on the Nanopore MinION. 10/96 strains failed Illumina quality control and were not sequenced. These will be repeated at a later date. All 96 strains have been tested for their phenotypic resistance to 27 different antibiotics (discussed in detail in chapter 3). These strains are summarized in **Table 2.1**.

Table 2.1. Summary of total *E. coli* isolated, sequenced, and AMR profiled as of October 2021.

Source	Total Isolates	Total Nanopore Sequenced Isolates	Total Nanopore + Illumina Sequenced Isolates	Total Isolates with AMR profiles
Systemic	394	63	52	58
Cecal	102	40	34	38

The average read quality score from Nanopore sequencing runs performed between July of 2019 and May of 2021 ranged from 7.6-12.8 (**Table 2.2**). As previously discussed, Nanopore sequencing technology produces reads with a lower quality score than Illumina sequencing (McCombie *et al*., 2019). Quality scores reflect the confidence in a given basecall by the Guppy basecalling software. Reads in highly repetitive regions of the genome will have lower quality scores because the basecalling software may struggle to determine the exact sequence of homopolymers (Purnell *et al*., 2008). We can perform additional manual filtering of the reads to remove any that are below a quality of 10 (10% error rate). Many of these errors can also be corrected during assembly if the amount of coverage is sufficient. *De novo* assemblers overlap sequencing reads and determine the base at each position based on consensus from all of the reads overlapping at that position. The number of reads and mean read length are critical for genome assembly. Enough reads to provide high coverage at every position of the genome will help mitigate Nanopore sequencing errors. Low coverage or read lengths shorter than repeat regions in the genome will result in fragmented assemblies.

Table 2.2. Summary of Nanopore MinION sequencing runs of genomic DNA from 13 barcoded *E. coli* samples performed between August 2019 and May 2021.

Date	Number of Reads	Mean Read Quality	Mean Read Length
8 July 2019	931,911	9.7	7,277.3
24 Sept 2019	514,107	11.2	9,829.6
11 Dec 2019	638,826	10.5	6,804.8
24 Jan 2020	1,487,290	9.4	7,853.8
4 Feb 2020	816,597	12.6	6,118.9
12 Feb 2020	1,975,468	8.8	6,295.3
24 Feb 2020	2,112,666	9.2	5,807.6
26 Feb 2020	2,918,972	7.6	3,003.6
30 Oct 2020	1,416,031	12.8	8,223.7
31 Oct 2020	1,814,449	9.4	10,971.4
16 Nov 2020	1,520,814	12.2	8,680.3
27 Jan 2021	2,030,102	8.8	9,369.1
14 May 2021	1,608,835	9.4	6,443.8

To troubleshoot our method for extracting DNA and sequencing avian *E. coli* genomes, we assembled rough draft genome assemblies for 96 isolates using Unicycler (version 0.4.8)

(Wick *et al*., 2017). Due to the diversity of *E. coli*, we decided to use a *de novo* assembly approach rather than a reference-based approach because it was unlikely that we would find a good reference strain for all of our *E. coli* isolates. *De novo* assembly provides a more accurate assembly for *E. coli* genomes compared to reference-based assembly due to the variability of the *E. coli* genome but may struggle to reconcile long repeats within the genome (Liao *et al*., 2018). Unicycler assembles long reads using a fast approach in which long reads that share parts of the same sequence are overlapped, then these pieces are arranged into contigs. Unicycler is faster than other programs because it does not have a consensus or polishing step after assembly, which means that the error rate of the assembly is the same as that of the sequencing reads (Wick *et al*., 2017). A summary of the characteristics of the 96 *E. coli* strains selected for draft genome assembly, as well as sequencing information and location on the lab portable hard drive, can be found in **Table 2.3**. In these 96 *E. coli* isolates, a mean read length of over 6kb was sufficient to assemble a closed, circular draft *E. coli* chromosome using Unicycler. With careful preparation, our *E. coli* sequencing methodology produces a sufficient quality and quantity of reads to assemble *E. coli* genomes.

Table 2.3. Summary of information about each *E. coli* strain sequenced on the Nanopore MinION between July 2019 and May 2021, as well as the name of the folder containing all the reads on the lab's portable hard drive.

Sample Name	Date Collected	Systemic/Cecal	Organ	Outbreak	Curli	Cellulose	# Reads	Location on HD
3862-H1	July 2019	Systemic	Heart	2	+	+	123,054	31 Oct 20
4957-3S1	Feb 2020	Systemic	Spleen	11	+	-	332,000	30 Oct 20
23315-H1	July 2019	Systemic	Heart	1	+	-	16,000	8 July 19
4957-C1	Feb 2020	Cecal	Cecum	11	+	-	124,000	30 Oct 20
6041-C2	Nov 2019	Cecal	Cecum	8	+	+	60,000	4 Feb 20
23315-C4	July 2019	Cecal	Cecum	1	+	+	140,000	31 Oct 20
6245-2L1	Nov 2019	Systemic	Liver	6	+	+	48,000	11 Dec 19
6245-1S2	Nov 2019	Systemic	Spleen	6	+	+	16,000	11 Dec 19
6245-1H1	Nov 2019	Systemic	Heart	6	+	+	44,000	11 Dec 19
6245-3L1	Nov 2019	Systemic	Liver	6	+	+	40,000	11 Dec 19
9226-1L2	Nov 2019	Systemic	Liver	5	+	+	208,000	12 Feb 20
9226-1S1	Nov 2019	Systemic	Spleen	5	+	+	76,000	12 Feb 20
9226-2S1	Nov 2019	Systemic	Spleen	5	+	+	64,000	12 Feb 20
9226-2L1	Nov 2019	Systemic	Liver	5	+	+	104,000	12 Feb 20
9226-3H1	Nov 2019	Systemic	Heart	5	+	+	52,000	12 Feb 20
9226-3L1	Nov 2019	Systemic	Liver	5	+	+	64,000	12 Feb 20
6245-C4	Nov 2019	Cecal	Cecum	6	+	+	60,000	11 Dec 19
6245-3H2	Nov 2019	Systemic	Heart	6	+	-	36,000	11 Dec 19
0012-1H1	Aug 2019	Systemic	Heart	4	+	-	104,000	24 Feb 20
0012-2L1	Aug 2019	Systemic	Liver	4	-	+	192,000	24 Feb 20
0012-3S1	Aug 2019	Systemic	Spleen	4	+	-	116,000	24 Feb 20
0012-C1	Aug 2019	Cecal	Cecum	4	+	-	192,000	24 Feb 20
0012-C5	Aug 2019	Cecal	Cecum	4	+	-	168,000	24 Feb 20
0012-C7	Aug 2019	Cecal	Cecum	4	+	+	220,000	24 Feb 20
9619-1H2	Aug 2019	Systemic	Heart	3	+	+	68,000	30 Oct 20
9619-2L1	Aug 2019	Systemic	Liver	3	+	+	120,000	24 Feb 20

9619-3S1	Aug 2019	Systemic	Spleen	3	+	+	152,000	24 Feb 20
9619-C1	Aug 2019	Cecal	Cecum	3	+	-	156,000	24 Feb 20
9619-C6	Aug 2019	Cecal	Cecum	3	+	+	136,000	24 Feb 20
9619-C8	Aug 2019	Cecal	Cecum	3	+	+	72,000	24 Feb 20
0205-1L3	Jan 2020	Systemic	Liver	9	-	+	84,000	26 Feb 20
0205-2L2	Jan 2020	Systemic	Liver	9	-	+	220,000	26 Feb 20
23315-C8	July 2019	Cecal	Cecum	1	-	+	516,000	24 Sep 19
23315-C5	July 2019	Cecal	Cecum	1	+	+	12,000	8 Jul 19
23315-C9	July 2019	Cecal	Cecum	1	-	-	88,000	14 May 21
23315-L3	July 2019	Systemic	Liver	1	+	-	80,000	8 Jul 19
3862-C9	July 2019	Cecal	Cecum	2	+	-	64,000	16 Nov 20
3862-S2	July 2019	Systemic	Spleen	2	+	+	516,000	24 Sept 19
3862-C11	July 2019	Cecal	Cecum	2	-	-	32,000	16 Nov 20
3862-C8	July 2019	Cecal	Cecum	2	+	-	112,000	14 May 21
3826-S1	July 2019	Systemic	Spleen	2	+	+	80,000	31 Oct 20
3862-L1	July 2019	Systemic	Liver	2	+	+	176,000	31 Oct 20
3862-C2	July 2019	Cecal	Cecum	2	-	-	480,000	24 Sep 19
3862-C6	July 2019	Cecal	Cecum	2	+	+	356,000	24 Sep 19
9413-2H1	Oct 2019	Systemic	Heart	7	+	+	132,000	4 Feb 20
9413-C4	Oct 2019	Cecal	Cecum	7	+	+	84,000	31 Oct 20
9413-C1	Oct 2019	Cecal	Cecum	7	-	-	112,000	16 Nov 20
9413-3S1	Oct 2019	Systemic	Spleen	7	+	+	152,000	31 Oct 20
9413-1H1	Oct 2019	Systemic	Heart	7	+	+	132,000	14 May 21
9413-2S2	Oct 2019	Systemic	Spleen	7	+	+	60,000	16 Nov 20
9413-1S2	Oct 2019	Systemic	Spleen	7	+	+	68,000	4 Feb 20
6041-1H1	Nov 2019	Systemic	Heart	8	+	+	88,000	4 Feb 20
6041-3S1	Nov 2019	Systemic	Spleen	8	+	-	92,000	4 Feb 20
6041-C6	Nov 2019	Cecal	Cecum	8	+	-	88,000	4 Feb 20
6041-C9	Nov 2019	Cecal	Cecum	8	+	+	84,000	4 Feb 20
6041-2L1	Nov 2019	Systemic	Liver	8	+	+	124,000	4 Feb 20
6245-1L1	Nov 2019	Systemic	Liver	6	+	+	240,000	16 Nov 20
6245-C6	Nov 2019	Cecal	Cecum	6	+	+	32,000	11 Dec 19
6245-2H3	Nov 2019	Systemic	Heart	6	+	+	68,000	11 Dec 19

6245-C1	Nov 2019	Cecal	Cecum	6	+	+	48,000	11 Dec 19
6245-C2	Nov 2019	Cecal	Cecum	6	+	+	132,000	24 Jan 20
6245-3S1	Nov 2019	Systemic	Spleen	6	+	+	100,000	11 Dec 19
4957-C5	Feb 2020	Cecal	Cecum	11	+	+	96,000	30 Oct 20
4957-2L1	Feb 2020	Systemic	Liver	11	+	+	104,000	30 Oct 20
4957-C3	Feb 2020	Cecal	Cecum	11	+	+	84,000	30 Oct 20
4957-1H1	Feb 2020	Systemic	Heart	11	+	+	85,000	30 Oct 20
4957-2L3	Feb 2020	Systemic	Liver	11	-	+	64,000	30 Oct 20
4957-1H2	Feb 2020	Systemic	Heart	11	-	+	124,000	14 May 21
4957-3S3	Feb 2020	Systemic	Spleen	11	+	+	132,000	16 Nov 20
4957-C6	Feb 2020	Cecal	Cecum	11	+	+	164,000	31 Oct 20
0205-C9	Jan 2020	Cecal	Cecum	9	+	+	192,000	26 Feb 20
0205-3H1	Jan 2020	Systemic	Heart	9	+	+	164,000	31 Oct 20
0205-C3	Jan 2020	Cecal	Cecum	9	+	-	232,000	26 Feb 20
0205-C7	Jan 2020	Cecal	Cecum	9	-	-	184,000	26 Feb 20
0205-3S1	Jan 2020	Systemic	Spleen	9	+	+	104,000	31 Oct 20
2402-1H1	Feb 2020	Systemic	Heart	10	+	+	176,000	26 Feb 20
2402-2L1	Feb 2020	Systemic	Liver	10	+	+	180,000	26 Feb 20
2402-3S1	Feb 2020	Systemic	Spleen	10	+	+	84,000	26 Feb 20
2402-C2	Feb 2020	Cecal	Cecum	10	-	-	88,000	26 Feb 20
2402-C3	Feb 2020	Cecal	Cecum	10	+	-	92,000	26 Feb 20
2402-C4	Feb 2020	Cecal	Cecum	10	+	+	104,000	14 May 21
7578-2H3	Feb 2020	Systemic	Heart	12	+	+	96,000	31 Oct 20
7578-1L2	Feb 2020	Systemic	Liver	12	+	+	284,000	16 Nov 20
7578-2L2	Feb 2020	Systemic	Liver	12	+	+	116,000	14 May 21
7578-2H2	Feb 2020	Systemic	Heart	12	+	+	20,000	16 Nov 20
7578-1H1	Feb 2020	Systemic	Heart	12	+	+	104,000	14 May 21
23315-S2	July 2019	Systemic	Spleen	1	+	-	128,000	14 May 21
9619-3L1	Aug 2019	Systemic	Liver	3	+	-	88,000	30 Oct 20
9619-2S1	Aug 2019	Systemic	Spleen	3	+	+	96,000	14 May 21
9226-C1	Nov 2019	Cecal	Cecum	5	-	+	140,000	24 Jan 20
9226-C2	Nov 2019	Cecal	Cecum	5	-	+	188,000	24 Jan 20
9226-C5	Nov 2019	Cecal	Cecum	5	-	+	96,000	24 Jan 20

9226-2H2	Nov 2019	Systemic	Heart	5	+	+	52,000	16 Nov 20
6041-1L2	Nov 2019	Systemic	Liver	8	+	+	192,000	14 May 21
6041-3L1	Nov 2019	Systemic	Liver	8	+	-	28,000	30 Oct 20
6245-C5	Nov 2019	Cecal	Cecum	6	+	+	68,000	24 Jan 20

2.3.2 Draft Long Read Assemblies Suggest Systemic and Cecal Isolates may be Genetically Different

After assembling the long-read sequences from our first colibacillosis outbreak, we used the Blast Ring Image Generator to compare the similarity of multiple draft *E. coli* sequences from this outbreak (Alikhan *et al.,* 2011). We compared *E. coli* isolates recovered from the organs of broilers with confirmed colibacillosis (shown in purple) to *E. coli* isolates from healthy broilers of the same flock (shown in yellow). Differences in the genome sequences are visualized as gaps. We noticed that the systemic isolates (purple) appeared to be highly similar, while there were regions of the cecal isolates that appeared to be quite different (**Fig 2.4**). Systemic isolates from this farm all belong to phylogroup C, while cecal isolates from this farm belonged to phylogroup A. This is discussed in more detail in section 2.3.3. While this was a very rough preliminary analysis, we decided to look further into the genomes to see if we could identify differences between the systemic and cecal *E. coli* populations.

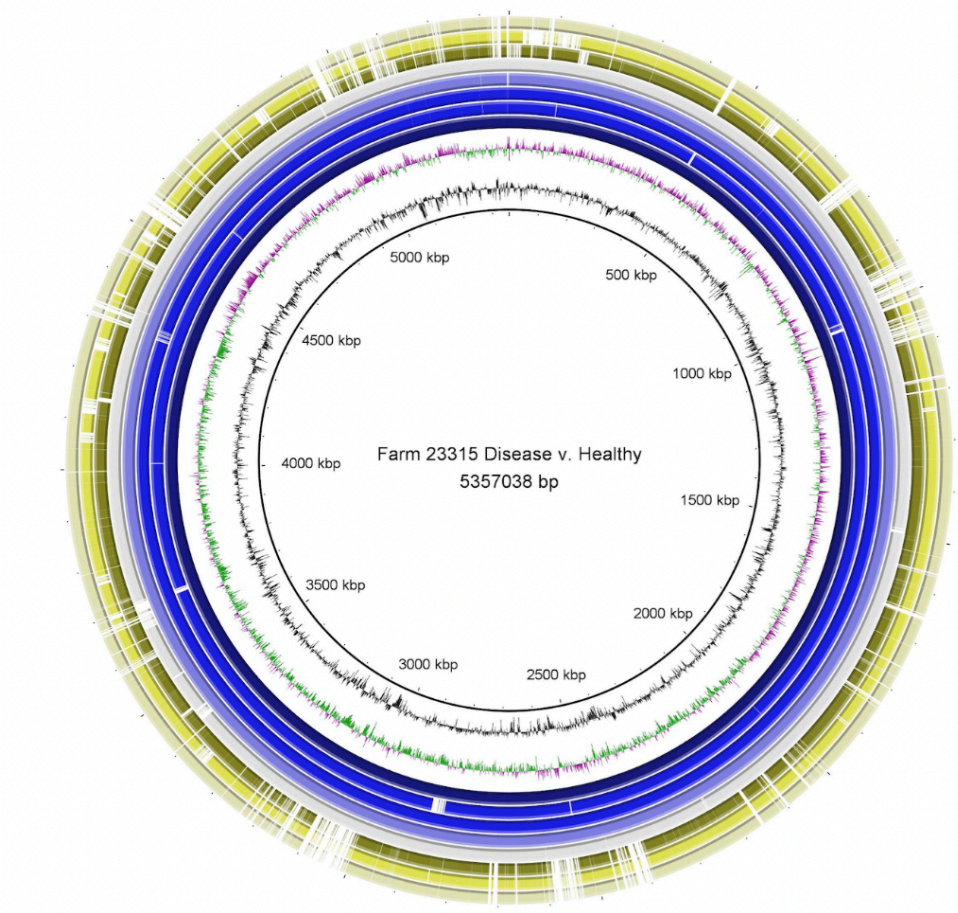

Figure 2.4. Comparison of the draft genome assemblies of the first eight *E. coli* isolates sequenced from a single colibacillosis outbreak (farm 23315). Isolates from diseased birds (purple) belong to phylogroup C, while isolates from healthy birds (yellow) belong to phylogroup A. Differences in the DNA sequences are represented as gaps. Image created using BRIG (Alikhan et al., 2011). The purple/green inner ring shows the GC skew of the genomes. The black ring represents the GC content at each position in the assembly.

2.3.3 Phylogeny, Plasmid Content, AMR, and Virulence Genes of Systemic and Cecal Isolates

We determined the phylogroup of each *E. coli* strain using the Clermontyping tool (http://clermontyping.iame-research.center/), a web-based platform that predicts phylogroups based on an *in silico* quadruplex PCR to screen the whole genome FASTA sequence of an isolate for the *arpA*, *chuA, yjaA* and TspE4.C2 genes, as described above in Section 1.2.1 (Beghain *et*

al., 2018). We found that the majority of Saskatchewan cecal strains fall into phylogroups A (60.5%) and B1 (15.8%), which contain commensal and environmental *E. coli* isolates, respectively (**Table 2.4**) (Duriez *et al*., 2001). In contrast, systemic isolates fell primarily into phylogroups G (32.8%), A (20.7%), and D (18.9%). Phylogroups G and D contains many pathogenic *E. coli* strains, including other groups of extraintestinal pathogenic *E. coli* (Clermont *et al*., 2000; Clermont *et al*., 2019). The wide genetic diversity identified in the pool of systemic isolates suggests that transmissible elements, such as plasmids, may play a role in infection (Sola-Ginez *et al*., 2015).

Table 2.4. A breakdown of *E. coli* phylogroups by sample origin. Systemic *E. coli* belong primarily to phylogroup G, a virulent and drug-resistant group, while cecal isolates primarily fall into phylogroup A, which commonly contains commensal and environmental isolates.

	A	**B1**	**B2**	**C**	**D**	**E**	**F**	**G**
Systemic	12	3	4	4	11	3	2	19
Cecal	23	6	0	2	3	3	0	1

Because plasmids are thought to play a role in APEC pathogenesis and survival, we screened *E. coli* draft assemblies for plasmid sequences using PlasFlow, a program that uses a neural network to differentiate plasmid from chromosome sequences and has a higher accuracy than BLAST-based programs such as Plasmid Finder (Krawczyk *et al.,* 2018). We found that systemic isolates possessed an average of 4 plasmids (range from 0-9), while cecal isolates possess an average of 1 plasmid (range from 0-4) (**Fig 2.5**).

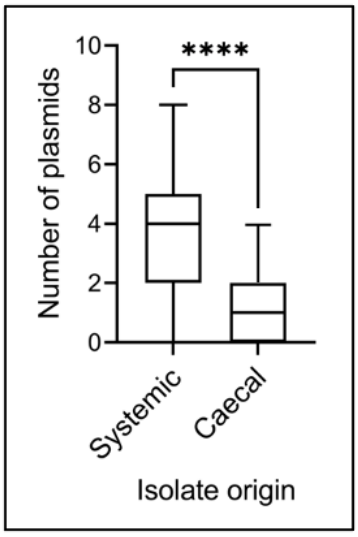

Figure 2.5. Plot of average plasmid number per cell for 96 *E. coli* isolates. Systemic *E. coli* were isolated from broilers with confirmed colibacillosis and cecal *E. coli* were isolated from the cecal contents of healthy broilers from the same flocks (****P<.0001). Significance determined using Mann Whitney U Test.

Next, we screened draft chromosome and plasmid sequences for virulence and antimicrobial resistance genes to see if we could identify genes unique to either group of isolates. We used a publicly available set of tools from the Centre for Disease Epidemiology: Virulence Finder and ResFinder (Carattoli *et al*., 2014; Kleinheinz *et al*., 2014). We used ResFinder to search for acquired AMR genes in each draft genome because it is a regularly curated and updated database containing many genes found in *E. coli* (Kleinheinz *et al*., 2014). Because of the high error rates of Nanopore sequencing, and a lack of any polishing or consensus step in the Unicycler assembly program, we did not search for AMR acquired via point mutations. This will be performed later using genomes assembled with both Illumina and Nanopore reads. We used Virulence Finder to screen draft assemblies for virulence genes for the same reason: this database is consistently updated and curated (Carattoli *et al*., 2014).

We identified few chromosomal antimicrobial resistance genes in our *E. coli* population, with the exception of *mdfA*, an outer membrane transporter that confers resistance to quaternary ammonium compounds, the active ingredient in a common commercial barn disinfectant (**Fig 2.6A and B**). A small number of systemic isolates possess chromosomal *TetA* genes, and four

cecal isolates possess chromosomal aminoglycoside or sulfonamide genes (**Fig 2.6A**). The remaining AMR genes identified in both groups were present on plasmids (**Fig 2.6C and D**). Over 25% of systemic isolates possessed plasmid-encoded tetracycline (*tet(A)*), beta-lactam (*blaTEM-1B*), and aminoglycoside (*aac(3)*-2d, *aph(3")*) resistance genes. Plasmid-mediated aminoglycoside resistance (*aph(6)-Id, aac(3)-IV)* was the most common type found in cecal isolates, at 9% abundance (**Fig 2.6D**). Almost half (41%) of cecal isolates lacked any plasmid-encoded AMR genes, compared to only 10% of systemic isolates.

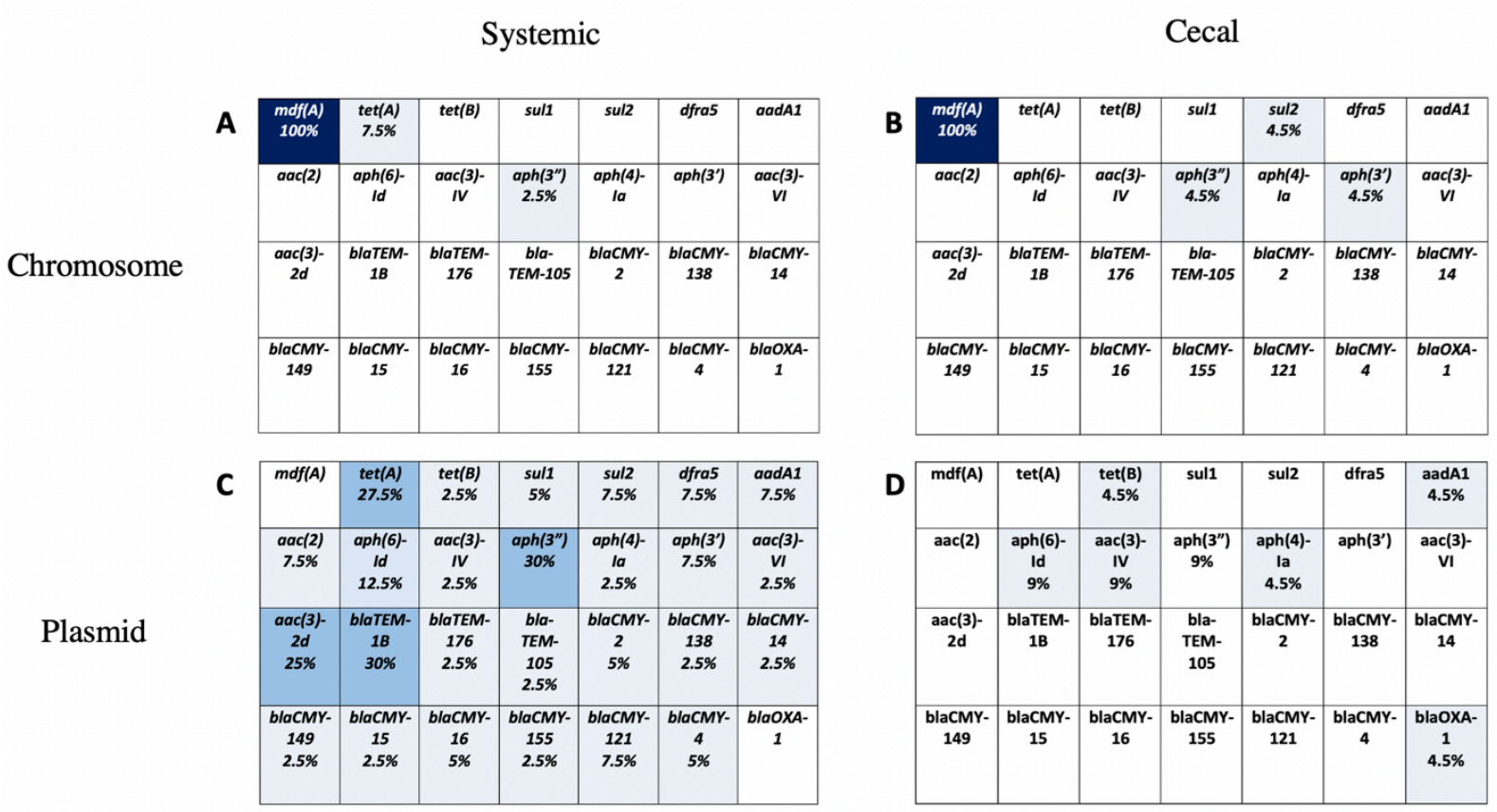

Figure 2.6. A. Percent abundance of AMR genes on the chromosome of systemic *E. coli* where light blue = >10%, medium blue = 10-30%, and dark blue = >30%. **B**. Percent abundance of AMR genes on the chromosome of cecal *E. coli*. **C**. Percent abundance of AMR genes on the plasmids of systemic *E. coli*, and **D**. Percent abundance of AMR genes on the plasmids of cecal *E. coli*.

We screened chromosome and plasmid sequences for virulence gene content and found that every plasmid containing an AMR gene also contained one or more virulence genes. Over 50% of systemic isolates possessed plasmid-mediated genes encoding increased serum survival (*iss*), a metal and manganese transporter implicated in hydrogen peroxide resistance (*sitA*), and the toxin hemolysin F (*hlyF*) **(Fig 2.7**). Only the plasmid-mediated complement resistance gene *traT* was common among both systemic and cecal isolates (**Fig 2.7C and D**). Cecal isolates appear to possess fewer virulence genes than systemic isolates, though we will need to sequence additional isolates in order to confirm this.

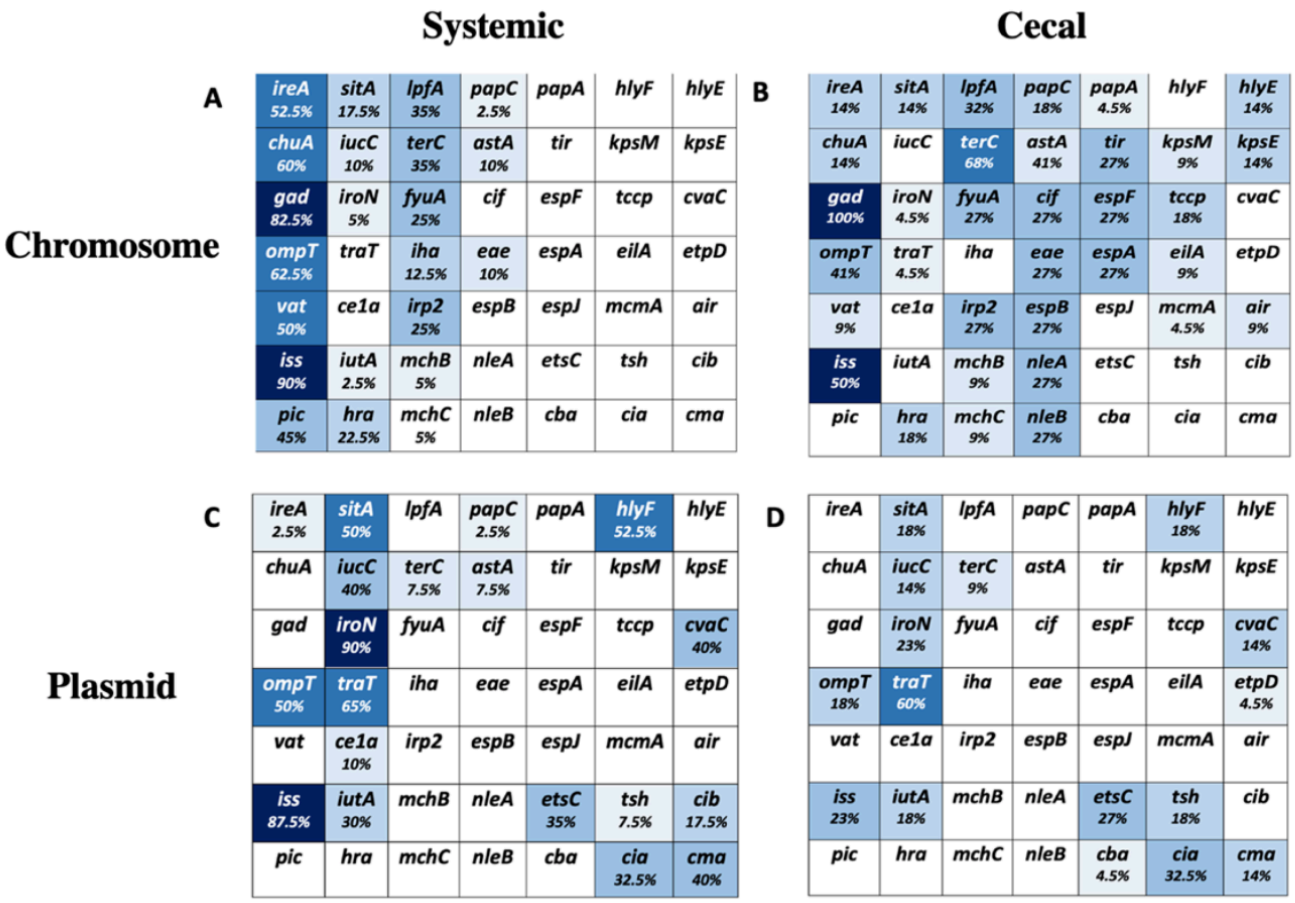

Figure 2.7. A. Percent abundance of virulence genes present on the chromosome of systemic *E. coli* where light blue = >10%, medium blue = 10-30%, and dark blue = >30%. **B**. Percent abundance of virulence genes present on the chromosome of cecal *E. coli*. **C.** Plasmid-encoded virulence genes from systemic *E. coli*, and **D**. plasmid-encoded virulence genes present in cecal *E. coli*.

Different sets of virulence genes appear to be enriched in systemic isolates, especially those found on plasmids. Plasmid-mediated colicin genes (*cia, cib, cma*) were frequently associated with beta-lactam resistance genes – 51% of systemic isolates possessed plasmids containing both colicin and beta-lactam resistance genes - though an analysis of a larger number of strains will be required to determine whether this is significant. Many chromosomal virulence genes are associated with pathogenicity islands, such as the *papC* and *lpfA* genes found in systemic isolates, and several effectors typically belonging to the EPEC belonging to the locus of enterocyte effacement found in cecal isolates (Deng *et al*., 2004; Najafi *et al*., 2018). We identified several plasmid-mediated genes present in over 50% of systemic isolates, but largely absent in cecal isolates. The significant difference in plasmid content may play a role in infection and ability to cause disease. Our *E. coli* come from a diverse group of genetic backgrounds, as determined by phylogenetic analysis, suggesting that plasmid-mediated virulence may play an important role in colibacillosis in broilers. However, this analysis is preliminary and biased towards known AMR and virulence determinants and will need to be redone in the future to confirm these findings.

2.3.4 Future Work: Hybrid Genomes and Larger Genomic Comparisons

I was interested in learning to use different tools to assemble and analyze long read sequencing data. Experimenting with these tools using rough draft genome assemblies uncovered potential differences in the two populations which helped us develop new questions and shape the direction of this thesis. However, it was always our intent to properly assemble hybrid genomes with this data and analyze it with an unbiased approach. A postdoctoral fellow in the White Lab, Dr. Haley Sanderson, will be using this Nanopore and Illumina sequencing data to generate highly accurate hybrid genome assemblies in the future. She will generate those genomes using long-read assemblies from the program Flye and polished with Illumina reads to increase the sequence accuracy (Kolmogorov *et al*., 2019). These polished genomes will be used in a larger, unbiased genomic analysis to compare virulence and AMR, as well as search for regions of the genome which may play a key role in virulence and environmental persistence.

2.4 References

Alikhan, N., Petty, N.K., Zakour, N.L.B., Beatson, S.A. (2011). BLAST ring image generator (BRIG): simple prokaryote genome comparisons. *BMC Genom*. 12(402). doi: 10.1186/1471-2164-12-402.

Anaconda Software Distribution. Computer Software. Vers 2-2.4.0. Anaconda, Nov. 2016. Web. <https://anaconda.com>.

Andrews, S. (2010). FASTQC. A quality control tool for high throughput sequence data. Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

Beghain, J., Bridier-Nahimas, A., Nagard, H., Denamur, E., Clermont, O. (2018). Clermontyping: an easy-to-use and accurate in silico method for *Escherichia* genus strain typing. *Microb. Genom*. doi: 10.1099/mgen.0.000192.

Bolger, A.M., Lohse, M., Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 30(15): 2114-2120. doi: 10.1093/bioinformatics/btu170.

Carattoli, A., Zankari, E., Garcia-Fernandez, A., Larsen, M.V., Lund, O., Villa, L., *et al*. (2014). *In Silico* detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob. Agents Chemother*. 58(7). doi: 10.1128/AAC.02412-14.

Chui, H., Chan, M., Hernandez, D., Chong, P., McCorrister, S., Robinson, A., Walker, M., *et al*. (2015). Rapid, sensitive, and specific *Escherichia coli* H antigen typing by matrix-assisted laser desorption ionization-time of flight-based peptide mass fingerprinting. *J. Clin. Microbiol*. 53(8): 2480-5. doi: 10.1128/JCM.00593-15.

De Coster, W., D'Hert, S., Schultz, D.T., Cruts, M., Van Broeckhoven, C. (2018). NanoPack: visualizing and processing long-read sequencing data.

Deng, W., Puente, J.L., Gruenheid, S., Li, Y., Vallance, B.A., Vasquez, A., Barba, J., Ibarra, A., *et al*. (2004). Dissecting virulence: systematic and functional analyses of a pathogenicity island. *PNAS*. 101(10): 3597-3602. doi: 10.1073/pnas.0400326101.

Desjardins, P., Conklin, D. (2010). NanoDrop microvolume quantitation of nucleic acids. *J. Vis. Exp*. (45): 2565. doi: 10.3791/2565.

Gilchrist, C.A., Turner, S.D., Riley, M.F., Petri, W.A., Hewlett, E.L. (2015). Whole-genome sequencing in outbreak analysis. *Clin. Rev. Microbiol*. 28(3): 541-563. doi: 10.1128/CMR.00075-13.

Goodwin, S., McPherson, J.D., McCombie, W.R. (2016). Coming of age: ten years of next-generation sequencing technologies. *Nat. Rev. Genet*. 17: 333-351. doi: 10.1038/nrg.2016.49.

Kleinheinz, K.A., Joensen, K.G., Larsen, M.V. (2014). Applying the ResFinder and VirulenceFinder web-services for easy identification of acquired antibiotic resistance and *E. coli* virulence genes in bacteriophage and prophage nucleotide sequences. *Bacteriophage*. 4: e27943. doi: 10.4161/bact.27943.

Kolmogorov, M., Yuan, J., Lin, Y., Pevzner, P. (2019). Asesmbly of long error-prone reads using repeat graphs. *Nat. Biotechnol*. 37: 540-546. doi: 10.1038/s41587-019-0072-8.

Kono, N., Arakawa, K. (2019). Nanopore sequencing: review of potential applications in functional genomics. *Develop. Growth Differentiation*. 61(5): 316-326. doi: 10.1111/dgd.12608.

Krawczyk, P.S., Lipinski, L., Dziembowski, A. (2018). PlasFlow: predicting plasmid sequences in metagenomic data using genome signatures. *Nucleic Acid Res*. 46(6): e35. doi: 10.1093/nar/gkx1321.

Leopold, S.R., Goering, R.V., Witten, A., Harmsen, D., Mellmann, A. (2014). Bacterial whole-genome sequencing revisited: portable, scalable, and standardized analysis for typing and detection of virulence and antibiotic resistance genes. *J. Clin. Microbiol*. 52(&): 2365-2370.

Lohman, N., Watson, M., Loose, M. (2018). A training bootcamp based on the Oxford Nanopore MinION. Web. <http://porecamp.github.io/2015/handbook.html>

Lopez, R., Chen, Y., Ang, S.D., Yekhanin, S., Makarychev, K., Racz, M.Z., Seelig, G., *et al*. (2019). DNA assembly for nanopore data storage readout. *Nat Comm*. 10: 2933. doi: 10.1038/s41467-019-10978-4.

Maghini, D.G., Moss, E.L., Vance, S.E., Bhatt, A.S. (2020). Improved high-molecular-weight DNA extraction, nanopore sequencing, and metagenomic assembly from the human gut microbiome. *Nature Prot*. 16: 458-471.

Maiden, M.C.J., van Rensburg, M.J.J., Bray, J.E., Earle, S.G., Ford, S.A., Jolley, K.A., McCarthy, N.D. (2013). MLST revisited: the gene-by-gene approach to bacterial genomics. *Nat. Rev. Microbiol*. 11: 728-736. doi: 10.1038/nrmicro3093.

Mardis, E., McCombie, W.R. (2017). Library quantification: fluorometric quantitation of double-stranded or single-stranded DNA samples using the Qubit system. *Cold Spring Harb. Protoc*. doi: 10.1101/pdb.prot094730.

McCombie, W.R., McPherson, J.D., Mardis, E.R. (2019). Next-generation sequencing technologies. *Cold Spring Harb. Perspect. Med*. doi: 10.1011/cshperspect.a036798.

Nagar, R., Schwessinger, B. (2018). DNA size selection (>3-4kb) and purification of DNA using an improved homemade SPRI beads solution. v.1. *Protocols.io* doi: 10.17504/protocols.io.n7hdhj6.

Najafi, A., Hasanpour, M., Askary, A., Aziemzadeh, M., Hashemi, N. (2018). Distribution of pathogenicity island markers and virulence factors in new phylogenetic groups of uropathogenic *Escherichia coli* isolates. *Folia Microbiol*. 63: 335-343. doi: 10.1007/s12223-017-0570-3.

Newman, D.M., Barbieri, N.L., Oliveira, A.L., Willis, D., Nolan, L.K., Logue, C.M. (2021). Characterizing avian pathogenic *Escherichia coli* (APEC) from colibacillosis cases, 2018. *PeerJ* 9: e11025. doi: 10.7717/peerj.11025.

Niemi, R.M., Mentu, J., Siitonen, A., Niemela, S.I. (2003). Confirmation of *Escherichia coli* and its distinction from *Klebsiella* species by gas and indole formation at 44 and 45°C. *J. Appl. Microbiol*. 95: 1242-1249. doi: 10.1046/j.1365-2672.2003.02125.x.

Purnell, R.F., Mehta, K.K., Schmidt, J.J. (2008). Nucleotide identification and orientation discrimination of DNA homopolymers immobilized in a protein nanopore. *Nano. Lett*. 8(9): 3029-3034.

Quainoo, S., Coolen, J.P.M., van Hijum, S.A.F.T., Huynen, M.A., Melchers, W.J.G., van Schaik, W., *et al*. (2017). Whole-genome sequencing of bacterial pathogens: the future of nosocomial outbreak analysis. *Clin. Microbiol. Rev*. 30(4): 1015-1063. doi: 10.1128/CMR.00016-17.

Quick, J., Lohman, N. (2019). DNA extraction strategies for Nanopore sequencing. In, Nanopore Sequencing, an Introduction. Branton, D., Deamer, D. (World Scientific), pp. 91-106.

Rang, F.J., Kloosterman, W.P., Ridder, J. (2018). From squiggle to basepair: computational approaches for improving nanopore sequencing read accuracy. *Genome Biol*. 19(90). doi: 10.1186/s13059-018-1462-9.

Sauget, M., Valot, B., Bertrand, X., Hocquet, D. (2017). Can MALDI-TOF mass spectrometry reasonably type bacteria? *Trends in Microbiol*. 25(6). doi: 10.1016/j.tim.2016.12.006.

Smith, A.D., Xuan, Z., Zhang, M.Q. (2008). Using quality scores and longer reads improves accuracy of Solexa read mapping. *BMC Bioinf*. 9(128). doi: 10.1186/1471-2105-9-128.

Sohn, J., Nam, J. (2018). The present and future of *de novo* whole-genome assembly. *Briefings in Bioinf*. 19(1): 23-40. doi: 10.1093/bib/bbw096.

Sola-Gines, M., Cameron-Veas, K., Badiola, I., Dolz, R., Majo, N., Dahbi, G., Viso, S., Mora, A., *et al*. (2015). Diversity of multi-drug resistant avian pathogenic *Escherichia coli* (APEC) causing outbreaks of colibacillosis in broilers during 2012 in Spain. *PLoS ONE* 10(11): e0143191. doi: 10.1371/journal.pone.1043191.

Utturkar, S.M., Klingeman, D.M., Land, M.L., Schadt, C.W., Doktycz, M.J., Pelletier, D.A., *et al*. (2014). Evaluation and validation of *de novo* and hybrid assembly techniques to derive high-quality genome sequences. *Bioinf*. 30(19): 2709-2716. doi: 10.1093/bioinformatics/btu391.

Vaughn, R.H., Osborne, J.T., Wedding, J.T., Tabachnick, J., Beisel, C.G., Braxton, T. (1950). The utilization of citrate by *Escherichia coli*. *J. Bacteriol*. 60(2): 119-127. PMID: 14774362.

Wagner, J., Coupland, P., Browne, H.P., Lawley, T.D., Francis, S.C., Parkhill, J. (2016). Evaluation of PacBio sequencing for full-length bacterial 16S rRNA gene classification. *BMC Microbiol*. 16(274). doi: 10.1186/s12866-016-0891-4.

Wick, R.R., Judd, L.M., Gorrie, C.L., Holt, K.E. (2017a). Completing bacterial genome assemblies with multiplex MinION sequencing. *Microb. Genom*. 3(10): e000132. doi: 10.1099/mgen.0.000132.

Wick, R.R., Judd, L.M., Gorrie, C.L., Holt, K.E. (2017b). Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comp. Biol*. 13(6): e1005595. doi: 10.1371/journal.pcbi.1005595.

Wunschel, S.C., Jarman, K.H., Petersen, C.E., Valentine, N.B., Wahl, K.L., Schauki, D., *et al*. (2005). Bacterial analysis by MALDI-TOF mass spectrometry: an inter-laboratory comparison. *J.A.S. Mass Spec*. 16(4): 456-462. doi: 10.1016/j.jasms.2004.12.004

CHAPTER 3. A DEDICATED PLASMID CAPTURE STEP IS REQUIRED TO CAPTURE THE TOTAL PLASMID CONTENT OF AVIAN *E. COLI* USING NANOPORE SEQUENCING

Submitting to Microbial Genomics, Spring 2022

3.1 Abstract

Non-conserved *E. coli* genes, also known as accessory genes, include virulence and antimicrobial resistance genes that are a significant human health concern. These accessory genes are frequently carried on mobile genetic elements such as plasmids. Plasmid analysis can be used to diagnose disease, track antimicrobial resistance genes and outbreaks, and identify novel metabolic and virulence factors. Long read sequencing platforms such as Oxford Nanopore Technologies and Pacific Biosciences provide tools to sequence and assemble plasmids in their entirety. Nanopore sequencing is particularly useful for public health surveillance of plasmid data due to the mobility and quick turnaround time for sequencing and analysis. Current plasmid sequencing protocols rely heavily on reference-based assembly, which requires the plasmids to have been identified and assembled previously. A preliminary analysis of 96 avian *E. coli* isolates suggests that *E. coli* isolated from broilers with confirmed colibacillosis possess significantly more plasmids than cecal isolates recovered from healthy broilers. To determine whether our whole genome sequencing protocol missed plasmids from these isolates, we developed and tested a methodology for sequencing and identifying plasmids from avian *E. coli*. We compared six different commercial plasmid extraction kits and two Nanopore library preparation kits to determine the optimum method for capturing a wide range of plasmids from avian *E. coli*. Using this method, we were able to identify four new plasmids that were missed using a standard genomic sequencing protocol. We identified plasmids ranging from 2kb to 300kb in length using a linear, ligation-based protocol to preserve long fragment lengths for *de novo* assembly. We determined that plasmid DNA from two commercial extraction kits is required to identify all small plasmids in two *E. coli* isolates. If plasmids truly are significantly associated with avian pathogenic *E. coli,* then a rapid plasmid identification protocol could have future implications as a diagnostic screening tool for pathogenic *E. coli* isolates.

3. 2 Introduction

Plasmids play a significant role in bacterial evolution. Mobile genetic elements are often the drivers of antimicrobial resistance, metabolic adaptation, and bacterial weaponry (De Toro et al., 2014). The *Escherichia coli* core genome, a conserved set of genes present in all *E. coli* isolates, consists of approximately 2,200 genes, which is less than half of the total gene content of a typical *E. coli* isolate (Woksepp et al., 2017). Many *E. coli* isolates acquire additional genes through horizontal gene transfer; these genes make up the accessory genome (Subedi et al., 2019). The size of the accessory genome is largely determined by the gain and loss of extrachromosomal DNA, including plasmids (Pitout, 2012). The *E. coli* genome contains hotspots of genomic recombination while still maintaining clonality in other regions (Kemmett et al., 2013). Previous research suggests that some lineages of *E. coli* are more receptive to genetic elements in those chromosomal hotspots, and thus may be more pathogenic (Kemmett et al., 2013).

Avian pathogenic *E. coli (*APEC) are responsible for a variety of acute and systemic diseases in poultry (Guabiraba and Schouler, 2015). These isolates are highly diverse and possess a highly varied set of virulence factors, making them difficult to classify prior to outbreaks of disease (Schouler *et al.,* 2012; Newman *et al.,* 2021). How APEC isolates differ from commensal *E. coli* residing in the chicken gut is unknown, due to considerable overlap in gene content, phylogeny, and serogroup (Cordoni *et al.,* 2016; Cunha *et al.,* 2017; Halfaoui *et al.,* 2017; DebRoy *et al.,* 2018). The presence of several APEC-associated may be a defining characteristic of pathogenic strains of *E. coli* (Mellata *et al.*, 2010). In addition to pathogenesis, plasmid-mediated antimicrobial resistance is common in many human and avian extraintestinal pathogenic isolates, and these plasmids typically contain additional virulence genes (Pitout, 2012; Johnson *et al.,* 2012). The identification of plasmid regions homologous to those of human *E. coli*, particularly those that cause neonatal meningitis, in avian isolates raises concerns about the potential for zoonotic infections (Fricke *et al.,* 2009; Johnson *et al.,* 2009; Logue *et al.,* 2017).

Plasmid sequencing is essential for understanding and tracking the spread of mobile antimicrobial resistance elements (George *et al.,* 2017; Weingarten *et al.*, 2018). One of the early challenges of sequencing bacterial plasmids was the size limitation presented by Roche-454 and

Illumina sequencing (Brouwer *et al*., 2015; Antipov *et al*., 2016). These technologies are limited to read lengths under 1000bp, which are sufficient for reference-based assembly but are difficult to fully assemble using *de novo* methods (Chhangawala *et al*., 2015; Sommerville *et al*., 2019). *De novo* assembly methods produce highly fragmented bacterial genomes and plasmids when repeat regions span a greater distance than the length of the sequencing reads (Liao *et al*., 2019; Sommerville *et al*., 2019). Advances in long read sequencing (Oxford Nanopore Technologies, Pacific Biosciences) have helped solve some of the challenges of bacterial *de novo* assembly (Lemon *et al*., 2017). These technologies can generate sequencing reads that are thousands of bases long (Lemon *et al*., 2017). Many bacterial genomes and plasmids are now sequenced using a hybrid method, which combines short and long reads to produce genomes with high sequence accuracy and correct gene order (Wick *et al*., 2017a; Lemon *et al*., 2017). However, hybrid strategies are not ideal for rapid diagnostic purposes due to the more time-consuming short read sequencing process (Li *et al*., 2018).

Rapid sequencing using the Oxford Nanopore MinION is an effective tool for identifying mobile resistance elements in clinical bacterial pathogens, however these are targeted protocols to sequence known plasmids (George *et al*., 2017; Li *et al*., 2018). To date, there is no standardized protocol for MinION sequencing of *E. coli* plasmids from avian or environmental sources. Our aim was to develop a protocol that could be used to sequence and identify the total plasmid content of avian *E. coli* isolates. Standard Nanopore sequencing protocols are optimized for sequencing of high molecular weight reads. These typically include a size selection step to remove small DNA fragments, which may make recovery and identification of small plasmids difficult (Wick *et al*., 2018; Lima *et al*., 2019).

Using a standard whole genome sequencing protocol, we identified significantly more plasmids in *E. coli* recovered from broilers with confirmed colibacillosis (systemic *E. coli*) than in *E. coli* isolated from the cecal contents of healthy broilers (cecal *E. coli)*. Almost all AMR genes identified in these isolates were present on plasmids. We tested six different commercial plasmid extraction kits and two different Nanopore library preparation methods to develop a protocol that could reproducibly identify a wide range of plasmids in two *E. coli* isolates. Using this methodology, we identified at least one new plasmid in both tested isolates which was missed using our previous whole genome sequencing protocol. This method captures plasmids ranging in size

from 2kb to 256kb. Finally, we determined that the average fragment length of a sample played the greatest role in the identification of plasmids and combining sequencing reads generated from two different plasmid kits was required to identify every plasmid in both isolates.

3. 3 Materials and Methods

3.3.1 Whole Genome Sequencing

96 *E. coli* isolates were isolated from colibacillosis outbreaks on Saskatchewan broiler farms according to the method described in section 3.3.1 of this thesis. These isolates were cultured from freezer stocks (LB broth + 50% glycerol) on LB agar and a single colony was grown overnight in LB broth with shaking. DNA was isolated according to the GenElute Bacterial Genomic DNA Kit (Sigma; #NA2110) and stored at -20°C. A size selection to eliminate fragments under 500 bp was performed using the protocol for short fragment elimination with NucleoMag magnetic beads (Macherey-Nagel) and eluted in sterile water. Barcodes (ONT, NBD104) and sequencing adapters were ligated to 12 *E. coli* genomic DNA samples using the Oxford Nanopore Ligation Sequencing Kit (SQK-LSK109). Multiplexed isolates were sequenced for 24 hours on a MinION R.9.4.1 flow cell. Barcodes were removed using Guppy debarcoder (ONT) and Porechop (Wick *et al*., 2017a). Draft genome of chromosome and plasmid sequences were assembled from MinION long reads using Unicycler (Wick *et al*., 2017b).

3.3.2 Screening for Plasmids

To screen for plasmids, draft genome contig files were analyzed by PlasFlow, a software which uses a trained neural network to distinguish plasmid from chromosomal sequences (Krawczyk *et al.,* 2018). Contigs containing suspected plasmids were confirmed via BLAST search. Plasmid Finder identified known *Enterobacteriaceae* plasmids in each draft assembly by detecting the sequence of the origin of replication and matching it to a replicon in the database (Carattoli *et al.,* 2014). We used a minimum cut-off of 80% nucleotide identity to determine the incompatibility type for each plasmid using Plasmid Finder.

3.3.3 Plasmid DNA Extraction

Two avian *E. coli* isolates from the same colibacillosis outbreak were previously sequenced using our laboratory MinION sequencing protocol (see chapter 2 methods). One isolate was recovered from the spleen of a broiler with confirmed systemic colibacillosis (4957-3S1), and the other was isolated from the cecal contents of a healthy broiler from the same flock (4957-C3). Both *E. coli* isolates were streaked on LB agar from freezer stocks and a single colony was incubated overnight in LB broth with shaking for 16 hours. Plasmid DNA was isolated according to the manufacturer's instructions from six commercial plasmid extraction kits (**Table 3.1**). DNA was stored at -20°C until sequencing. Using the Ligation Sequencing Kit (SQK-LSK109), barcodes (NBD104) and adapters were ligated to any nicked or linear plasmid DNA. A separate library preparation using the Rapid Barcoding Kit (SQK-RBK004) inserted transposon-tagged barcodes into plasmid DNA. The MinION run using the Ligation Sequencing Kit samples was terminated at 14 hours due to a lack of available DNA present on the flow cell. The Rapid Barcoding Kit sequencing run was terminated after 24 hours. Both samples were sequenced on new R.9.4.1 flow cells.

Table 3.1 List of plasmid extraction kits used to extract plasmid DNA from avian *E. coli* isolates

Kit Name	Manufacturer	Product Number
GenElute Plasmid Miniprep Kit	Millipore Sigma	PLN70
NucleoSpin Plasmid Mini Kit	Macherey-Nagel	740588.50
Presto Mini Plasmid Kit	Geneaid	PD100
Monarch Plasmid Miniprep Kit	NEB	T1010S
Plasmid Midi Kit	Qiagen	12143
GeneJET Plasmid Miniprep Kit	Thermo Fisher	K0502

3.3.4 *De Novo* Assembly/Identification of Plasmids

Barcodes were removed from both sets of reads using Guppy (ONT) and Porechop (Wick *et al*., 2017). Read length information and histograms of read lengths for each sequencing run were generated using NanoStat and NanoPlot, respectively (De Coster *et al*., 2018). De *novo* assembly was performed using Flye, a long-read assembler optimized for plasmid sequences (Kolmogorov *et al*., 2019). Read mapping to determine the amount of chromosomal contamination was performed using Samtools and Minimap2 (Li *et al*., 2009; Li, 2018). After assembly, plasmid

sequence and type, as well as virulence and AMR genes were identified using the tools and methods described in section 3.3.2.

3.4 Results

3.4.1 Systemic Isolates Possess Significantly More Plasmids than Cecal Isolates

Contigs from 96 draft *E. coli* genomes (58 systemic, 38 cecal) were screened for plasmid content using PlasFlow. Systemic isolates possessed an average of 4 plasmids (range from 0-8), while cecal isolates possessed an average of 1 plasmid (range from 0-4).

3.4.2 A Dedicated Plasmid Capture Step Identified 4 New Plasmids Missed During WGS

To confirm the plasmid-related differences observed in systemic and cecal isolates, we needed to determine whether we were missing plasmids during WGS. We extracted plasmid DNA from two isolates from a single outbreak in February 2020: a systemic isolate, 4957-3S1, and a cecal isolate, 4957-C3, both of which had been sequenced previously. We previously identified three plasmids in isolate 4957-3S1: a 137 kb plasmid belonging to the IncFIB incompatibility group, a 167 kb IncB/O/K/Z plasmid, and a 256 kb IncH12A plasmid. Isolate 4957-C3 possessed two plasmids: a 138 kb IncFIB plasmid, and a 96 kb p0111 plasmid. A summary of the plasmids found in the original WGS assemblies can be found in **Table 3.2**.

Table 3.2. Reference Assemblies of virulence plasmids present in WGS assemblies of a systemic (4957-3S1) and a cecal (4957-C3) isolate from DNA extracted using a genomic DNA purification kit and the Nanopore Ligation Sequencing Kit. Plasmid names beginning with an 'S' represent systemic plasmids, while plasmids beginning with a 'C' represent cecal plasmids.

Reference Assembly	Plasmid Name	Incompatibility Group	Size
4957-3S1	S1	IncFIB	137kb
	S2	IncB/O/K/Z	167kb
	S3	IncH12A	256kb
4957-C3	C1	IncFIB	138kb
	C2	p0111	96kb

The plasmid capture step identified four new plasmids ranging in size from 2-100 kb: three in 4957-3S1 and one in 4957-C3 (**Table 3.3**). We found two small plasmids, ColpV (2 kb) and Col_1 (2.3 kb) in 4957-3S1, as well as a large 97 kb p0111 plasmid. In 4957-C3 we identified a 2 kb ColpV plasmid.

Table 3.3. Summary of the plasmid content of the systemic (4957-3S1) and cecal (4957-C3) isolates after targeted sequencing was performed.

Isolate	*Plasmid Name	Incompatibility Group	Size
4957-3S1	S1	IncFIB	137 kb
	S2	IncB/O/K/Z	167 kb
	S3	IncHI2A	256 kb
	S4	**ColpV**	**2 kb**
	S5	**p0111**	**97 kb**
	S6	**Col_1**	**2.3 kb**
4957-C3	C1	IncFIB	137 kb
	C2	p0111	96 kb
	C3	**ColpV**	**2 kb**

*Plasmids that were identified using targeted plasmid purifications followed by DNA sequencing are in **bold**. These plasmids were not recovered in standard genomic DNA sequencing.

3.4.3 The Ligation Sequencing Kit produces fewer but longer reads

The standard Nanopore Ligation Sequencing Kit (Oxford Nanopore Technologies; #SQK-LSK109) requires that input DNA be linearized prior to library preparation to successfully ligate

barcodes and adapters for sequencing. We hypothesized that during routine plasmid purification some proportion of the circular plasmid DNA molecules would become 'nicked' and could have barcode primers ligated directly without performing additional fragmentation steps. To determine whether we could directly sequence plasmid DNA without a dedicated fragmentation step we performed sequencing runs using two Nanopore library kits: the Ligation Sequencing Kit, which ligates adapters onto linear, double-stranded DNA, and the Rapid Barcoding Kit (Oxford Nanopore Technologies, #SQK-RBK004), which uses a transposon-based barcoding system to fragment and linearize high molecular weight and circular DNA before attaching sequencing adapters (**Fig 3.1**).

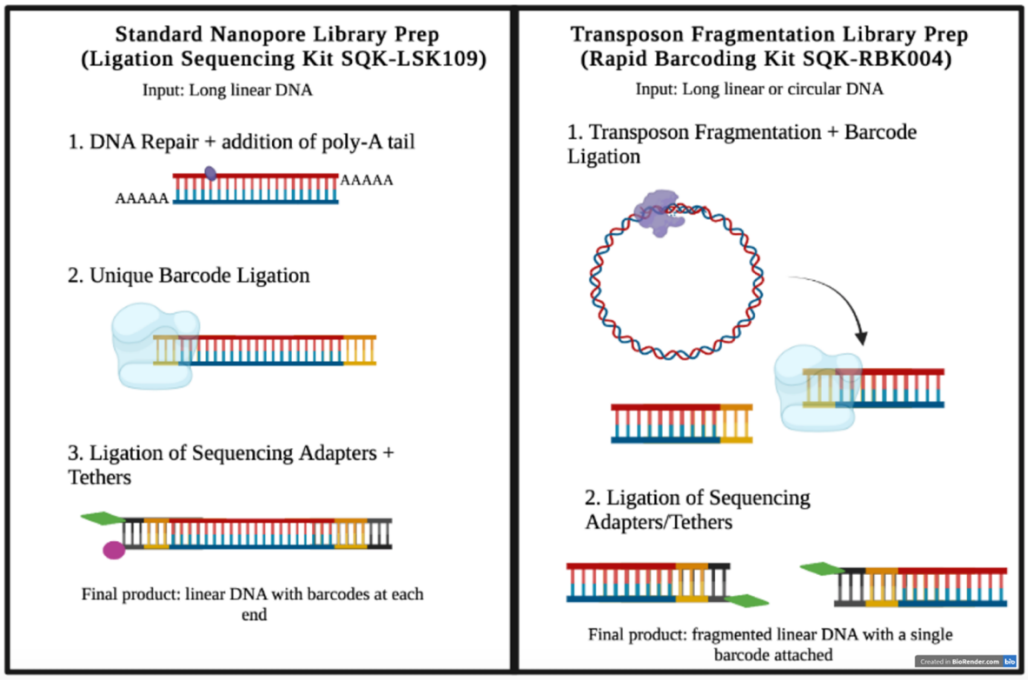

Figure 3.1. Illustration of the methods used in the standard Nanopore Ligation Sequencing Kit (SQK-LSK109) and the transposon-based Rapid Barcoding Kit (SQK-RBK004). Image created using BioRender.

The MinION sequencing run using the ligation-barcoded DNA was terminated at 15h, sequencing could not continue past this point due to an insufficient amount of linear, barcoded-DNA. In a second sequencing run using the rapid-barcoded DNA, fragmentation linearized a sufficient amount of plasmid DNA to proceed to the normal 24h point before terminating the run.

Sequencing plasmids with the Ligation Sequencing Kit (LSK) produced fewer reads than the Rapid Barcoding Kit (RBK), but the lack of a fragmentation step meant that reads from the LSK sequencing run had longer mean read lengths (**Fig 3.2A**). The amount of linear plasmid DNA eluted from every plasmid extraction kit was sufficient to generate plasmid assemblies: the smallest number of reads from any kit without additional fragmentation was 16,000 (**Table 3.4**). Commercial plasmid extraction kits varied greatly in plasmid fragmentation: some commercial plasmid kits produced heavily fragmented DNA (Sigma, Macherey Nagel, and GeneAid) which produced hundreds of thousands of sequencing reads, while others (NEB, Qiagen) eluted largely intact plasmids and produced tens of thousands of reads (**Table 3.4**). Avoiding an enzymatic fragmentation step produced longer reads that are advantageous for *de novo* assembly. The RBK kit is optimized for fragmentation of DNA larger than 30kb in size. All of the commercial plasmid purification kits produced sequencing reads that are considerably shorter than 30kb (**Fig 3.2**). As a result, the RBK kit with the plasmid DNA eluate produced more than two times as many sequencing reads compared to the LSK kit. However, this resulted in a shorter mean read length (**Fig 3.2B**). In addition to greater mean read lengths, library preparation using the LSK kit produced a slightly higher mean read quality in 5/6 sequencing runs (**Table 3.4**). The first step of the LSK protocol includes a DNA repair step to repair nicks and gaps in the reads, which improves the overall sequence quality. This step is not included in the RBK protocol.

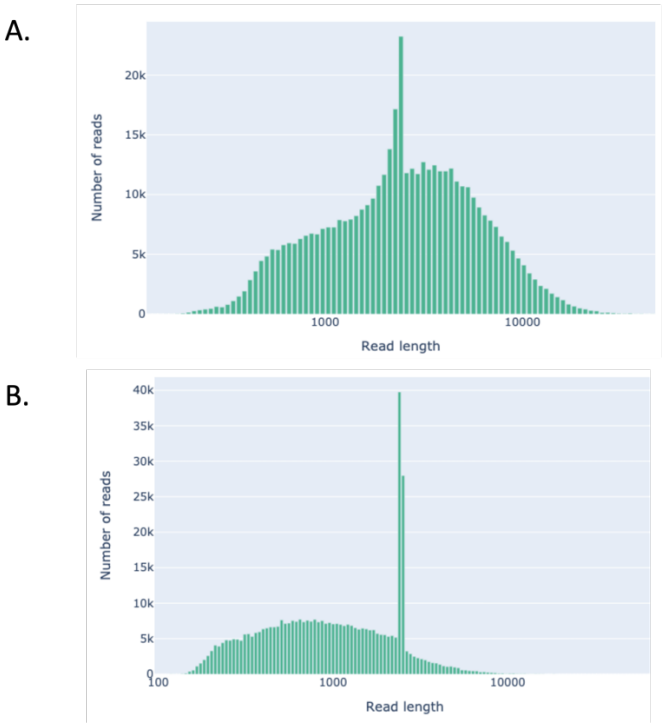

Figure 3.2. Histogram of read lengths and number of reads generated using DNA extracted from strain 4957-3S1 using the Thermo Fisher GeneJet Extraction kit prepared with **A**. Ligation Sequencing Kit protocol (LSK), and **B**. Rapid Barcoding Kit protocol (RBK).

Table 3.4. Summary of read numbers and read quality using six commercial plasmid kits and a ligation-based (LSK109) or transposon-based (RBK004) Nanopore library preparation method.

Plasmid Kit	Isolate	LSK109 # Reads	Mean Read Quality	RBK004 # Reads	Mean Read Quality
Sigma	Systemic	116,000	10.8	88,401	10.2
	Cecal	400,000	11.2	153,344	10.4
Macherey-Nagel	Systemic	32,000	10.8	6,835	10.1
	Cecal	192,000	11.1	13,108	10.4
GeneAid	Systemic	100,000	10.7	49,787	10.3
	Cecal	408,000	11.2	165,547	10.6
NEB	Systemic	20,000	10.6	140,090	10.4
	Cecal	20,000	10.8	268,113	10.7
Qiagen	Systemic	16,000	9.5	184,000	10.4
	Cecal	24,000	10.4	348,287	10.7
Thermo Fisher	Systemic	80,000	10.8	176,589	10.3
	Cecal	140,000	11.1	145,569	10.4

3.4.4 Longer reads generate a more complete *de novo* plasmid assembly

The fragmentation step used in the RBK protocol produced very small DNA fragments (mean read lengths of under 2kb) that were difficult for the long-read assembly algorithm to reassemble (**Fig 3.3**). Commercial plasmid kits that eluted mostly intact, circular plasmid DNA (NEB, Qiagen, Thermo Fisher) fared best using the RBK protocol (Thompson *et al.*, 2008). Overall, large plasmids with repeat regions were highly fragmented and difficult to assemble from RBK-barcoded DNA. RBK assemblies from each of the six commercial plasmid extraction kits consistently failed to identify the small 2kb plasmids in isolate 4957-3S1. In isolate 4957-C3, RBK assemblies were more successful– reads from 5 out of 6 kits identified the new 2kb plasmid. In contrast, the LSK samples produced more complete assemblies. Reads from 5/6 commercial plasmid extraction kits + LSK-barcoding were sufficient to assemble and identify 80% of the plasmids in the systemic isolate, while all three new plasmids were identified in DNA from 3/6 commercial kits (**Table 3.5**).

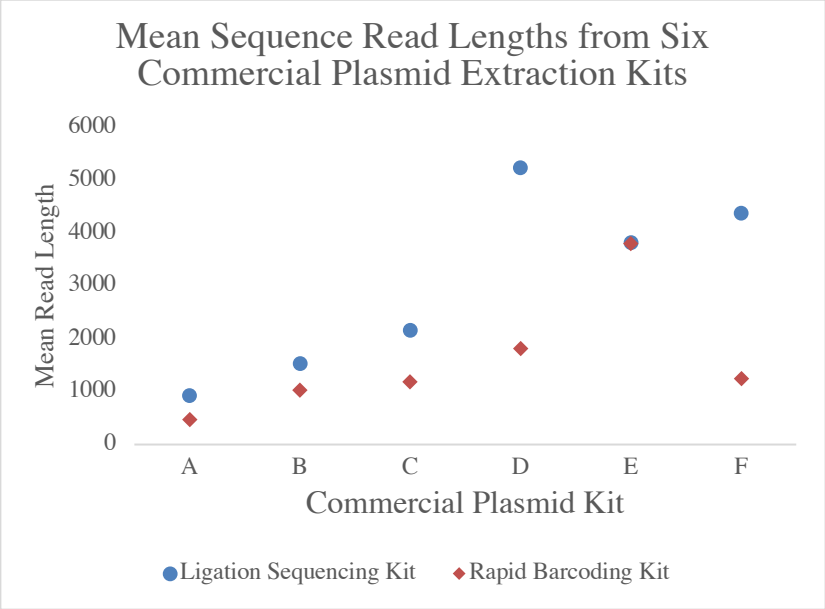

Figure 3.3. The mean read lengths of sequencing reads produced using six commercial plasmid extraction kits (A. Sigma, B. Macherey-Nagel, C. GeneAid, D. NEB, E. Qiagen, F. Thermo Fisher) using DNA extracted from systemic isolate 4957-3S1. Blue dots represent DNA prepared with the standard ligation sequencing kit (SQK-LSK109) and red dots represent DNA prepared using the transposon-based rapid barcoding kit (SQK-RBK004).

Table 3.5. Summary of assemblies and plasmids for each plasmid prep kit using Standard Prep and Transposon-based preps, including # of reads and percentage mapping to the chromosome.

E. coli isolate + plasmid prep kit	Standard reads	*% chrom mapping	Tn Reads	*% chrom mapping	Standard plasmids identified	Tn Plasmids Identified
4957-3S1						
Sigma	116,000	25.8	88,401	38.3	S1 **S4 S5 S6**	S1 **S5**
Macherey-Nagel	32,000	15	6,835	19.7	S1 S2 **S5**	
GeneAid	100,000	32.8	49,787	33.3	S1 S2 S3 **S4 S5 S6**	S1 S3 **S5**
NEB	20,000	49.5	140,090	29.7	S1 S2 S3 **S4 S6**	S1 S2 S3 **S5**
Qiagen	16,000	75	184,000	31.4	S1 S2 S3 **S4 S5**	S1 S3 **S5**
Thermo Fisher	80,000	42.9	176,589	36.1	S1 S2 S3 **S4 S6**	S1 S2 S3 **S5**
GeneAid + NEB	120,000	--	--	--	S1 S2 S3 **S5**	--
NEB + Qiagen	36,000	--	--	--	S1 S2 S3 **S4 S5 S6**	--
NEB + TF	100,000	--	--	--	S1 S2 S3 **S4 S5 S6**	--
Qiagen + TF	96,000	--	--	--	S1 S2 S3 **S4 S5 S6**	--
4957-C3						
Sigma	400,000	11.7	153,344	10		C1 C2 **C3**
Macherey-Nagel	192,000	7	13,108	8.1	C1 C2 **C3**	**C3**
GeneAid	408,000	11	165,547	15.7	**C3**	C1 C2 **C3**
NEB	20,000	44.4	268,113	7.7	C1 C2	C1 C2 **C3**
Qiagen	24,000	60.4	348,287	8.2	C1 C2 **C3**	C1 C2
Thermo Fisher	140,000	41.6	145,569	43.1	C1 C2 **C3**	C1 C2 **C3**
GeneAid + NEB	428,000	--	--	--	C2 **C3**	--
NEB + Qiagen	44,000	--	--	--	C1 C2 **C3**	--
NEB + TF	160,000	--	--	--	C1 C2 **C3**	--
Qiagen + TF	164,000	--	--	--	C1 C2 **C3**	--

* The proportion of reads that map back to the original *E. coli* isolate chromosomal assembly.

Sigma, Macherey Nagel, and GeneAid kits, which produced reads with the shortest mean read lengths, consistently failed to identify small plasmids (ColpV and Col_1) when they were prepared with the RBK enzymatic barcoding step. These kits produced highly fragmented assemblies which could not reconcile repeat regions within the large plasmids (S1, S2, and S3). A plasmid contig matching the S6 (ColpV) plasmid could consistently be mapped to a repeat region

in the larger S1 (IncFIB) plasmid in these assemblies (**Fig 3.4**). Plasmid sizes varied when using the RBK kit due to assembly errors in unresolved repeat regions.

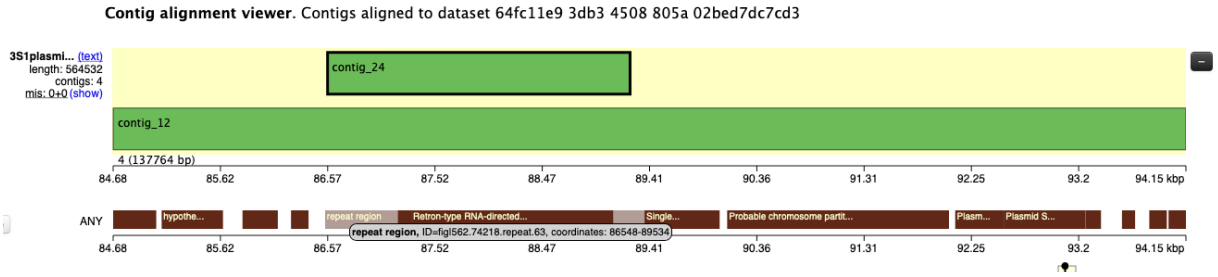

Figure 3.4. Quast genome assembly comparison showed that the S6 (ColpV) plasmid assembly from the GeneAid plasmid kit (contig 24) mapped to a repeat region in the larger S1 IncFIB plasmid (contig 12).

3.4.5 No single commercial kit + library preparation identified all the plasmids in either isolate

No combination of a single commercial plasmid extraction kit + library prep kit was able to consistently assemble and identify all six plasmids found in 4957-3S1 and 3 plasmids in 4957-C3. Combining the reads from two plasmid kits prepared with the LSK kit was sufficient to assemble all the identified plasmids in both isolates (**Table 3.5**). Any combination of commercial NEB, Qiagen, and Thermo Fisher kits, which eluted primarily intact plasmid DNA and produced mean read lengths of over 3kb, was able to assemble all the plasmids in both the systemic and the cecal isolate (**Table 3.5**). Because DNA prepared with the RBK kit was heavily fragmented, and plasmids were assembled using a long-read assembly program, combining multiple kits did not improve the identification of small plasmids in either isolate.

3.4.6 Commercial plasmid kits varied in chromosomal contamination

It is inevitable that some amount of chromosomal DNA will be present in the plasmid preparations. To determine the amount of chromosomal contamination in each kit, we mapped the sequencing reads to the previously assembled reference genome for each isolate as described in the methods section. An average of 40% of the reads mapped back to the chromosome of the systemic isolate, while an average of 29% of reads mapped back to the cecal chromosome (**Table 3.5**). A

large proportion of the systemic reads that matched the chromosome mapped to mobile genetic elements integrated into the *E. coli* chromosome.

3.5 Discussion

We tested six commercial plasmid extraction kits and two Nanopore library preparation kits to develop a methodology for capturing a wide range of plasmids from avian *E. coli* isolates. We determined that size selection during Nanopore library preparation favours whole genome *de novo* assembly but hinders the assembly and identification of smaller plasmids. Using our laboratory's standard WGS protocol, which includes a size selection step to remove DNA fragments smaller than 500bp to optimize sequencing quality, we identified three plasmids in 4957-3S1 and two plasmids in 4957-C3. After a dedicated plasmid DNA extraction, we identified an additional three plasmids in 4957-3S1 and one plasmid in 4957-C3, clearly demonstrating that an additional plasmid extraction step is required when using a long-read optimized WGS protocol.

The DNA extraction and library preparation methods used for the original WGS run favoured long fragment lengths and using this method we were able to sequence plasmids of 100 kb or greater in size. However, we were unable to capture a wide size range of plasmid diversity using this method. Standard Nanopore sequencing kits, such as the Ligation Sequencing kit (SQK-LSK109) are optimized for input DNA that is sheared to 8kb or longer, because the ligation reactions are optimized based on an expected number of molecules of DNA in a sample, rather than a particular DNA mass (Kono and Arakawa, 2019). A fragment size of 8kb or larger produces an optimal number of adapter-ligated DNA for Nanopore sequencing. Samples with highly fragmented DNA will produce much poorer sequencing outputs because only a small fraction of the sample will actually be sequenced. To identify and assemble plasmids using *de novo* Nanopore sequencing, longer DNA fragment lengths are required. Plasmid DNA extraction kits that produced short, highly fragmented plasmid DNA (Sigma, Macherey Nagel, and GeneAid) performed poorly during assembly, regardless of the type of library preparation kit used. Previous studies have shown that some commercial plasmid extraction kits produce largely intact, supercoiled plasmid DNA, while others are more likely to fragment plasmids (Thompson *et al.,* 2008). In particular, spin column purification tends to shear DNA more readily than other extraction methods due to the physical force required to move the DNA through the filter (Krsek and Wellington, 1999).

Similarly, enzyme fragmentation with the Rapid Barcoding kit (SQK-RBK004) greatly increased the number of sequencing reads for each plasmid sample, but at the cost of read length and *de novo* assembly quality. Large plasmids (90kb and greater) were still typically identified after enzymatic fragmentation, but small plasmids (2kb) were lost. We anticipated that a certain proportion of purified plasmid molecules would become nicked during extraction and library prep, producing sufficient linear ends to use a double stranded, ligation-based library preparation method (SQK-LSK109). Using the kits from NEB, Qiagen, and Thermo Fisher, we were able to recover enough nicked, linear plasmids to successfully identify large and small plasmids in both isolates without using a fragmentation step. We have determined that the best method to capture plasmids from avian *E. coli* is as follows: use two commercial plasmid extraction kits (NEB and Thermo Fisher) to produce enough long, linear DNA fragments for sequencing (**Fig 3.5**). Use the Nanopore Ligation Sequencing Kit (LSK109) to prepare samples for Nanopore sequencing. This kit preserves long fragment lengths for post-sequencing *de novo* plasmid assembly and identification with Flye and Plasmid Finder.

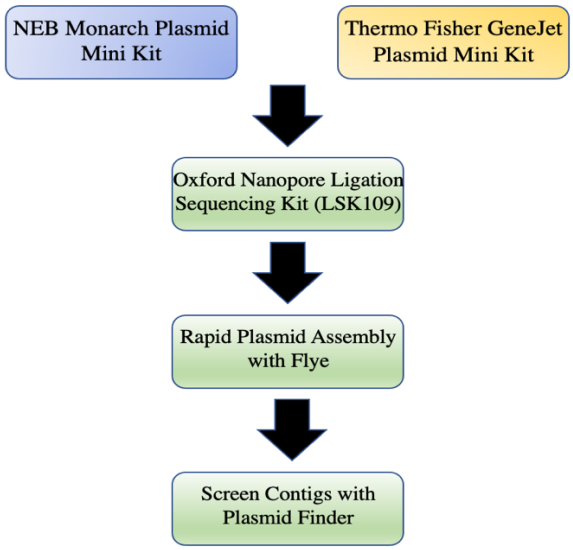

Figure 3.5. Schematic outlining the methodology for sequencing new plasmids in avian *E. coli* isolates using two plasmid extraction kits and the Ligation Sequencing Kit from Oxford Nanopore.

Characterizing the accessory genome of *E. coli* isolates, including plasmids, provides key information about potential pathogenesis and zoonotic risks (Sarowska *et al*., 2019). Large virulence plasmids are common among ExPEC (Nicholson *et al*., 2016; Sarowska *et al*., 2019). Sections of virulence plasmids which significantly increase fitness during infection have been incorporated into the UPEC genome (Cusumano *et al*., 2010). NMEC isolates often possess a conserved set of plasmids containing a core set of virulence genes which increase fitness *in vivo* and avoid host defenses (Nicholson *et al*., 2016). Large plasmids have previously been implicated in APEC pathogenicity, and their removal often results in attenuation in lethality assays (Johnson *et al*., 2008; Mellata *et al*., 2010; Mellata, 2013; Schouler *et al*., 2014). Given the strong sequence similarity between ExPEC virulence plasmids, APEC plasmids likely play a key role in invasion and infection of the host (Skyberg *et al*., 2007; Nicholson *et al*., 2016). The presence of virulence genes on these plasmids helps to facilitate their spread to a diverse group of *E. coli* via horizontal gene transfer (Johnson *et al*., 2008). In our study population of 96 *E. coli* isolates, systemic isolates possessed significantly more plasmids than cecal isolates. Perhaps the mobilization of genes on these plasmids is important for APEC fitness and survival during pathogenesis or in the broiler barn environment (Wawire *et al*., 2021).

Plasmids identified in commensal fecal *E. coli* from poultry typically belong to different incompatibility groups than those found in APEC isolates (Johnson *et al*., 2007). Two plasmids belonging to the same incompatibility group have the same origin of replication and cannot stably coexist in a bacterial cell over multiple generations, possibly because of competition for factors required for replication (Velappan *et al*., 2007). For example, IncFIB plasmids are more commonly associated with APEC strains than with commensal strains (Johnson *et al*., 2007). We identified an IncFIB plasmid in both isolates, but because they were isolated from broilers in the same barn, it is possible that this plasmid was widely circulating among bacteria in the environment but was not sufficient to cause disease in the absence of additional factors (Ruano-Gallego *et al*., 2020). The presence of small plasmids lacking virulence and AMR genes may also play a role in pathogenesis. Their exact mechanism is unknown, but they may be important for transmission of mobile elements. Small plasmids may play a role in transposon-mediated horizontal gene transfer (He et al., 2015). Additionally, small, high copy plasmids may counteract fitness defects from other antimicrobial resistance plasmids present in bacterial isolates by facilitating the coexistence of plasmid-mediated traits for multiple generations (Rodriguez-Beltran et al., 2018). The presence of

these plasmids in our *E. coli* isolates may provide a fitness advantage in the barn environment or during infection. Analysis of additional strains will be required to understand the role that plasmids play in outbreaks of Colibacillosis.

To confirm the difference we observed in the plasmid content of systemic and cecal *E. coli* isolated from Saskatchewan farms, we first needed to determine whether we were missing plasmids in our original WGS sequencing and draft assemblies. Bacterial DNA sequencing is typically performed in one of two ways: via a hybrid approach of Illumina sequencing plus a long-read instrument, or solely using Illumina sequencing (Bertelli and Greub, 2013; Kwong *et al*., 2015; Lohman and Pallen, 2015). Illumina sequencing produces high quality data with accurate basecalls, but the short read lengths make *de novo* genome assembly incredibly difficult (Kwong *et al*., 2015). In contrast, Nanopore sequencing provides real-time data, but the lower base calling accuracy of the long reads can introduce small errors in sequence accuracy in a genome assembly (Deamer *et al*., 2016). Previous studies investigating *E. coli* plasmids used an Illumina-based or hybrid approach to capture a wide variety of plasmids in *E. coli* samples (Gonzalez-Escalona *et al*., 2019; Taylor *et al*., 2019; Grevskott *et al*., 2020). A recent study by Wick *et al*. determined that the Oxford Nanopore MinION can capture both large and small plasmids (2021). However, there is no standardized methodology for capturing and assembling novel plasmids from *E. coli* isolates. Read-mapping-based methods are sufficient when an appropriate reference sequence is available, but due to the incredible diversity of *E. coli* plasmids this method is not always practical (de Toro *et al*., 2014; Arrendono-Alonso *et al*., 2017). In the absence of an appropriate reference, using long-read technologies such as Nanopore sequencing can produce a *de novo* assembly which provides a wealth of information about novel plasmids and can be used to target and sequence highly similar plasmids in the future (Rozov *et al*., 2017; Roosaare *et al*., 2018). Nanopore sequencing provides a rapid platform sufficient for *de novo* assembly. If systemic *E. coli* isolated from colibacillosis outbreaks on Saskatchewan broiler flocks truly possess significantly more plasmids than cecal isolates, this methodology could be used as a rapid preliminary diagnostic test for colibacillosis outbreaks. Rapid identification of antimicrobial resistance plasmids using this method can also be applied as a surveillance tool to monitor resistance trends on Saskatchewan farms.

3.6 References

Al-Kandari, F., Woodward, M.J. (2018). Genotypic and phenotypic diversity differences of presumptive commensal and avian pathogenic *E. coli*. *British Poul. Sci*. 60(1): 79-86. doi: 10.1080/00071668.2018.154415.

Antipov, D., Hartwick, N., Shen, M., Raiko, M., Lapidus, A., Pevzner, P.A. (2016). plasmidSPAdes: assembling plasmids from whole genome sequencing data. *Bioinformatics*. 32(22): 3380-3387. doi: 10.1093/bioinformatics/btw493.

Arredondo-Alonso, S., Willems, R.J., van Schaik, W., Schurch, A.C. (2017). On the (im)possibility of reconstructing plasmids from whole-genome short-read sequencing data. *Microb Genom*. 3(10): e000128. doi: 10.1099/mgen.0.000128.

Bertelli, C., Greub, G. (2013). Rapid bacterial genome sequencing: methods and applications in clinical microbiology. *Clin Microbiol and Infect*. 19(9): 803-813. doi: 10.1111/1469-0691.12217.

Bortolaia, V., Kaas, R.S., Ruppe, E., Roberts, M.C., Schwarz, S., Cattoir, V., *et al*. (2020). ResFinder 4.0 for predictions of phenotypes and genotypes. *J. Antimicrob. Chemo*. 75(12): 3491-3500. doi: 10.1093/jac/dkaa345.

Brandariz-Fontes, C., Camacho-Sanchez, M., Vila, C., Vega-Pla, J.L., Rico, C., Leonard, J.A. (2015). Effect of the enzyme and PCR conditions on the quality of high-throughput DNA sequencing results. *Nat. Sci. Rep*. 5: 8056. doi: 10.1038/srep08056.

Brouwer, M.S.M., Tagg, K.A., Mevius, D.J., Iredell, J.R., Bossers, A., Smith, H.E., Partridge, S.R. (2015). IncI shufflons: assembly issues in the next-generation sequencing era. *Plasmid*. 80:111-117. doi: 10.1016/j.plasmid.2015.04.009.

Carrattoli, A., Zankari, E., Garcia-Fernandez, A., Larsen, M.V., Lund, O., Villa, L., *et al*. (2014). *In Silico* detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob. Agents Chemother*. 58(7): 3895-3903. doi: 10.1128/AAC.02412-14.

Carattoli, A., Hasman, H. (2020). PlasmidFinder and In Silico pMLST: identification and typing of plasmid replicons in whole genome sequencing (WGS). *Methods Mol. Biol*. 2075: 285-295. doi: 10.1007/978-1-4939-9877-7_20.

Chhangawala, S., Rudy, G., Mason, C.E., Rosenfeld, J.A. (2015). The impact of read length on quantification of differentially expressed genes and splice junction detection. *Genome Biol*. 16(131). doi: 10.1186/s13059-015-0697-y.

Cusumano, C.K., Hung, C.S., Chen, S.L., Hultgren, S.J. (2010). Virulence plasmid harbored by uropathogenic *Escherichia coli* functions in acute stages of pathogenesis. *Infect Immun*. 78(4): 1457-1467. doi: 10.1128/IAI.01260-09.

Deamer, d., Akeson, M., Branton, D. .(2016). Three decades of nanopore sequencing. *Nat Biotechnol*. 34: 518-524. doi: 10.1038/nbt.3423.

De Coster, W., D'Hert, S., Schultz, D.T., Cruts, M., Van Broeckhoven, C. (2018). NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics*. 34(15): 2666-2669. doi: 10.1093/bioinformatics/bty149.

De Toro, M., Garcillan-Barcia, P., De La Cruz, F. (2014). Plasmid diversity and adaptation analyzed by massive sequencing of *Escherichia coli* plasmids. *Microbiol Spectrum* 2(6): PLAS-0031-2014. doi: 10.1128/microbiolspec.PLAS-0031-2014.

Fricke, W.F., McDermott, P.F., Mammel, M.K., Zhao, S. (2009). Antimicrobial resistance-conferring plasmids with similarity to virulence plasmids from avian pathogenic *Escherichia coli* strains in *Salmonella enterica* serovar Kentucky isolates from poultry. *Appl. Environ. Microbiol*. 75(18): 5963-5971. doi: 10.1128/AEM.00786-09.

George, S., Pankhurst, L., Hubbard, A., Votintseva, A., Stoesser, N., Sheppard, A.E., Mathers, A., Norris, R., *et al*. (2017). Resolving plasmid structures in *Enterobacteriaceae* using the MinION nanopore sequencer: assessment of MinION and MinION/Illumina hybrid data assembly approaches. *Microb. Genom*. 3. doi: 10.1099/mgen.0.000118.

Goldstein, S., Beka, L., Graf, J., Klassen, J.L. (2019). Evaluation of strategies for the assembly of diverse bacterial genomes using MinION long-read sequencing. *BMC Genom*. 20(23). doi: 10.1186/s12864-018-5381-7.

Gonzalez-Escalona, N., Allard, M.A., Brown, E.W., Sharma, S., Hoffman, M. (2019). Nanopore sequencing for fast determination of plasmids, phages, virulence markers, and antimicrobial resistance genes in Shiga toxin-producing *Escherichia coli*. *PLoS ONE* 14(7): e0220494. doi: 10.1371/journal.pone.0220494.

Grevskott, D.H., Salva-Serra, F., Moore, E.R.B., Marathe, N.P. (2020). Nanopore sequencing reveals genomic map of CTX-M-type extended-spectrum B-lactamases carried by *Escherichia coli* strains isolated from blue mussels (*Mytilus edulis*) in Norway. *BMC Microbiol* **20**: 134. doi: 10.1186/s12866-020-01821-8.

He, S., Hickman, A.B., Varani, A.M., Siguier, P., Chandler, M., Dekker, J.P., et al. (2015). Insertion sequence IS26 reorganizes plasmids in clinically isolated multidrug-resistant bacteria by replicative transposition. *mBio* 6: e00762-15. doi: 10.1128/mBio.00762-15.

Johnson, T.J., Wannemuehler, Y.M., Johnson, S.J., Logue, C.M., White, D.G., Doetkott, C., Nolan, L.K. (2007). Plasmid replicon typing of commensal and pathogenic *Escherichia coli* isolates. *Appl. Environ. Microbiol*. 73(6): 1976-1983. doi: 10.1128/AEM.02171-06.

Johnson, T.J., Nolan, L.K. (2009). Pathogenomics of the virulence plasmids of *Escherichia coli*. *Microbiol. Molec. Biol. Rev*. 73(4): 750-774. doi: 10.1128/MMBR.00015-09.

Johnson, T.J., Logue, C.M., Johnson, J.R., Koskowski, M.A., Sherwood, J.S., Barnes, H.J., DebRoy, C., Wannemuehler, Y.M., Obata-Yasuoka, M., Spanjaard, L., Nolan, L.K. (2012). Associations between multidrug resistance, plasmid content, and virulence potential among extraintestinal pathogenic and commensal *Escherichia coli* from humans and poultry. *Food. Pathog. And Dis*. 9(1). doi: 10.1089/fpd.2011/0961.

Kleinheinz, K.A., Joensen, K.G., Larsen, M.V. (2014). Applying the ResFinder and VirulenceFinder web-services for easy identification of acquired antibiotic resistance and *E. coli* virulence genes in bacteriophage and prophage nucleotide sequences. *Bacteriophage*. 4: e27943. doi: 10.4161/bact.27943.

Kologorov, M., Yuan, J., Lin, Y., Pevzner, P. (2019). Assembly of long error-prone reads using repeat graphs. *Nature Biotechnol*. doi: 10.1038/s41587-019-0072-8.

Kono, N., Arakawa, K. (2019). Nanopore sequencing: review of potential applications in functional genomics. *Dev Growth and Diff*. 61(5): 316-326. doi: 10.1111/dgd.12608.

Kwong, J.C., McCallum, N., Sintchenko, V., Howden, B.P. (2015). Whole genome sequencing in clinical and public health microbiology. *Molec Diagnostics Microbiol*. 47(3): 199-210. doi: 10.1097/PAT.0000000000000235.

Johnson, T.J., Wannemuehler, Y., Johnson, S.J., Stell, A.L., Doetkott, C., Johnson, J.R., Kim, K.S., Spanjaard, L., Nolan, L.K. (2008). Comparison of extraintestinal pathogenic *Escherichia coli* strains from human and avian sources reveals a mixed subset representing potential zoonotic pathogens. *Appl. Environ. Microbiol*. 74(22): 7043-7050.

Lemon, J.K., Khil, P.P., Frank, K.M., Dekker, J.P. (2017). Rapid Nanopore sequencing of plasmids and resistance gene detetion in clinical isolates. *J. Clin. Microbiol*. 55(12): 3530-3543. doi: 10.1128/JCM.01069-17.

Li, H., Handsaker, B., Wyosker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., *et al*. (2009). The sequence alignment/Map format and SAMtools. *Bioinformatics*. 25(16): 2078-2079. doi: 10.1093/bioinformatics/btp352.

Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*. 34(18): 3094-3100. doi: 10.1093/bioinformatics/bty191.

Li, R., Xie, M., Dong, N., Lin, D., Yang, X., Wong, M.H.Y., Chang, E.W., Chen, S. (2018). Efficient generation of complete sequences of MDR-encoding plasmids by rapid assembly of MinION barcoding sequencing data. *GigaSci*. 7(3). doi: 10.1093/gigascience/gix132.

Liao, Y., Cheng, H., Wu, H., Kuo, S., Lauderdale, T.Y., Chen, F. (2019). Completing circular bacterial genomes with assembly complexity by using a sampling strategy from a single MinION run with barcoding. *Front. Microbiol.* 4. doi: 10.3389/fmicb.2019.02068.

Lima, L., Marchet, C., Caboche, S., da Silva, C., Istsace, B., *et al.* (2019). Comparative assessment of long-read error correction software applied to Nanopore RNA-sequencing data. Briefings in Bioinformatics, *Oxford University Press* (OUP). pp.1-18. doi: 10.1093/bib/bbz058-hal-02394395.

Logue, C.M., Wannemuehler, Y., Nicholson, B.A., Doetkott, C., Barbieri, N.L., Nolan, L.K. (2017). Comparative analysis of phylogenetic assignment of human and avian ExPEC and fecal commensal *Escherichia coli* using the (previous and revised) Clermont phylogenetic typing methods and its impact on avian pathogenic *Escherichia coli* (APEC) classification. *Front. Microbiol.* 8:283. doi: 10.3389/fmicb.2017.00283.

Lohman, N.J., Pallen, M.J. (2015). Twenty years of bacterial genome sequencing. *Nat Rev Microbiol.* 13: 787-794. doi: 10.1038/nrmicro3565.

Mellata, M., Ameiss, K., Mo, H., Curtiss, R. (2010). Characterization of the contribution to virulence of three large plasmids of avian pathogenic *Escherichia coli* $_{X}$7122 (O78:K80:H9). *Infect Immun.* 78(4): 1528-41. doi: 10.1128/IAI.00981-09.

Mellata, M. (2013). Human and avian extraintestinal pathogenic *Escherichia coli*: infections, zoonotic risks, and antibiotic resistance trends. *Foodborn Pathog. Dis.* 10(11): 916-932. doi: 10.1089/fpd.2013.1533

Nicholson, B.A., West, A.C., Mangiamele, P., Barbieri, N., Wannemuehler, Y., Nolan, L.K., Logue, C.M., Li, G. (2016). Genetic characteristics of ExPEC-like virulence plasmids among a subset of NMEC. *PLoS ONE* 11(1): e0147757. doi: 10.1371/journal.pone.0147757.

Pitout, J.D.D. (2012). Extraintestinal pathogenic *Escherichia coli*: a combination of virulence with antibiotic resistance. *Front. Microbiol.* 19. doi: 10.3389/fmicb.2012.00009.

Platt, A.R., Woodhall, R.W., George, A.L. (2018). Improved DNA sequencing quality and efficiency using an optimized fast cycle sequencing protocol. *BioTechniques.* 43: 58-62. doi: 10.2144/000112499.

Rodriguez-Beltran, J., Hernandez-Beltran, J.C.R., DelaFuente, J., Escudero, J.A., Fuentes-Hernandez, A., MacLean, R.C., Pena-Miller, R., San Millan, A. (2018). Multicopy plasmids allow bacteria to escape from fitness trade-offs during evolutionary innovation. *Nat Ecol and Evol.* 2: 873-881. doi: 10.1038/s41559-018-0529-z.

Roosaare, M., Puustusmaa, M., Mols, M., Vaher, M., Remm, M. (2018). PlasmidSeeker: identification of known plasmids from bacterial whole genome sequencing reads. *PeerJ* 6:e4588. doi: 10.7717/peerj.4588.

Rozov, R., Kav, A.B., Bogumil, D., Shterzer, N., Halperin, E., Mizrahi, I., Shamir, R. (2017). Recycler: an algorithm for detecting plasmids from *de novo* assembly graphs. *BMC Bioinf*. 33(4): 475-482. doi: 10.1093/bioinformatics/btw651.

Sarowska, J., Futoma-Koloch, B., Jama-Kmiecik, A., Frej-Madrzak, M., Ksiazczyk, M., Bugla-Ploskonska, G., *et al*. (2019). Virulence factors, prevalence, and potential transmission of extraintestinal pathogenic *Escherichia coli* isolated from different sources: recent reports. *Gut Pathog*. 11(10). doi: 10.1186/s13099-019-0290-0.

Skyberg J.A., Johnson, T.J., Johnson, J.R., Clabots, C., Logue, C.M., Nolan, L.K. (2007). Acquisition of avian pathogenic *Escherichia coli* plasmids by a commensal *E. coli* isolate enhances its abilities to kill chicken embryos, grow in human urine, and colonize the murine kidney. *Molec Pathog*. 74(11): 6287-6292. doi: 10.1128/IAI.00363-06.

Sommerville, V., Lutz, S., Schmid, M., Frei, D., Moser, A., Irmler, S., Frey, J.E., Ahrens, C.H. (2019). Long-read based de novo assembly of low-complexity metagenome samples results in finished genomes and reveals insights into strain diversity and an active phage system. *BMC Microbiol*. 19(143). doi: 10.1186/s12866-019-1500-0.

Subedi, D., Kohli, G.S., Vijay, A.K., Willcox, M.D.P., Rice, S.A. (2019). Accessory genome contributes to the virulence and resistance of the ocular isolate of *Pseudomonas aeruginosa:* a complete genome analysis. *PLoS ONE* 14(4): 30215038. doi: 10.1371/journal.pone.0215038.

Taylor, T., Volkening, J.D., DeJesus, E., Simmons, M., Dimitrov, K.M., Tillman, G.E., *et al*. (2019). Rapid, multiplexed, whole genome and plasmid sequencing of foodborne pathogens using long-read nanopore technology. *Sci Reports* **9**: 16350. doi: 10.1038/s41598-019-5242-x.

Thompson, M.D., Aukema, K.G., O'Bryan, D.M., Rader, S.D., Murray, B.W. (2008). Plasmid sonication improves sequencing efficiency and quality in the Beckman Coulter CEQ system. *BioTechniques* 45: 327-329. doi: 10.2144/000112902.

Velappan, N., Sblattero, D., Chasteen, L., Pavlik, P., Bradbury, A.R.M. (2007). Plasmid incompatibility: more compatible than previously thought? *Prot. Engineer. Design Select*. 20(7): 309-313. doi: 10.1093/protein/gzm005.

Wawire, S.A., Reva, O.N., O'Brien, T.J., Figueroa, W., Dinda, V., Shivoga, W.A., Welch, M. (2021). Virulence and antimicrobial resistance genes are enriched in the plasmidome of clinical *Escherichia coli* isolates compared with wastewater isolates from western Kenya. *Infect. Genet. Evol*. 91: 104784. doi: 10.1016/meegid.2021.104784.

Weingarten, R.A., Johnson, R.C., Conlan, S., Ramsburg, A.M., Dekker, J.P., Lau, A.F., *et al*. 2018. Genomic analysis of hospital plumbing reveals diverse reservoir of bacterial plasmids conferring carbapenem resistance. *mBio*. 9(1): e02011-17. doi: 10.1128/mBio.02011-17.

Wick, R.R., Judd, L.M., Gorrie, C.L., Holt, K.M. (2017a). Completing bacterial genome assemblies with multiplex MinION sequencing. *Microb. Genom.* 3(10): e000132. doi: 10.1099/mgen.0.000132.

Wick, R.R., Judd, L.M., Gorrie, C.L., Holt, K.E. (2017b). Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comp. Biol.* 13(6): e1005595. doi: 10.1371/journal/pcbi.1005595.

Wick, R.R., Judd, L.M., Holt, K.E. (2018). Deepbinner: demultiplexing barcoded Oxford Nanopore reads with deep convolutional neural networks. *PLoS Comp. Biol.* 14(11): e1006583. doi: 10.1371/journal.pcbi.1006583.

Wick, R.R., Judd, L.M., Wyres, K.L., Holt, K.E. (2021). Recovery of small plasmid sequences via Oxford Nanopore sequencing. *Microb Genom.* 7(8). doi: 10.1099/mgen.0.000631.

Woksepp, H., Ryberg, A., Berglind, L., Schon, T., Soderman, J. (2017). Epidemiological characterization of a nosocomial outbreak of extended spectrum B-lactamase *Escherichia coli* ST-131 confirms the clinical value of core genome multilocus sequence typing. *APMIS* 125: 1117-1124. doi: 10.111/apm.12753.

Zanarki, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., *et al*. (2012). Identification of acquired antimicrobial resistance genes. *J. Antimicrob. Chemother.* 67(11): 2640-2644. doi: 10.1093/jac/dks261.

Chapter 4 is a co-first author manuscript written by me and Dr. Sivaranjani Murugesan, a postdoctoral fellow in the White Lab. I performed the *E. coli* isolation, curli and cellulose testing, the phenotypic AMR testing, and the crystal violet biofilm assays. Dr. Murugesan was responsible for growing and testing *E. coli* biofilms on polystyrene pegs and determining the MIC and MBC concentrations of planktonic and biofilm cells in the presence of commercial disinfectants.

CHAPTER 4: BIOFILM FORMATION AND RESISTANCE TRENDS IN *E. COLI* STRAINS ASSOCIATED WITH OUTBREAKS OF COLIBACILLOSIS ON SASKATCHEWAN BROILER FLOCKS

Submitting to Frontiers in Microbiology, November 2021

Murugesan Sivaranjani[a#], Madeline McCarthy[a,b#], Michelle Sniatynski[a,c], Linzhi Wu[a], Jo-Anne Dillon[a,b], Joseph E. Rubin[c] and Aaron P. White[a,b*]

[a] Vaccine and Infectious Disease Organization, Saskatoon, Saskatchewan, Canada.

[b] Department of Biochemistry, Microbiology and Immunology, University of Saskatchewan, Saskatoon, Saskatchewan, Canada.

[c] Department of Veterinary Microbiology, University of Saskatchewan, Saskatoon, Saskatchewan, Canada.

[#] these authors contributed equally

4.1 ABSTRACT

The global poultry industry has grown to the extent that the number of chickens on Earth now exceeds the number of humans. *Escherichia coli* infections in poultry cause significant morbidity and economic losses for producers each year. We screened 94 *E. coli* isolates from 12 colibacillosis outbreaks on Saskatchewan farms for antimicrobial resistance and biofilm formation. 56 isolates were from broilers with confirmed colibacillosis (systemic *E. coli*), and 38 isolates were from healthy broilers in the same flocks (cecal *E. coli*). Resistance to penicillins, tetracyclines, and aminoglycosides was common in isolates from all 12 outbreaks, while cephalosporin resistance varied by outbreak. Most *E. coli* were able to form biofilms in at least one of three growth media (1/2 TSB, M63, and BHI broth). In general, systemic *E. coli* strains were more drug-resistant and more likely to form biofilms in nutrient-rich media (BHI) than in minimal media (1/2 TSB and M63). However, on a strain-to-strain basis, we determined that there was no clear correlation between antimicrobial resistance and biofilm formation. The 21 strongest biofilm forming strains were equally comprised of drug resistant and susceptible isolates. We also tested the bactericidal efficacy of four different disinfectants against planktonic (single) cells and biofilm cells of both systemic and cecal isolates. All selected disinfectants effectively killed the planktonic cells at concentrations lower than the supplier recommended concentration, however the resistance of biofilm cells was 2- to 64-fold higher. The bactericidal efficacies of disinfectants Virocid and Virkon did not vary between the systemic and cecal biofilms, but there were notable differences between the strains treated with DDAC and H_2O_2. Our results speak to the complex lifestyles and evolutionary pressures leading to diversity in *E. coli* poultry isolates.

Keywords: Antibiotics, Antimicrobial Resistance, Avian Pathogenic *Escherichia coli,* Biofilm, Colibacillosis, Disinfectants, Planktonic cells, Poultry.

4.2 INTRODUCTION

Colibacillosis is an umbrella term for acute, systemic infections caused by avian pathogenic *Escherichia coli* (APEC) (Dziva and Stevens, 2008; Kemmett *et al.,* 2013; Gabiraba and Schouler, 2015). Colibacillosis can present as multiple local and systemic diseases including acute septicemia, air sacculitis, chronic respiratory diseases, cellulitis, pericarditis, peritonitis and salphingitis (Nolan *et al.,* 2020). Outbreaks among flocks cause significant economic losses in the poultry industry due to carcass condemnation, reduced egg laying, morbidity, and mortality, as well as costs incurred for disinfection and antimicrobial treatment (Dziva and Stevens, 2008; Mellata, 2013; Graveline *et al.,* 2014; Guabiraba and Schouler, 2015). Colibacillosis is thought to begin as a respiratory tract infection in broilers that inhale fecal dust contaminated with APEC (Dziva and Stevens, 2008; Guabiraba and Schouler, 2015). Due to a lack of consistent phenotypic and genotypic characteristics, the gold standard for APEC classification is the chicken lethality assay (Wooley *et al.,* 2000; Collingwood *et al.,* 2014; Mageiros *et al.,* 2021). APEC strains, which encompass a large number of *E. coli* sequence types, are one of the subpathotypes of ExPEC (Extraintestinal pathogenic *E. coli*) together with uropathogenic *E. coli* (UPEC), sepsis-associated *E. coli* (SEPEC) and neonatal meningitis *E. coli* (NMEC) (Johnson *et al.,* 2008). Several studies have suggested that APEC and human ExPEC strains share similar virulence-associated genes despite their sources of isolation, indicating the increased potential for zoonotic *E. coli* infections in humans (Johnson *et al.,* 2008; Jakobsen *et al.,* 2011). Although they cause significant damage to broiler flocks, little is known about the ecological niche of the APEC strains. Do they persist in the barn environment, or do they colonize broilers and later escape the intestinal niche? Many of the details surrounding colibacillosis infections remain a mystery.

The primary treatment option for Colibacillosis is antimicrobial therapy (Sargeant *et al.,* 2019). Many antimicrobials that are critical for human medicine are used in the global poultry industry to prevent and treat bacterial diseases (Landoni and Albarellos, 2015). Antimicrobials have been used for decades to enhance growth and prevent infection in food animals (Xiong *et al.,* 2018), but this overuse of antimicrobials can lead to the emergence of antimicrobial resistance (AMR) among bacterial pathogens. This is thought to be one of the most significant contributors to the current AMR crisis in human clinical settings (Cantas and Suer, 2014; Michael *et al.,* 2014; Medina *et al.,* 2020). *E. coli* are resident microbes in the intestinal flora of most food animals (Partridge *et al.,* 2018). In the gut, they are exposed to billions of other

bacteria and viruses that may harbor antimicrobial resistant mobile genetic elements. *E. coli* may be a significant reservoir of AMR on farms (Osman *et al.,* 2018; Duy *et al.,* 2020). For example, *E. coli* that possess extended-spectrum β-lactamases (ESBL) are resistant to penicillins and certain cephalosporins and because the resistance genes are located on mobile genetic elements, ESBL *E. coli* are now globally disseminated (Thaden *et al*., 2016;). Infections caused by ESBL *E. coli* are associated with higher economic costs, longer hospital stays, and increased mortality compared to non-ESBL *E. coli* (Tumbarello *et al.,* 2007; Malande *et al.,* 2019). Poultry are a known reservoir of ESBL *E. coli,* and transmission from livestock to humans has been demonstrated via multi-locus sequence typing (MLST) (Leverstein *et al.,* 2011; Overdevest *et al.,* 2011).

Most APEC strains are also likely to form biofilms, an important factor contributing to increased persistence and survival of APEC within broiler farms (Reisner *et al.,* 2006; Rodrigues *et al.,* 2019). Biofilms are a physiological state where bacterial cells become irreversibly attached to surfaces and enclosed in a self-secreted extracellular polymeric matrix composed of proteins, polysaccharides, and nucleic acids (Hung *et al*., 2013; Sanchez-Vizuete *et al.,* 2015). Growth in biofilms offers certain advantages to APEC including facilitating the exchange of AMR and virulence-associated genes via horizontal gene transfer between bacteria of same or different species (van den Bogaard and Stobberingh, 2000). In addition, the biofilm matrix provides protection for the embedded cells, resisting detrimental effects of antibiotics, disinfectants, and the host immune system, all of which makes it difficult to eradicate biofilm-related infections in clinical and food production settings (Reisner *et al.,* 2006; Kester and Fortune, 2014; Rodrigues *et al.,* 2019). Despite the increasing occurrence of AMR, bacteria seem to remain susceptible to disinfectants when used correctly. As most disinfectant products contain multiple active ingredients with different mechanisms of action, it is unlikely that resistance will develop when used according to manufacturer's recommendations (Davin-Regli and Pages, 2012). However, disinfectants are often used incorrectly in practice; the use of inappropriate concentrations, expired products, or the inactivation of compounds due to the presence of inorganic matter on insufficiently cleaned surfaces are common errors. The exposure of bacteria to sub-inhibitory concentrations may result in reduced susceptibility to disinfectants (Soumet *et al.,* 2016), which can cause significant problems for producers at multiple steps in the food production pipeline. Further, insufficient cleaning of barn surfaces or in meat packaging facilities may allow biofilm formation to occur on

equipment. It is very important to understand the lifestyle of APEC strains to prevent dissemination of disease in birds and potentially to humans. We believe that studying trends in biofilm formation and antimicrobial/disinfectant resistance in isolates from Saskatchewan farms will provide producers with more information to better manage outbreaks of disease.

We isolated 94 *E. coli* (56 from diseased birds and 38 from healthy birds) from 12 different colibacillosis outbreaks from Saskatchewan farms and analysed these isolates for biofilm formation, and resistance to 27 antimicrobial compounds and four disinfectants to ascertain whether AMR and biofilm characteristics differ among *E. coli* isolates from diseased birds (systemic *E.coli*) and cecal isolates from healthy birds (cecal *E. coli*).

4.3 MATERIALS AND METHODS

4.3.1 Sample Collection and *E. coli* Isolation and Identification

Local producers with farms within a 2-hr driving distance of Saskatoon submitted broilers with suspected colibacillosis to the Poultry Extension Team at the Western College of Veterinary Medicine, University of Saskatchewan. Each submission (i.e., typically three birds) was considered a new outbreak. Necropsy was performed and *E. coli* infection of the heart, liver, and spleen was confirmed by microbial isolation followed by MALDI-TOF mass spectrometry (Prairie Diagnostic Services; http://pdsinc.ca/). Portions of the heart, liver, and spleen of diseased birds were reserved and stored at 4°C until positive *E. coli* diagnosis. Organs were placed in 2 ml Eppendorf Safe-Lock tubes containing 1 ml of phosphate-buffered saline (PBS) and a 5-mm steel bead (Qiagen #69989) and were homogenized for 5 min at 30 Hz using a mixer mill (Retsch; MM400). Aliquots of the organ homogenates were spread on MacConkey Agar and colonies allowed to grow overnight at 37°C. Suspected *E. coli* colonies were confirmed via positive indole and negative Simmons-citrate biochemical tests. For each confirmed *E. coli* outbreak, 3-4 healthy broilers from the same flocks were obtained within a few days from the same producers. *E. coli* strains were isolated from the cecal contents of each healthy bird, following the same procedure as described above. In total, 12 outbreaks over a two-year period yielded 94 *E. coli* isolates, consisting of 56 disease-causing or "systemic" isolates and 38 healthy or "cecal" isolates. Isolates with visually distinct colony morphologies were selected for further analysis. In the end, each isolate was considered distinct,

even if recovered from the same outbreak, due to having a unique combination of biofilm and AMR profiles.

4.3.2 Bacterial Growth and Media Conditions

Bacterial isolates were stored in 20% glycerol at -80°C for long-term preservation. Prior to each experiment, isolates were inoculated from frozen stocks onto Tryptic soy agar (TSA; Becton Dickinson) and were incubated for 18 h at 37°C. A single colony was inoculated into 5 mL of LB broth and grown overnight at 37°C with shaking at 200 rpm. Overnight cultures were used for all subsequent experiments.

4.3.3 Curli and Cellulose Testing

Overnight cultures were standardized to an optical density of 1.0 at 600 nm. Four microlitre aliquots were spotted on Congo Red Agar (20 ug/mL of Congo Red dye in 1% tryptone, 0.5% yeast extract, 1.5% agar; pH 7.2-7.4), to visualize the rdar morphotype and curli production (PMID: 1677357 - Collinson et al., 1991) and onto Calcofluor White agar (200mg/L fluorescent brightener #28 (Sigma-Aldrich) in 1% tryptone, 0.5% yeast extract, 1.5% agar) to visualize cellulose production (PMID: 11929533 – Solano et al).

4.3.4 Crystal Violet Assay

Crystal violet assays were performed using the method of George O-Toole *et al*., 2005), with details outlined previously (Liu et al., 2018). For each isolate, 20 μL of standardised inoculum was used to inoculate wells of three 96-well plates, containing 180 μL of three biofilm media: 1) 1/2 tryptic soy broth (TSB), 2) M63 media (ammonium sulfate, potassium monophosphate, and ferrous sulfate), and 3) brain heart infusion (BHI) broth. Each isolate was inoculated into six replicate wells and grown at 28°C for 24 hours; negative control wells contained 180uL of broth and were not inoculated. After 24 hours, liquid was gently removed from the wells and each plate was washed twice with 200 μL of PBS. Bacterial cells and extracellular biofilm matrix were fixed by adding 200 μL of 99% methanol to each well and incubating at room temperature for 15 minutes. The methanol was removed, and plates were left to air dry for 30 minutes, prior to staining with 200 μL of 0.2% crystal violet for 5 minutes. Excess crystal violet was gently rinsed from each well using dH_2O. Plates were left to air dry before addition of 160 μL of 33% glacial acetic acid. The

absorbance of the resulting solution in each well was measured at 570 nm using a spectrophotometer. The classification system developed by Stepanovic *et al*. was used to describe biofilm formation by each strain (2008). The cut-off OD (OD_c) was calculated as three standard deviations above the average OD of the negative control wells; this value was compared to the mean OD for each strain. The biofilm categories were: strong (mean OD > 4*OD_c); moderate (4*OD_c > mean OD > 2* OD_c); weak (2* OD_c > mean OD > OD_c); or none (OD_c > mean OD).

4.3.5 Biofilm Formation in the MBEC Assay® Biofilm Inoculator Plates

Biofilm formation and the measurement of disinfectant susceptibility of APEC biofilms were performed using the MBEC Assay® biofilm inoculator 96-well plates (Innovotech Inc., Alberta, Canada) according to previously described method (Harrison *et al.,* 2010). A standardized inoculum of 1.5 x 10^7 CFU/mL was added to biofilm growth media to a final volume of 150 µL per well. The MBEC (minimum biofilm eradication concentration) plate lid containing 96 polystyrene pegs was placed into the 96-well microtitre plate base, sealed with parafilm and incubated at 37°C for 24 h with slight rocking on a tilting platform shaker. After incubation, peg lids were removed from the base plate and washed twice with sterile PBS to remove loosely attached cells. Biofilm formation on pegs was quantified by viable cell counts following disruption of the biofilm by sonication for 30 min with a bath sonicator (Branson #3510, Canada). Dislodged biofilm cells were serially diluted and grown on Mueller Hinton (MH; Becton Dickinson) agar plates (Liu *et al.*, 2018; Sivaranjani *et al.,* 2021).

4.3.6 Determining the Minimum Inhibitory Concentration (MIC) and Minimum Bactericidal Concentration (MBC) for disinfectants

The MIC was determined by broth microdilution according to the Clinical and Laboratory Standards Institute (CLSI) standard for antimicrobial susceptibility testing (CLSI, 2020). Serial two-fold dilutions of each disinfectant (**Table B2**) were prepared in MH broth to a final volume of 90 µL per well. 10 µL aliquots of culture, representing 1.5 x 10^7 bacterial cells were added to each appropriate well containing MH broth with disinfectant, and the plates were covered and incubated at 37°C with slight rocking. After 24 h of growth, the OD at 600 nm of each well was measured by an xMarkTM Microplate Absorbance Spectrophotometer (Bio-Rad). The MIC was recorded as the concentration in the wells where bacterial growth was visibly inhibited. To determine the MBC, 100 µl of aliquots from MIC wells were sub-cultured on MH agar and incubated for 24 h. The

lowest concentration that resulted in no bacterial growth on MH agar plates was determined as the MBC.

4.3.7 Disinfectant Susceptibility Testing

(i) Planktonic cells

Cells from overnight cultures were washed twice with sterile PBS and resuspended in 5 mL of fresh Mueller Hinton (MH) broth. The MH broth cultures were normalized to an optical density of 1.0 at 600 nm and the starting concentration of bacteria (CFU/mL) for each culture before disinfectant challenge was determined by viable cell counts (**Table B2**). Serial two-fold dilutions of each disinfectant were prepared in MH broth with a final volume of 90 µL in each well of a 96-well microtitre plate. 10 µL of inoculum was added to 90 µL of MH broth containing disinfectant and incubated at 37°C for 30 min. To test the viability of cells within each well, 10 µL aliquots were removed and sub-cultured into fresh MH broth and the remaining 90 µL was spread on a fresh MH agar plate. Both the sub-cultures were incubated at 37°C for 24 h and observed for growth after 24 h. The lowest concentration resulting in no regrowth on both the liquid and agar plates was recorded as the bactericidal concentration of disinfectant against planktonic cells from each strain.

(ii) Biofilm cells

After 24 h growth, biofilms formed on polystyrene pegs were transferred to a 96-well microtitre plate in which serial double dilutions of disinfectants were prepared in MH broth. Prior to disinfectant challenge, 6 control pegs were broken off from each plate, sonicated for 30 min in recovery media and the resuspended cells were serially diluted to determine the average starting number of biofilm cells for each strain. The remaining pegs were exposed to disinfectant for 30 min, rinsed twice in sterile PBS and cells were dislodged by bath sonication as described above. For each disinfectant concentration, cells from six replicate pegs were serially diluted and grown on MH agar plates to determine the number of viable cells remaining. The bactericidal effect of disinfectants on biofilm cells was determined as the lowest concentration of disinfectant resulting in viable cell counts at or below the detection limit of 125 CFU per mL.

4.3.8 Antimicrobial Susceptibility Testing

Antimicrobial susceptibility of all strains was determined via broth microdilution using the Gram-negative panel for the MicroScan system (Beckman Coulter), testing the following 24 agents: amikacin, amoxicillin-clavulanic acid, ampicillin-sulbactam, ampicillin, aztreonam, cefazolin, cefepime, cefotaxime, cefoxitin, ceftazidime, ceftriaxone, cefuroxime, ciprofloxacin, doripenem, ertapenem, gentamicin, imipenem, levofloxacin, meropenem, piperacillin-tazobactam, tetracycline, tigecycline, tobramycin, trimethoprim-sulfamethoxazole, cephalothin, and nitrofurantoin. The results were interpreted according to the Clinical and Laboratory Standards Institute Guidelines (2020). MIC breakpoints were as follows: ampicillin (≥32 μg/mL), cefazolin (≥8 μg/mL), cephalothin (≥32 μg/mL), ceftazidime (≥16 μg/mL), ceftriaxone (≥4 μg/mL), cefuroxime (≥32 μg/mL), cefotaxime (≥4 μg/mL), cefepime (≥16 μg/mL), amoxicillin + clavulanic acid (≥32 μg/mL), ampicillin + sulbactam (≥32 μg/mL), cefoxitin (≥32 μg/mL), aztreonam (≥16 μg/mL), gentamicin (≥8 μg/mL), tobramycin (≥16 μg/mL), nalidixic acid (≥64 μg/mL), tetracycline (≥16 μg/mL), colistin (≥4 μg/mL), chloramphenicol (≥32 μg/mL), ciprofloxacin (≥1 μg/mL), doripenem (≥4 μg/mL), ertapenem (≥2 μg/mL), imipenem (≥4 μg/mL), meropenem (≥4 μg/mL), nitrofurantoin (≥128 μg/mL).

Chloramphenicol, Colistin, and Nalidixic acid testing was performed using agar dilution (Wiegand *et al*., 2008). For each antibiotic, petri plates were prepared containing LB agar (lysogeny broth, 1.5% agar) plus antibiotic ranging from 0.5 ug/mL to 32 μg/mL. Isolates were streaked on blood agar (Oxoid #CM0055) and incubated for 18 hours at 37°C. A 0.5 MacFarland standard was made for each isolate and isolates were spotted onto antibiotic plates using a 96-pin microplate replicator. Plates were incubated at 35°C for 24 hours before reading MIC values. For quality control, *E. coli* ATCC 25922, *Staphylococcus aureus* ATCC 29213 and *Pseudomonas aeruginosa* ATCC 27853 were included as appropriate.

4.3.9 Statistics

Graphing of data and statistical analyses were performed using GraphPad Prism software v.8.0.2 (San Diego, CA, USA). Data from the disinfectant susceptibility assays were logarithmically transformed and analysed using the Kruskal-Wallis test with post-hoc analysis via Dunn's multiple tests to compare the log differences between the disinfectant treated and untreated control groups. Statistical differences in biofilm biomass, as measured by crystal violet staining, were determined using an ordinary-one way ANOVA test [$F = 14.28$, $p > 0.0001$]. Distribution of

systemic isolates was analyzed using a chi-square test to determine whether the frequency of positive biofilm formation in each media was statistically significant [chi-square statistic: 15.74, df=2, p>0.0001].

4.4 RESULTS

4.4.1 Antimicrobial resistance is common in avian *E. coli* isolated from broilers

A total of 94 *E. coli* were isolated from birds with colibacillosis (n = 56; systemic *E. coli*) and from presumed healthy birds of the same flock (n = 38; cecal *E. coli*). These *E. coli* represented 12 outbreaks on Saskatchewan farms between 2019 and 2020. Antimicrobial testing showed that 32 43 isolates were susceptible to all 27 tested antimicrobials, while 32 were resistant to three or more drugs. Of these, nine isolates were resistant to 9 or more drugs from six different classes (**Table 4.1**). Resistance to Tetracycline (29.8%), Ampicillin (24.6%), Gentamicin (23.4%), and Tobramycin (19.1%) was the most observed. No resistance to colistin, chloramphenicol, or the carbapenems was detected.

4.4.2 Systemic *E. coli* are more drug-resistant than cecal *E. coli*

Systemic isolates were resistant to a significantly greater number of antimicrobials than cecal isolates: 71% of systemic *E. coli* were resistant to at least one drugs compared to 29% of cecal *E. coli* (X^2, $p < 0.001$). 71% of cecal isolates were susceptible to all tested antimicrobials. Tetracycline resistance was common in both groups, but resistance to ampicillin, gentamycin, and tobramycin was 20-30% higher in systemic isolates (**Table 4.1**). Systemic isolates also had higher MIC values for first and third generation cephalosporins, compared to cecal isolates (**Table 4.1**). Resistance to cefoxitin was found exclusively in systemic isolates, while Cefepime resistance was exclusively identified in cecal isolates.

Table 4.1. Minimum inhibitory concentration (MIC) distribution for systemic (n=56) and cecal (n=38) *E. coli* isolates cultured from Saskatchewan broilers.

Drug	Isolate Type	0.5	1	2	4	8	16	32	64	128	% Resistant	MIC 50	MIC 90
Ampicillin	Systemic			19	17			20			35.7	4	32
	Cecal			18	16			4			10.5	4	32
Cefazolin	Systemic			42	7	1		6			10.7	2	32
	Cecal			35							7.9	2	2
Cephalothin	Systemic					35	11	10			17.8	8	32
	Cecal					27	8	3			7.9	8	8
Ceftazidime	Systemic			52				4			7.1	2	32
	Cecal			38							0	2	2
Ceftriaxone	Systemic		51			1		4			8.9	1	8
	Cecal		35					3			7.9	1	1
Cefuroxime	Systemic				49	3		4			7.1	4	32
	Cecal				31	4		3			7.9	4	8
Cefotaxime	Systemic			50	2	3	1				8.9	2	2
	Cecal			35	3						7.9	2	2
Cefepime	Systemic			56							0	2	2
	Cecal			35			3				7.9	2	2
Amox/Clav	Systemic					50		6			10.7	4	32
	Cecal					37		1			2.6	8	8
Amp/Sul	Systemic				36		11	9			16.1	4	32
	Cecal			36			1	1			2.6	1	1
Cefoxitin	Systemic			43	9			4			7.1	2	4
	Cecal			25	13						0	2	4
Aztreonam	Systemic			51	1	3	1				1.8	2	4
	Cecal			31	4		3				7.9	2	4
Gentamicin	Systemic		28	8	1		19				33.9	2	16
	Cecal		18	8	8		4				10.5	2	16
Tobramycin	Systemic		13	24	2	2	15				26.8	2	16
	Cecal		9	21	4	1	3				7.9	2	8
Nalidixic Acid	Systemic		1	29	18	8					0	2	8
	Cecal			14	13	9			2		5.2	4	8
SXT	Systemic	49			7						12.5	0.5	4
	Cecal	38									0	0.5	0.5
Tetracycline	Systemic			37				19			33.9	2	32
	Cecal			30				8			21.1	2	32
Colistin	Systemic		54	2							0	1	1
	Cecal		34	4							0	1	2
Nalidixic Acid	Systemic			25	20	12			1		1.8%	4	8
	Cecal			15	13	9			1		2.6%	4	8
Chloramphenicol	Systemic				30	26					0	2	4

	Cecal			12	26			0	4	4

Cells corresponding to concentrations tested are not shaded and resistance breakpoints are marked with dark bars. The number of isolates inhibited at each concentration is noted in each cell. Isolates not inhibited by the highest concentration of each drug are included in the next concentration above the highest tested concentration (SXT = trimethoprim-sulfamexazole).

4.4.3 Systemic isolates form greater biofilms in rich media

All *E. coli* isolates were screened for biofilm-forming ability in three different liquid media: ½ TSB, M63, and BHI broth. Crystal violet staining was used to quantify and classify the level of biofilm. To determine whether the origin of the isolates played a role in biofilm formation, we compared systemic and cecal isolates (**Fig 4.1**). Overall, the systemic isolates generated significantly more biomass in nutrient rich BHI broth, as compared to the other two media (**Fig 4.1A**). In contrast, cecal isolates did not show a significant preference for any one media type (**Fig 4.1B**).

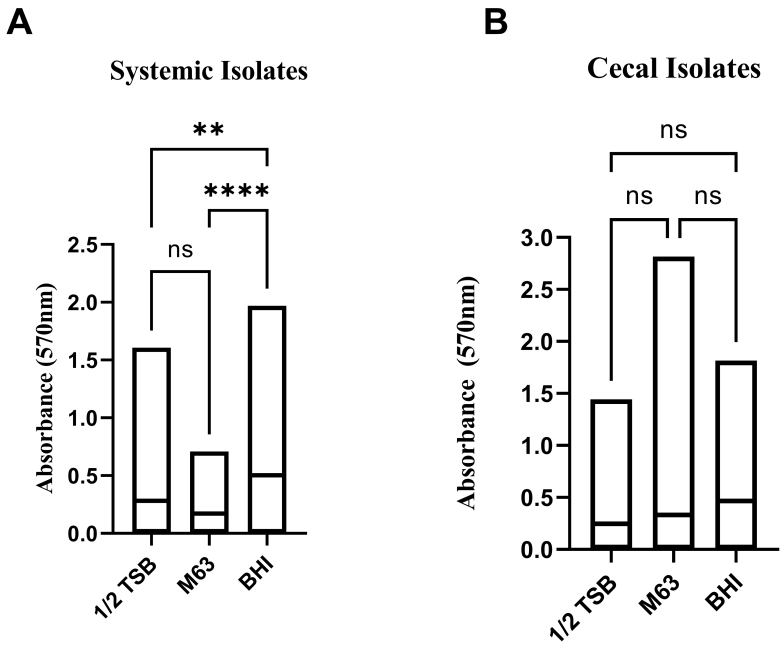

Figure 4.1. Quantitation of biofilm biomass for systemic and cecal *E. coli* isolates grown in three different growth media. Bars represent the mean absorbance at 570 nm of crystal violet that was used to stain biofilm cells and extracellular matrix from 56 systemic *E. coli* (**A**) or 38 cecal *E. coli* (**B**). Statistical significance between different growth media was tested by one-way ANOVA (****, $p < 0.0001$; **, $p < 0.01$; ns, $p > 0.05$).

To classify the level of biofilm formation for individual isolates, we used the formula described by Stepanovic et al. (2008) (see Materials and Methods). Isolates were categorized as positive if they formed strong or moderate biofilms, while isolates that produced none or weak biofilms were categorized as negative. Overall, a significantly greater number of total isolates (63.8%) formed positive biofilms in nutrient rich BHI broth (X^2, $p < 0.005$) (**Fig 4.2**). 56% of systemic isolates formed strong biofilms in BHI broth, which was significant compared to positive biofilm formation in ½ TSB and M63 media (X^2, $p < 0.005$) (**Fig 4.2A**). Cecal isolates also formed the greatest biofilms in BHI broth: almost 35% produced strong biofilms (**Fig 4.2B**). However, unlike systemic isolates, cecal isolates did not form significantly greater biofilms in BHI broth when compared to nutrient-poor media. Systemic isolates formed poor biofilms in minimal media: 71.4% were negative for biofilm in ½ TSB broth and 76.8% were negative for biofilm in M63 broth (**Fig 4.2A**). In contrast, only 27.8% of systemic strains were negative for biofilm formation in BHI. A roughly equal proportion of cecal isolates (52.6% in M63 broth and 55.3% in BHI broth) formed positive biofilms in rich and minimal media, suggesting that these strains are equally capable of forming biofilms in minimal and nutrient-rich conditions (**Fig 4.2**). Approximately 30% of cecal isolates formed no biofilm in each medium tested (**Fig 4.2**). In comparison, in M63 and BHI broth less than 20% of systemic isolates formed no biofilm.

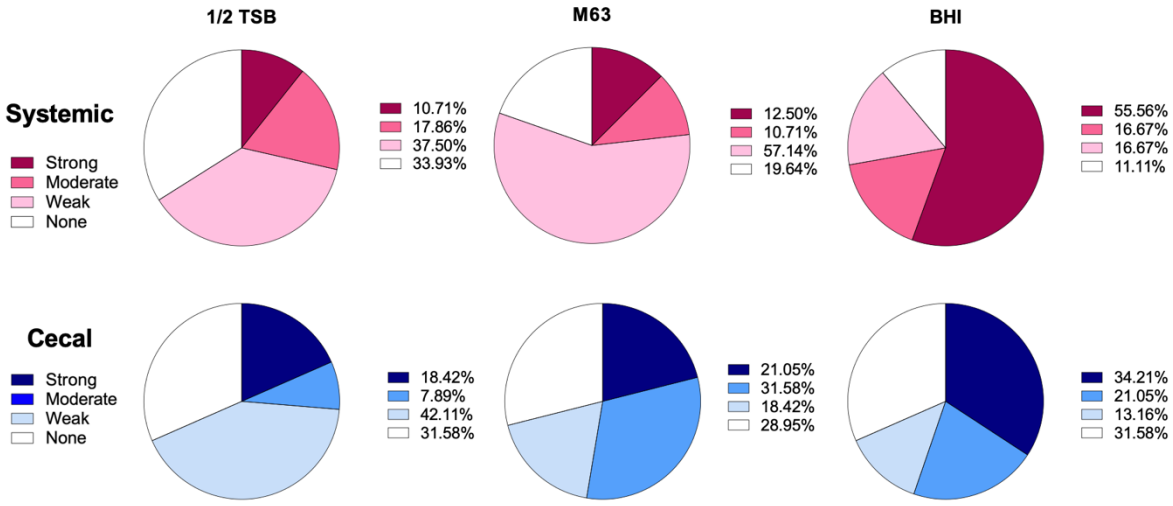

Figure 4.2. Quantification of Systemic (n=56) and Cecal (n=38) *E. coli* biofilm formation in 96 well plates in ½ TSB, M63, and BHI broth. Systemic isolates are represented by shades of red/pink while cecal isolates are blue. Biofilm classification was measured using the formula by Stepanovic et al (30) where the cut-off OD (OD_c) of the negative control wells (three standard deviations above the average OD of the negative control wells) was compared to the mean OD for each strain. Strong = mean OD > 4*OD_c; Moderate = 4*OD_c > mean OD > 2* OD_c); Weak = 2* OD_c > mean OD > OD_c; or None = OD_c > mean OD.

We also screened the 94 *E. coli* isolates for curli and cellulose production since these polymers are central to biofilm formation in the host and the environment (Chuppaya et al., 2019; MacKenzie et al., 2017 review; White et al. 2011). 62 isolates produced both curli and cellulose, distributed as 73% (41/56) of systemic isolates and 55% (21/38) of cecal isolates (**Table 4.2**). 20 isolates produced only curli, consisting of 10 systemic and 10 cecal isolates. Nine isolates produced only cellulose, consisting of five systemic and four cecal isolates. Only five isolates were negative for both biofilm polymers and all were cecal *E. coli*.

Table 4.2. Cellulose and curli production by systemic and cecal *E. coli* isolates.

	CURLI+	CELLULOSE+	BOTH	NEITHER
SYSTEMIC (56)	10	5	41	0
CECAL (38)	10	4	21	5
TOTAL	20	9	62	5

4.4.4 Identifying the strongest biofilm forming *E. coli* isolates and testing their resistance to commercial disinfectants.

We performed additional screening of 54 isolates (37 systemic and 17 cecal) that were classified as strong biofilm formers by crystal violet staining. We tested their ability to adhere and form biofilms on polystyrene pegs in MBEC biofilm inoculator plates. Rather than measure biomass by staining, we focused on live cell counts. 21 isolates (11 systemic + 10 cecal) formed biofilms with cell counts greater than 10^7 CFU (**Fig 4.3**; stars), which we considered an appropriate cut off for good biofilm formation, based on previous MBEC assays (Liu et al., 2018).

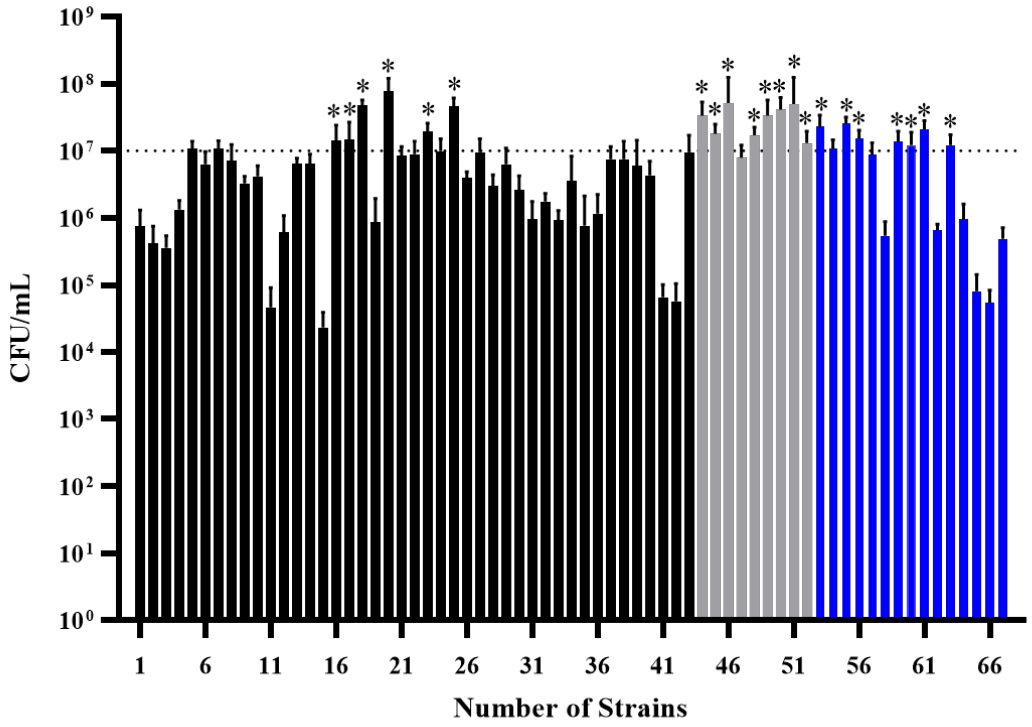

Figure 4.3. Screening of 54 strong biofilm producing *E. coli* isolates for their ability to form biofilms on polystyrene pegs. Strains that formed strong biofilms in more than one medium were tested in both media types. Cells were dislodged and enumerated from each peg (n = 6) after biofilm growth for 24 h in BHI (Black bar), M63 (Grey bar) or ½ TSB (Blue bar) media. Isolates that formed biofilms at a cell density > 10^7 CFU/mL were considered the best biofilm formers (21 strains denoted by star (*) are strong biofilm formers).

The 21 strongest biofilm-forming strains consisted of 11 systemic isolates and 10 cecal isolates (**Table 4.3**). 9/11 systemic isolates were resistant to one or more antimicrobials, while all the cecal isolates were susceptible to all 27 antimicrobials.

Table 4.3. Preferred biofilm growth media, disease status, and antimicrobial resistance phenotype of 21 *E. coli* strains that formed biofilms with cell counts greater than 10^7 CFU/mL.

Media	Strain	Disease Status	AMR Phenotype
½ TSB	9226-2S1	Systemic	AmoxKclav/AmpSulbactam/Ampicillin/Aztreonam/ Cefazolin/Cefotaxime/Cefoxitin/Ceftazidime/Cefriaxone/ Cefuroxime/Gentamicin/Tet/Tobramycin
	9226-2L1	Systemic	Amp Sulbactam/ Ampicillin/Gentamicin/Tobramycin
	6245-C1	Cecal	Susceptible to all
	4957-2L3	Systemic	Susceptible to all
	6245-C2	Cecal	Susceptible to all
	6245-C5	Cecal	Susceptible to all
	9226-3L1	Systemic	AmoxKclav/AmpSulbactam/Ampicillin/Aztreonam/ Cefazolin/Cefotaxime/Cefoxitin/Ceftazidime/Cefriaxone/ Cefuroxime/Gentamicin/ Tobramycin
	9226-3H1	Systemic	AmoxKclav/AmpSulbactam/Ampicillin/Aztreonam/ Cefazolin/Cefotaxime/Cefoxitin/Ceftazidime/Cefriaxone/ Cefuroxime/Gentamicin/Tet/Tobramycin
	9226-2H2	Systemic	AmoxKclav/AmpSulbactam/Ampicillin/Aztreonam/ Cefazolin/Cefotaxime/Cefoxitin/Ceftazidime/Cefriaxone/ Cefuroxime/Gentamicin/Tet/Tobramycin
M63	9619-3L1	Systemic	Susceptible to all
	0012-C1	Cecal	Susceptible to all
	0012-C5	Cecal	Susceptible to all
	9413-1S2	Systemic	Amp Sulbactam/Ampicillin/Gentamycin/Tobramycin
	6041-C2	Cecal	Susceptible to all
	6041-C6	Cecal	Susceptible to all
	6041-C9	Cecal	Susceptible to all
	6041-3L1	Systemic	AmoxKclav/AmpSulbactam/Ampicillin/Aztreonam/ Cefazolin/Cefotaxime/Cefoxitin/Ceftazidime/Cefriaxone/ Cefuroxime/Gentamicin/Tet/Tobramycin
BHI	6245-2H3	Systemic	Amp Sulbactam/Ampicillin/Gentamicin/Tobramycin
	6245-C1	Cecal	Susceptible to all
	6245-C2	Cecal	Susceptible to all
	4957-C5	Cecal	Susceptible to all
	4957-C3	Cecal	Susceptible to all
	9619-1H2	Systemic	Amp Sulbactam/Ampicillin/Tet/Trimethoprim+Sulfamexazone
	6041-3L1	Systemic	AmoxKclav/AmpSulbactam/Ampicillin/Aztreonam/ Cefazolin/Cefotaxime/Cefoxitin/Ceftazidime/Cefriaxone/ Cefuroxime/Gentamicin/Tet/Tobramycin

We screened 12 of the strongest biofilm-forming isolates (7 systemic + 5 cecal) that were selected from the 21 isolates identified in Fig. 4, for resistance to four commercial disinfectants. Each disinfectant was able to kill planktonic cells of all *E. coli* isolates at low concentrations, ranging from 0.00156%-0.003125% for Virocid, 0.125%-0.25% for Virkon, 1.88ppm-3.75ppm for DDAC and 0.03125%-0.0625% for H_2O_2 (**Table 4.4**). The MBC values of all four disinfectants were either equal or one serial dilution higher than the corresponding MIC values (**Table 4.4**). However, we deemed these results to be misleading because MIC and MBC determination requires exposure to test antimicrobials for 24 h, whereas most common disinfectants require contact times of only minutes to achieve bactericidal effects.

Table 4.4. Determination of MIC and MBC as recommended by CLSI guidelines.

S.No	Strain ID	Virocid (%)		Virkon™ (%)		DDAC (ppm)		H_2O_2 (%)	
		MIC	MBC	MIC	MBC	MIC	MBC	MIC	MBC
Systemic *E.coli* isolates									
1	9226-B3H1	0.003	0.003	0.25	0.25	3.75	7.5	0.0625	0.125
2	4957-B2L3	0.0016	0.003	0.125	0.25	3.75	3.75	0.03	0.03
3	6245-B2H3	0.003	0.006	0.25	0.25	1.88	3.75	0.0625	0.0625
4	9619-H12	0.003	0.006	0.25	0.25	3.75	3.75	0.0625	0.0625
5	6041-B3L1	0.003	0.006	0.25	0.25	3.75	7.5	0.03	0.03
6	9619-L31	0.003	0.003	0.25	0.25	3.75	3.75	0.0625	0.0625
7	9413-B1S2	0.003	0.003	0.25	0.25	3.75	7.5	0.03125	0.0625
Cecal *E.coli* isolates									
8	6245-C2	0.003	0.003	0.25	0.25	3.75	3.75	0.0625	0.0625
9	4957-C5	0.003	0.003	0.25	0.25	3.75	3.75	0.0625	0.0625
10	6245-C5	0.003	0.003	0.25	0.25	3.75	3.75	0.0625	0.0625
11	0012-C1	0.003	0.003	0.25	0.25	3.75	3.75	0.03	0.0625
12	6041-C2	0.003	0.003	0.25	0.25	3.75	3.75	0.0625	0.125

We repeated the disinfectant screening on planktonic and biofilm cells of each of the 12 *E. coli* strains with a 30 min exposure. The starting cell numbers for each strain ranged between 10^6 and 10^8 CFU (**Table B3**). Virkon and Virocid disinfectants exhibited effective bactericidal activity against both cell types of all 12 isolates even at low concentrations (**Table 4.5; Figures B3 and B4**). The concentration of Virocid required to achieve full bactericidal effect was 1/2- to 1/32- fold lower than the suppliers' recommended concentration, and for Virkon it was reduced approximately by 1/2- to 1/8- fold (**Table 4.5**). The quaternary ammonium disinfectant DDAC killed planktonic cells of the 12 *E. coli* isolates with bactericidal concentrations ranging from 4.7 to 9.4 ppm (**Table 4.4**). However, biofilm cells of all 12 isolates had 2- to 64-fold increased resistance towards DDAC, and one of the systemic isolates was not killed even at the manufacturer recommended concentration of 300 ppm (**Table 4.5; Figure B5**). Higher concentrations of DDAC were required to eradicate the biofilms of 3 systemic isolates as compared to the five cecal isolates (**Table 4.5**). The results with hydrogen peroxide were similar to DDAC, as planktonic cells of all tested isolates were susceptible to low concentrations, but biofilms cells had 4- to 64-fold increased resistance (**Table 4.5; Figure B5**). Again, 4 systemic isolates were more resistant and required 4% H_2O_2 for the complete eradication of biofilms, whereas all five cecal isolates were susceptible to concentrations of 2% or less (**Table 4.5; Figure B6**).

Table 4.5. Bactericidal effect of disinfectants on planktonic and biofilm cells after 30 min of contact time

S.No	Strain ID	Bactericidal effect of disinfectants concentrations on planktonic cells (PC) and biofilm cells (BC)							
		Virocid %		Virkon (%)		DDAC (ppm)		H_2O_2 (%)	
		PC	BC	PC	BC	PC	BC	PC	BC
Systemic *E. coli* isolates									
1	9226-B3H1	0.0156	0.125	0.25	0.5	9.4	300	0.0625	2
2	4957-B2L3	0.008	0.008	0.25	0.5	9.4	37.5	0.0625	0.5
3	6245-B2H3	0.0156	0.125	0.25	0.5	9.4	300[#]	0.0625	4
4	9619-H12	0.008	0.008	0.25	0.5	9.4	18.75	0.0625	4
5	6041-B3L1	0.0156	0.03125	0.25	0.5	9.4	75	0.0625	4
6	9619-L31	0.0156	0.125	0.25	0.5	9.4	75	0.0625	4
7	9413-B1S2	0.008	0.0156	0.25	0.25	9.4	150	0.0625	2
Cecal *E. coli* isolates									
8	6245-C2	0.0156	0.0625	0.25	0.5	9.4	75	0.0625	1
9	4957-C5	0.0156	0.0156	0.25	0.5	4.7	18.75	0.0625	0.25
10	6245-C5	0.0156	0.03125	0.25	0.5	9.4	75	0.0625	1
11	0012-C1	0.008	0.0156	0.25	0.5	9.4	75	0.0625	0.5
12	6041-C2	0.0156	0.0156	0.25	0.25	9.4	75	0.0625	2

[#] The minimum recommended concentration was not enough to achieve the bactericidal effect on the biofilm cells.

4.4.5 Trends in Antimicrobial Resistance by Outbreak

As a final comparison of the *E. coli* strains analyzed, we divided each group of isolates from outbreak to outbreak. This perhaps has more biological relevance than comparing the large groups of systemic and cecal strains to each other, since in almost all cases outbreaks occurred on different poultry farms. We identified 20 unique antibiotic resistance profiles across 12 Saskatchewan colibacillosis outbreaks (**Table 4.6**). The most common resistance pattern was ampicillin + gentamicin + tobramycin, which was identified in isolates from 7 of 12 outbreaks. In all but three outbreaks, we identified systemic *E. coli* isolates that were resistant to at least four drugs. While the cecal isolates were generally susceptible, we identified at least one drug resistant cecal isolate in half of the outbreaks. In two outbreaks, cecal isolates were identified that were resistant to more than four drugs. This comparison again revealed an overall trend for systemic *E. coli* isolates to be resistant to a greater number of antibiotics than cecal isolates, but there was a lot of diversity and no absolute patterns.

Table 4.6. AMR profiles for systemic and cecal *E. coli* isolated from colibacillosis outbreaks on Saskatchewan broiler farms.

OUTBREAK	#ISOLATES (Systemic/ Cecal)	AMR Profiles* SYSTEMIC	AMR Profiles* CECAL
2019-1	3 S; 4 C	**3** Tet	**3** Amp Sulbactam/Ampicillin/Aztreonam/Cefazolin/Cefepime/ Cefotaxime/Ceftriaxone/Cefuroxime/Tetracycline **1** Susceptible
2019-2	4 S; 5 C	**1** Aztreonam **3** Susceptible	**1** Tet **4** Susceptible
2019-3	5 S; 3 C	**1** Amp Sulbactam/ Ampicillin/Tet/Trimeth&Sulfa **1** Amox K Clav/Amp Sulbactam/Ampicillin/Aztreonam/Cefazolin/Cefotaxime/Cefoxitin/Ceftazidime/ceftriaxone /Cefuroxime/Gentamicin/Tetracycline/Tobramycin **1** Gentamicin/Tet/Tobramycin **2** Susceptible	**3** Susceptible
2019-4	3 S; 3 C	**1** Amp Sulbactam/Ampicillin/Tet/Trimeth&Sulfa **2** Susceptible	**3** Susceptible
2019-5	7 S; 3 C	**4** Amox Kclav/Amp Sulbactam/Ampicillin/Aztreonam/Cefazolin/Cefotaxime/Cefoxitin/ Ceftazidime/Ceftriaxone/Cefuroxime/Gentamicin/Tet/Tobramycin **2** Amp Sulbactam/Ampicillin/Gentamicin/Tobramycin **1** Tet	**1** Amox Kclav/Amp Sulbactam/Ampicillin/Pip-Tazo **2** Susceptible
2019-6	8 S; 5 C	**4** Tet **3** Amp Sulbactam/Ampicillin/Gentamicin/Tobramycin **1** Trimeth&Sulfa	**1** Gentamicin/Tobramycin/Nitrofurantoin **1** Tet **3** Susceptible
2019-7	5 S; 2 C	**2** Amp Sulbactam/Ampicillin/Gentamicin/Tobramycin **1** Gentamicin/Tet **1** Tet **1** Susceptible	**1** Aztreonam **1** Susceptible
2019-8	4 S; 3 C	**1** Amox Kclav/Amp Sulbactam/Ampicillin/Aztreonam/Cefazolin/Cefotaxime/Cefoxitin/Ceftazidime/ Ceftriaxone/Cefuroxime/Gentamicin/Tobramycin **1** Amp Sulbactam/Ampicillin/Gentamicin/Tobramycin **1** Gentamicin/Tobramycin **1** Trimeth&Sulfa	**3** Susceptible
2020-1	4 S; 3 C	**1** Amp Sulbactam/Ampicillin/Tet/Trimeth&Sulfa **1** Amox Kclav/Amp Sulbactam/Ampicillin/Tet/Trimeth&Sulfa **2** Susceptible to all	**3** Susceptible
2020-2	3 S; 3 C	**3** Susceptible	**3** Gentamicin/Tet/Tobramycin
2020-3	6 S; 4 C	**2** Amp Sulbactam/Ampicillin/Gentamicin/Tet/Tobramycin **1** Nitrofuranotin **3** Susceptible	**4** Susceptible
2020-4	5 S	**4** Amp Sulbactam/Ampicillin/Gentamicin/Tobramycin **1** Susceptible	

*The number in bold listed before each antimicrobial resistance profile correlates to the number of isolates with that profile.

4.5 DISCUSSION

In this study, we analyzed 94 *E. coli* isolates that were obtained from 12 different colibacillosis outbreaks that occurred on Saskatchewan broiler farms during a 2-year period. The comparison between 56 systemic or disease-causing *E. coli* strains and 38 cecal isolates from presumed healthy birds in the same flocks revealed an overall trend where the systemic isolates had a greater degree of AMR and were more likely to form biofilms than cecal isolates. This trend was not absolute and there were exceptions depending on the outbreak. The patterns of AMR and biofilm formation tended to be unique within each outbreak and between systemic and cecal isolates within each outbreak. Overall, there was a lot of phenotypic diversity (Schouler *et al*., 2012; Clermont *et al*., 2013; Riley, 2014; Cordoni *et al*., 2016). The differences observed between the systemic and cecal *E. coli* isolates indicated that they represent distinct groups of *E. coli* within the broiler farm environment. If the cecal *E. coli* were the source of disease, we would have expected to identify conserved AMR patterns within each outbreak, but we did not. The significance of the differences between systemic and cecal *E. coli* are difficult to understand in the context of *E. coli* biology without knowing the exact pressures that these groups of *E. coli* are subject to (Reisner *et al*., 2006; Weissman *et al*., 2006; Bruzuskiewicz *et al*., 2009; Coombes, 2009; Rodriguez-Beltran *et al*., 2015; Azevedo *et al*., 2016; O'Boyle and Roe, 2021).

One of the goals for performing phenotypic analysis was to determine if there was a correlation between the level of AMR and biofilm-forming ability of individual *E. coli* strains, regardless of source. Despite the trend for systemic isolates having more biofilm formation and higher levels of antibiotic resistance, there was no statistically significant relationship between AMR status and biofilm production in either the systemic or cecal group of *E. coli* strains analyzed. A roughly equal proportion of systemic and cecal isolates formed strong biofilms despite systemic isolates being more drug resistant. Nine of 11 systemic isolates that formed the strongest biofilms, in 96-well plates and on polystyrene pegs, were multi-drug resistant, while the remaining two systemic isolates and 10 cecal isolates that formed strong biofilms were susceptible to all tested antimicrobials. This indicated that there was no direct correlation between the two phenotypes.

Despite biofilm formation not being correlated with increased AMR, it is well established that this growth state is important for other aspects of persistence and survival (Harms *et al*., 2016; Jung *et al*., 2019; Yan and Bassler, 2019). The biofilm environment offers bacteria a source of nutrients as well as protection from external stressors such as desiccation, antimicrobials, and disinfectants. In our study, systemic isolates formed significantly greater biofilms in BHI broth than in any other media, with 72% forming a positive biofilm. In contrast, less than 25% of systemic strains formed a positive biofilm in relatively nutrient-poor media. The significance of the media type is not fully understood, however, this preference for nutrient-rich media and disease-causing *E. coli* isolates has been shown before. For cecal isolates, there was no significant preference for the type of biofilm media, with approximately 50% of cecal isolates producing a biofilm in two media: M63 and BHI broth. This is consistent with previous reports that *E. coli* isolates of fecal origin can form biofilms in both minimal and nutrient-rich media (Skyberg *et al*., 2007; Nielsen *et al*., 2018). Previous studies have shown that APEC strains form negligible biofilms in most media (Skyberg *et al*., 2007; Nielsen *et al*., 2018). However, the strains characterized in those studies were obtained from geographically diverse collections of APEC strains isolated from birds with various manifestations of disease. To our knowledge, our study is the first to analyze APEC strains isolated exclusively from birds with confirmed sepsis in a small geographic area. This may explain why a significant number of our systemic strains formed biofilms in nutrient-rich media. Perhaps *E. coli* strains responsible for a majority of systemic colibacillosis cases in Saskatchewan outbreaks possess genetic characteristics that favour biofilm formation in nutrient-rich media (Mellata *et al*., 2010; Mellata *et al*., 2012).

Cleaning with disinfectants is one the most common, practical intervention strategies to eliminate bacterial contamination in the poultry barn setting. Many disinfectants are used in the poultry industry, with active ingredients belonging to aldehydes (e.g. formaldehyde), halogens (e.g. sodium hypochlorite), surface active agents (e.g. quaternary ammonium compounds [QAC]), oxidizing agents (e.g. H_2O_2), phenols and alcohols (Sagripanti and Bonifacino, 2000). The continuous exposure of residual disinfectants at the lower concentration may increase the bacterial tolerance (Wieland *et al.,* 2017). Any increase in tolerance could increase the bacterial adaptive resistance to antibiotics and enhance their survival fitness against various environmental stresses (Ortega Morente *et al.,* 2013). We used the Calgary biofilm device to evaluate the killing efficacy of four commonly used disinfectants on the adherent population of cells from 12 of the

strongest biofilm forming isolates. Each of the selected disinfectants is formulated with single or multiple active ingredients, in which KleenGlow (DDAC) and Virocid have QAC in common with the remaining active ingredients' varying from each other.Virkon formulated with multiple ingredients includes peroxygens, surfactants and other ingredients. It kills bacteria by oxidising sulfur bonds in proteins and enzymes and resulting in cell wall rupture. Virocid is also a broad-spectrum disinfectant, formulated with synergistic blend of two different QAC (i.e., single chain and twin chain quaternary ammonia) along with glutaraldehyde and isopropanol. QACs interacts with negatively charged cell membrane to elicit bactericidal action, while glutaraldehyde synergistically interacts with functional thiol and amine groups of proteins (Gilbert and Moore, 2005; Masadeh *et al*., 2013). It is believed that the interaction of active components with matrix proteins may facilitate the effective penetration of biofilm and synergistically kills the embedded bacterial cells (Osland *et al*., 2020). Our results showed that Virocid and Virkon had excellent killing at low concentrations against the planktonic and biofilm cells of all test strains, presumably due to the synergistic blend of multiple ingredients with non-specific bactericidal mode of action. In contrast, DDAC and H_2O_2 showed a large gap in killing where up to 64-fold higher concentrations were necessary to eradicate biofilm cells. The germicidal action of DDAC depends on numerous factors including the length of the N-alkyl chain, their combination with other active ingredients, pH, concentration used and the different target pathogens (He *et al*., 2013; Gilbert and Moore, 2005).

H_2O_2 is a strong oxidizing agent that has been shown to damage bacterial DNA, proteins, and cellular membranes (Bell *et al*., 1997). It is well-known that bacteria residing in the inner matrix of a biofilm are protected from bactericidal concentrations of chemical agents, providing prime opportunities to develop resistance. Several studies have implicated that the sub-inhibitory concentrations of QACs may have a potential to select the emergence of AMR among pathogens, which could be raised from cross-resistance between QACs and other medically important antibiotics such as ampicillin, ceftazidime, cefotaxime, and ciprofloxacin (Nhung *et al.,* 2015, Soumet *et al.* 2016; Nasr *et al.* 2018). Wieland *et al.* reported that the regular use of DDAC at low residual concentration would be able to select antibiotic resistance among *E. coli* and *Enterococcus* spp., which was confirmed by correlating the increased MIC values of DDAC with increased MIC values for several antibiotics, and high-level resistance was observed against aminoglycoside in enterococci (2017). These reports are in line with our results, that most of our

isolates showed relatively common resistance to aminoglycoside and several other antibiotics. There was also a clear difference in susceptibility between systemic and cecal biofilms, as systemic isolates required higher concentration of DDAC and H_2O_2 for complete killing of biofilm cells. Our results suggest that the disinfectants with multiple ingredients (Virocid and Virkon), with different mode of action are more effective against both systemic and cecal biofilms than the disinfectants with single mode of action (DDAC and H_2O_2).

Resistance to therapeutic antibiotics among *E. coli* isolated from poultry is of serious concern for both human and veterinary medicine. Recently, Canada has restricted the use of medically important antibiotics in food animals from being used as a growth promotor and requiring producers to obtain a veterinary prescription to treat infections in their flocks as an attempt to reduce the AMR among pathogens (Government of Canada, 2021). However, studies suggest that resistant bacteria can persist in the environment and spread AMR genes to commensal bacteria or between premises even with declining or no antimicrobial usage (Ge *et al.*, 2005; Agunos *et al*., 2013). Reasons for this fitness advantage are likely to be complex but not limited to horizontal transmission of AMR genes between bacteria, improper cleaning and disinfection, biofilm formation and co-selection of resistance to certain antimicrobials. 71% of systemic isolates were resistant to between one and nine tested antimicrobial compounds. In contrast, only 29% of *E. coli* recovered from the cecal contents of uninfected birds were drug resistant or multi-drug resistant (MDR). Previous studies have demonstrated that resistance to specific drugs is correlated with the presence of ExPEC virulence genes (Johnson *et al*., 2006; Johnson *et al*., 2012). Resistance to trimethoprim-sulfamexazole was exclusive to systemic isolates – this drug is significantly associated with the possession of the *afa* gene (Johnson *et al*., 2012). Afa is a fimbrial adhesin which is found in uropathogenic *E. coli* (UPEC) and is associated with recurrent infection (Bien *et al*., 2011). Most *E. coli* strains tested in our study demonstrated resistance to Ampicillin, Tetracycline, Gentamycin and Tobramycin, and displayed MDR as classified by resistance to as many as three different antimicrobial classes. These resistance trends are comparable to those previous reports, with most clinical *E. coli* isolates recovered from either diseased or uninfected chickens exhibiting resistance to aminoglycosides, tetracyclines and sulfa drugs (Zhao *et al*., 2005; Sáenz *et al.,* 2003, Yang *et al.,* 2004; Varga *et al.,* 2019).

AMR profiles appeared in multiple outbreaks, suggesting that many of the antimicrobial resistance genes identified may be plasmid-mediated genes circulating in broiler barns in central Saskatchewan. For example, *E. coli* isolates with the same 9 drug resistance profile were isolated from four outbreaks, while resistance to ampicillin, gentamicin, and tobramycin was present in over 50% of outbreaks. This may be indicative of transmission between farms, though further genetic characterization would be required to confirm this. Given the small geographic location of our study population and only two hatchling suppliers servicing central Saskatchewan, vertical transmission of APEC may also play a role in some of the similarities observed between outbreaks (Giovanardi *et al.,* 2005). These systemic isolates may establish biofilms in the broiler barn environment and act as reservoirs for virulence and antimicrobial resistance genes in future outbreaks. *E. coli* can survive for years in the barn environment, particularly in dust (Schulz *et al.*, 2016), perhaps supporting the prevailing hypothesis that colibacillosis infections occur via the inhalation of contaminated dust (Lutful Kabir, 2010). We also found that *E. coli* from Saskatchewan colibacillosis outbreaks had outbreak-specific media preferences for biofilm formation, suggesting that outbreak isolates may share properties that influence biofilm formation in a particular media type (Skyberg *et al.,* 2007; Nielsen *et al.,* 2018). In some cases, systemic and cecal isolates from a single outbreak exclusively formed biofilms in only a single medium. However, analysis of more outbreaks and a greater number of isolates is required to see if any trends are consistent. While biofilm preference was not correlated with any particular AMR profile, previous studies have identified biofilm formation as a key tool for the persistence of antimicrobial resistance genes in broiler barns over (Zhai *et al.,* 2020). The genetic components that promote biofilm formation in different media types may be linked to survival in the presence of different disinfecting agents (Zhai *et al.,* 2020). A survey of the cleaning and disinfecting strategies used by Saskatchewan producers would help us to understand the role that biofilm formation plays in the persistence of outbreak-specific *E. coli* populations.

Proper biosecurity is critical for the management of colibacillosis on broiler farms. While biofilm formation and antimicrobial resistance are not directly correlated, both likely play a role in the survival and pathogenicity of systemic isolates on farms. Outbreak tracking may provide better information about the population of *E. coli* causing colibacillosis in a particular area. Bacterial persistence in the environment plays a key role in other recurrent bacterial infections, and tracking

outbreaks makes it possible to identify *E. coli* reservoirs (Pereira da Silva and Pereira De Martinis, 2012; Wibmann *et al*., 2021).

4.6 References

Agunos, A., L´eger, D., Avery, B. P., Parmley, E. J., Deckert, A., Carson, C.A., & Dutil, L. (2013). Ciprofloxacin-resistant Campylobacter spp. in retail chicken, western Canada. *Emerg Infect Dis*. 19(7), 1121–1124.

Azevedo, M., Sousa, A., Moura de Sousa, J., Thompson, J.A., Proenca, J.T., Gordo, I. (2016). Trade-offs of *Escherichia coli* adaptation to an intracellular lifestyle in macrophages. *PLoS ONE* 11(1): e0146123. doi 10.1371/journal.pone.0146123.

Bell, K. Y., Cutter, C.N., Sumner, S.S. (1997). Reduction of foodborne microorganisms on beef carcass tissue using acetic acid, sodium bicarbonate and hydrogen peroxide spray washes. *Food Microbiol*. 14:439–448.

Bien, J., Sokolova, O., Bozko, P. (2012). Role of uropathogenic *Escherichia coli* virulence factors in development of urinary tract infection and kidney damage. *Int. J. Nephrol*. 2012: 681473. doi: 10.1155/2012/681473.

Brzuskiewicz, E., Gottschalk, G., Ron, E., Hacker, J., Dobrindt, U. (2009). Adaptation of pathogenic *E. coli* to various niches: genome flexibility is key. *Microb Pathog*. 6: 110-125. doi: 10.1159/000235766.

Cantas, L., Suer, K. (2014). Review: the important bacterial zoonoses in "One Health" concept. *Front Public Health* 2:144. doi: 10.3389/fpubh.2014.001144.

Ceri, H., Olson, M. E., Stremick, C., Read, R. R., Morck, D., & Buret, A. (1999). The Calgary Biofilm Device: new technology for rapid determination of antibiotic susceptibilities of bacterial biofilms. *J Clin Microbiol*. 37(6), 1771-1776.

Chuppava, B., Keller, B., El-Wahab, A., Surie, C., Visscher, C. (2019). Resistance reservoirs and multi-drug resistance of commensal *Escherichia coli* from excreta and manure isolated in broiler houses with different flooring designs. Front. Microbiol. 10:2633. doi: 10.3389/fmicb.2019.02633.

Clinical and Laboratory Standards Institute. Performance standards for antimicrobial susceptibility testing. *CLSI Document M100-30*. Wayne, PA: CLSI; 2020.

Clermont, O., Christensen, J.K., Denamur, E., Gordon, D.M. (2013). The Clermont *Escherichia coli* phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. *Environ. Microbiol. Rep*. 5(1): 58-65.

Collingwood, C., Kemmett, K., Williams, N., Wigley, P. (2014). Is the concept of avian pathogenic Escherichia coli as a single pathotype fundamentally flawed? Front. Vet. Sci. 1:5. doi:10.3389/fvets.2014.00005.

Coombes, B.K. (2009). Type III secretion systems in symbiotic adaptation of pathogenic and non-pathogenic bacteria. *Trends Microbiol*. 17(3): 89-94. doi: 10.1016/j.tim.2008.11.006.

Cordoni, G., Woodward, M.J., Wu, H., Alanazi, M., Wallis, T., La Ragione, R.M. (2016). Comparative genomics of European avian pathogenic *E. coli* (APEC). *BMC Genomics*. 17, 960. doi: 10.1186/s12864-016-3289-7.

Davin-Regli, A., Pagès, J. (2019) Cross-resistance between biocides and antimicrobials: an emerging question. *Rev Sci Tech*. 31(1):89–104.

Duy, P.T., Nguyen, T.N.T., Thuy, D.V., The, H.C., Alcock, F., Boinett, C., *et al*. (2020). Commensal *Escherichia coli* are a reservoir for the transfer of XDR plasmids into epidemic fluoroquinolone resistant *Shigella sonnei*. *Nat Microbiol*. 5: 256-264. doi: 10.1038/s41564-019-0645-9.

Dziva, F., Stevens, M.P. (2008). Colibacillosis in poultry: unravelling the molecular basis of virulence of avian pathogenic Escherichia coli in their natural hosts. Avian Pathol. 37(4):355-66. doi: 10.1080/03079450802216652.

Ge, B., McDermott, P.F., White, D.G., Meng, J. (2005). Role of efflux pumps and topoisomerase mutations in fluoroquinolone resistance in Campylobacter jejuni and Campylobacter coli. *Antimicrob Agents Chemother*. 49:3347–54.

Gilbert, P., Moore, L.E. (2005) Cationic antiseptics: Diversity of action under a common epithet. *J Appl Microbiol*. 99, 703–715.

Government of Canada. (2021). Responsible use of medically important antimicrobials in animals.<https://www.canada.ca/en/public-health/services/antibiotic-antimicrobial-resistance/animals/actions/responsible-use-antimicrobials.html> [Accessed 20 July 2021].

Graveline, R., Garneau, P., Martin, C., Mourez, M., Hancock, M.A., Lavoie, R., Harel, J. (2014), Leucine- responsive regulatory protein Lrp and PapI homologues in116uence phase variation of CS31Ambriae. *J Bacteriol* 196(16): 2944-2953.

Guabiraba, R., Schouler, C. (2015). Avian colibacillosis: still many black holes. FEMS Microbiol. Lett. 362(15): fnv118. doi: 10.1093/femsle/fnv118.

He, J., Soderling, E., Vallittu, P.K., Lassila, L.V. (2013). Investigation of double bond conversion, mechanical properties, and antibacterial activity of dental resins with different alkyl chain length quaternary ammonium methacrylate monomers (QAM). *J Biomater Sci Polym Ed*. 24:565–573.

Hung, C., Zhou, Y., Pinker, J.S., Dodson, K.W., Crowley, J.R., Heuser, *et al*. (2013). *Escherichia coli* biofilms have an organized and complex extracellular matrix structure. mBio. 4(5). doi: 10.1128.mBio.00645-13.

Jakobsen, L., Garneau, P., Kurbasic, A., Bruant, G., Stegger, M., Harel, J., Jensen, K.S., Brousseau, R., Hammerum, A.M., Frimodt-Moller, N. (2011). Microarray-based detection of extended virulence and antimicrobial resistance gene profiles in phylogroup B2 Escherichia coli of human, meat and animal origin. *J Med Microbiol* 60:1502–1511.

Johnson, T.K., Wannemuehler, Y., Johnson, S.J., Stell, A.L., Doetkott, C., Johnson, J.R., *et al*. (2008). Comparison of extraintestinal pathogenic *Escherichia coli* strains from human and avian sources reveals a mixed subset representing potential zoonotic pathogens. Appl. Environ. Microbiol. 74(22). doi: 10.1128/AEM.01395-08.

Kemmett, K., Humphrey, T., Rushton, S., Close, A., Wigley, P., Williams, N.J. (2013). A longitudinal study simultaneously exploring the carriage of APEC virulence associated genes and the molecular epidemiology of faecal and systemic *E. coli* in commercial broiler chickens. PLoS One. 8(6): e67749. doi: 10.1371/journal.pone.0067749.

Kester, J.C., Fortune, S.M. (2014). Persisters and beyond: mechanisms of phenotypic drug resistance and drug tolerance in bacteria. *Crit Rev Biochem Mol Biol* 49:91-101.

Landoni, M.F., Albarellos, G. (2015). The use of antimicrobial agents in broiler chickens. *Vet J* 205:21–7.

Liu, F., Hansra, S., Crockford, G., Koster, W., Allen, B.J., Blondeau, J.M., Lainesse, C., White, A.P. (2018). Tetrasodium EDTA is effective at eradicating biofilms formed by clinically relevant microorganisms from patients' central venous catheters. *mSphere*. 3(6): e00525-18. doi: 10.1128/mSphere.00525-18.

Ludden, C., Raven, K.E., Jamrozy, D., Gouliouris, T., Blane, B., Coll, F., Goffau, M. *et al*. (2019). One Health genomic surveillance of *Escherichia coli* demonstrates distinct lineages and mobile genetic elements in isolates from humans versus livestock. *mBio* 10(1). doi: 10.1128/mBio.02693-18.

Mageiros, L., Meric, G., Bayliss, S.C., *et al*. (2021). Genome evolution and the emergence of pathogenicity in avian *Escherichia coli*. Nat. Commun. 12:765. doi: 10.1038/s41467-021-20988-w.

Malande, O.O., Nuttal, J., Pillay, V., Bamford, C., Eley, B. (2019). A ten-year review of ESBL and non-ESBL *Echerichia coli* bloodstream infections among children at a tertiary referral hospital in South Africa. *PLoS ONE* 14(9): e0222675. doi: 10.1371/journal. pone0222675.

Masadeh, M.M., Gharaibeh, S.F., Alzoubi, K.H., Al-Azzam, S.I., Obeidat, W.M. (2013). Antimicrobial activity of common mouthwash solutions on multidrug-resistance bacterial biofilms. *J. Clin. Med. Res.* 5, 389–394.

Mellata, M., Ameiss, K., Mo, H., Curtiss, R. (2010). Characterization of the contribution to virulence of three large plasmids of avian pathogenic *Escherichia coli* $_{X}$7122 (O78:K80:H9). *Infect Immun*. 78(4): 1528-41. doi: 10.1128/IAI.00981-09.

Mellata, M., Maddux, J.T., Nam, T., Thomson, N., Hauser, H., Stevens, M.P., Mukhopadhyay, S., *et al*. (2012). New insights into the bacterial fitness-associated mechanisms revealed by the characterization of large plasmids of an avian pathogenic *E. coli*. *PLoS ONE*. 7(1): 329481. doi: 10.1371/journal.pone.0029483.

Mellata, M. (2013). Human and avian extraintestinal pathogenic Escherichia coli: infections, zoonotic risks, and antibiotic resistance trends. Foodborne Pathog. Dis. 10(11):916-32. doi: 10.1089/fpd.2013.1533.

Mellata, M., Ameiss, K., Mo, H., Curtiss, R. (2020). Characterization of the contribution to virulence of three large plasmids of avian pathogenic Escheria coli X7122 (O78:K80:H9). Infect. Immun. 78(4). doi: 10.1128/IAI.00981-09.

Merritt, J.H., Kadouri, D.E., O'Toole, G.A. (2005). Growing and analyzing static biofilms. *Curr. Prot. Microbiol*. doi: 10.1002/9780471729295.mc01b01s00.

Michael, C.A., Dominey-Howes, D., Labbate, M. (2014). The antimicrobial resistance crisis: causes, consequences, and management. *Front Public Health* 2: 145. doi: 10.3389/fpubh.2014.00145.

Nasr, A.M., Mostafa. S., Mostafa, H., Arnaout, H., Elshimy, A.A. (2018). The effect of exposure to sub-inhibitory concentrations of hypochlorite and quaternary ammonium compounds on antimicrobial susceptibility of Pseudomonas aeruginosa. *Amer J Infect Control* 46 (7):e57–e63.

Nhung, N. T., Thuy, C. T., Trung, N. V., Campbell, J., Baker, S., Thwaites, G., Carrique-Mas, J. (2015). Induction of antimicrobial resistance in Escherichia coli and non-typhoidal Salmonella strains after adaptation to disinfectant commonly used on farms in Vietnam. *Antibiotics* 4(4), 480-494.

Nielsen, D.W., Klimavicz, J.S., Cavender, T., Wannemuehler, Y., Barbieri, N.L., Logue, C.M., *et al*. (2018). The impact of media, phylogenetic classification, and *E. coli* pathotypes on biofilm formation in extraintestinal and commensal *E. coli* from humans and animals. Front. Microbiol. 9:902. doi: 10.3389/fmicb.2018.00902.

Nolan, L.K., Vaillancourt, J.P., Barbieri, N.L., Logue, C.M. (2020). Colibacillosis in: Diseases of Poultry, D.E. Swayne (Ed.), Wiley Online Library, Ames, IA, pp. 770- 830.

Ortega Morente, E., Fernandez-Fuentes, M.A., Grande Burgos, M.J., Abriouel, H., Pérez Pulido, R., Galvez, A. (2013). Biocide tolerance in bacteria. *Int. J. Food Microbiol*. 162:13–25.

Osland, A.M., Vestby, L.K., & Nesse, L.L. (2020). The Effect of Disinfectants on Quinolone Resistant E. coli (QREC) in Biofilm. *Microorganisms* 8(11), 1831.

Osman, K.M., Kappell, A.D., Elhadidy, M., ElMougy, F., El-Ghany, W.A.A., Orabi, A., *et al*. (2018). Poultry hatcheries as potential reservoirs for antimicrobial-resistant *Escherichia coli*: a risk to public health and food safety. *Nat Sci Rep*. 8(5859). doi: 10.1038/s41598-018-23962-7.

Partridge, S.R., Kwong, S.M., Firth, N., Jensen, S.O. (2018). Mobile genetic elements associated with antimicrobial resistane. *Clin. Microbiol. Rev*. 31(4): e00088-17. doi: 10.1128/CMR.00088-17.

Reisner, A., Krogfelt, K.A., Klein, B.M., Zechner, E.L., Molin, S. (2006). In vitro biofilm formation of commensal and pathogenic *Escherichia coli* strains: impact of environmental and genetic factors. J. Bacteriol. 188(10):3572-81. doi: 10.1128/JB.188.10.3572-3581.2006.

Rodrigues, S.V., Laviniki, V., Borges, K.A., Furian, T.Q., Moraes, H.L.S., *et al*. (2019). Biofilm formation by avian pathogenic Escherichia coli is not related to in vivo pathogenicity. *Curr. Microbiol*. 72(6): 194-199. doi: 10.1007/s00284-018-1608-8.

Romling, U. (2005). Characterization of the rdar morphotype, a multicellular behaviour in Enterobacteriaceae. Cell Mol. Life Sci. 62(11):1234-46. doi: 10.1007/s00018-005-4557-x.

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. *J Antimicrob Chemo*. 51(4), 1001-1005.

Sagripanti, J.L., & Bonifacino, A. (2000). Resistance of Pseudomonas aeruginosa to liquid disinfectants on contaminated surfaces before formation of biofilms. *J AOAC Intl*. 83(6), 1415-1422.

Sanchez-Vizuete, P., Orgaz, B., Aymerich, S., Le Coq, D., Briandet, R. (2015). Pathogens protection against the action of disinfectants in multispecies biofilms. Front. Microbiol. 6:705. doi: 10.3389/fmicb.2015.00705.

Sargeant, J., Bergevin, M., Churchill, K., Dawkins, K., Deb, B., Dunn, J., *et al*. (2019). The efficacy of antibiotics to control colibacillosis in broiler poultry: a systematic review. Animal Health Res. Reviews. 20(2): 263-273. doi: 10.1017/S146625319000264.

Schouler, C., Schaeffer, B., Bree, A., Mora, A,. Dahbi, G., Biet, F., Oswald, E., Mainil, J., Blanco, J., Moulin-Schouler, M. (2012). Diagnostic strategy for identifying avian pathogenic *Escherichia coli* based on four patterns of virulence genes. *J. Clin. Microbiol*. 50(5): 1673-78. doi: 10.1128/JCM.05057-11.

Skyberg, J.A., Siek, K.E., Doetkott, C., Nolan, L.K. (2007). Biofilm formation by avian *Escherichia coli* in relation to media, source and phylogeny. J. Appl. Microbiol. doi: 10.1111/j.1365-2672.206.03076.x

Soumet, C., Méheust, D., Pissavin, C., Le, G.P., Frémaux, B., Freurer, C., *et al*. (2016). Reduced susceptibilities to biocides and resistance to antibiotics in food associated bacteria following exposure to quaternary ammonium compounds. *J Appl Microbiol*. 121:1275–81.

Stepanovic, S., Vukovic, D., Savic, D.B., Svabic-Vlahovic, M. (2008). A modified microtiter-plate test for quantification of staphylococcal biofilm formation. J. Microbiol. Methods. 40(2): 175-9. doi: 10.1016/s0167-7012(00)00122-6.

Thaden, J.T., Fowler, V.G., Sexton, D.J., Anderson, D.J. (2016). Increasing incidence of extended-spectrum B-lactamase-producing *Escherichia coli* in community hospitals throughout the Southeastern United States. *Infect Control Hosp. Epidemiol*. 37:49-54. doi: 10.1017/ice.2015.239.

Tumbarello, M., Sanguinetti, M., Montuori, E., Trecarichi, E.M., Posteraro, B., Fiori, B. *et al*. (2007). Predictors of mortality in patients with bloodstream infections caused by extended-spectrum B-lactamase-producing *Enterobacteriaceae:* importance of inadequate initial antimicrobial treatment. *Antimicrob. Agents Chemother*. 51: 1987-1994. doi: 10.1128/AAC.01509-06.

van den Bogaard, A.E., Stobberingh, E.E. (2000). Epidemiology of resistance to antibiotics: links between animals and humans. *Int J Antimicrob Agents* 14:327–335.

Varga, C., Guerin, M. T., Brash, M. L., Slavic, D., Boerlin, P., & Susta, L. (2019). Antimicrobial resistance in fecal Escherichia coli and Salmonella enterica isolates: a two-year prospective study of small poultry flocks in Ontario, Canada. *BMC Vet Res* 15(1), 1-10.

Wieland, N., Boss, J., Lettmann, S., Fritz, B., Schwaiger, K., Bauer, J. (2017). Susceptibility to disinfectants in antimicrobial-resistant and -susceptible isolates of Escherichia coli, Enterococcus faecalis and Enterococcus faecium from poultry–ESBL/ AmpC-phenotype of E. coli is not associated with resistance to a quaternary ammonium compound, DDAC. *J Appl Microbiol*. 122:1508–17.

Wooley, R.E., Gibbs, P.S., Brown, T.P., Maurer, J.J. (2000). Chicken embryo lethality assay for determining the virulence of avian *Esherichia coli*. *Avian Dis*. 44(2): 318-324.

Xiong, W., Sun, Y., Zeng, Z. (2018). Antimicrobial use and antimicrobial resistance in food animals. *Environ. Sci. and Pollution Res*. 25: 18377-18384. doi: 10.1007/s11356-018-1852-2.

Yang, H., Chen, S., White, D. G., Zhao, S., McDermott, P., Walker, R., Meng, J. (2004). Characterization of multiple-antimicrobial-resistant Escherichia coli isolates from diseased chickens and swine in China. *J Clin Microbiol* 42(8), 3483-3489.

Zhai, R., Fu, B., Shi, X., Sun, C., Liu, Z., Wang, S., *et al*. (2020). Contaminated in-house environment contributes to the persistence and transmission of NDM-producing bacteria in a Chinese poultry farm. *Environ*. *International* 139: 105715. doi: 10.1016/j.envint.2020.105715.

Zhao, S., Maurer, J. J., Hubert, S., De Villena, J. F., McDermott, P. F., Meng, J, White, D. G. (2005). Antimicrobial susceptibility and molecular characterization of avian pathogenic Escherichia coli isolates. *Vet Microbiol* 107(3-4), 215-224.

CHAPTER 5: DISCUSSION

5.1 Summary and Relationship Between Manuscripts

5.1.1 Summary of Chapter 2

Chapter 2 describes the methodology developed to sequence the genomes of *E. coli* extracted from the tissues of Saskatchewan broilers. Over a two-year period, I have isolated 496 *E. coli* from 18 colibacillosis outbreaks on Saskatchewan broiler farms. This includes 394 systemic *E. coli* isolated from 54 birds with confirmed colibacillosis, and 102 cecal *E. coli* isolated from 34 visually healthy birds from the same barns. The diagnosis of colibacillosis is confirmed post-necropsy by MALDI-TOF mass spectrometry, and samples of the heart, liver, and spleen of diseased birds are released to me for immediate bacterial isolation. Homogenized portions of each organ are incubated on MacConkey agar to select for gram-negative organisms, while *E. coli* are relatively distinguishable on this medium based on their general colony morphology and lactose-fermenting capability.

To date, 104 *E. coli* isolates from 12 colibacillosis outbreaks have been sequenced on the Oxford Nanopore MinION sequencer. Our laboratory uses a modified version of the Porecamp Vancouver Nanopore Sequencing protocol, which is optimized for higher DNA inputs and longer fragment lengths than the standard Nanopore sequencing kit protocol (Lohman *et al.,* 2018). The Porecamp protocol includes a magnetic bead clean-up step to remove DNA fragments under 500bp prior to library preparation. This helps to increase the quality and average read length of a sequencing run. DNA was stored post-magnetic bead clean-up and the same DNA was used for Illumina sequencing. 86 *E. coli* isolates have been sequenced on an Illumina HiSeq 400. Hybrid *E. coli* genomes will be assembled by Dr. Haley Sanderson, a postdoctoral fellow in the White Lab. These hybrid genomes will be used for further analysis.

Draft genomes of 96 *E. coli* isolates were generated using long Nanopore reads to confirm that the isolation, extraction, and library preparations methods would produce sufficient reads to assemble *E. coli* genomes. These draft genomes were used to determine the phylogroup of each *E. coli* isolate. 60.5% of cecal isolates fell into phylogroup A, followed by 15.8% in phylogroup B1. Systemic isolates were split between three primary groups: phylogroup G (32.8%), phylogroup A

(20.7%), and phylogroup D (19%). A preliminary comparison of systemic and cecal isolates from outbreak #1 suggested that there are numerous genetic differences between these two groups.

Screening the draft genomes for plasmid content determined that systemic isolates possess significantly more plasmids than cecal isolates. These plasmids contain many antimicrobial resistance genes, in particular beta-lactam, tetracycline, and aminoglycoside resistance genes. Cecal isolates often lacked plasmids possessing AMR genes. The chromosomes of the 96 isolates contained very few resistance genes. Virulence genes were more widespread in both isolates, though over 50% of systemic isolates possessed a group of plasmid-mediated virulence genes involved in iron acquisition and evasion of the host immune system. These differences suggest that systemic and cecal isolates may belong to two distinct populations of *E. coli*.

5.1.2 Summary of Chapter 3

The aim of chapter 4 was to develop a comprehensive methodology to capture and sequence all plasmids from avian *E. coli* isolates. Our genome sequencing protocol includes a size-selection step which removes DNA fragments below 500bp before sequencing. Without a dedicated plasmid purification step, we could not say whether we were capturing the entire plasmid content of each isolate. We used six popular commercial plasmid extraction kits to test whether the choice of kit affected downstream assembly results. Because Nanopore sequencing requires linear DNA, we tested two library preparation kits: the standard ligation-based kit (LSK109) which requires linear ends for ligation, and a transposon-based (RBK004) barcoding kit which does not require linear DNA. We hypothesized that some percentage of plasmids would be nicked during extraction, producing linear ends for the ligation-based library preparation method.

Using a dedicated plasmid capture step, we were able to recover an additional four plasmids from two *E. coli* isolates: three from a systemic isolate, and one from a cecal isolate. These plasmids ranged in length from 2kb to 265kb and contained a mix of virulence and antimicrobial resistance genes. Overall, the LSK109 kit performed better than the RBK004 kit. The average read lengths from the LSK109 sequencing run were longer than those from the RBK004 kit. Even with an incubation time of only one minute, the transposon-based method produced highly fragmented plasmid DNA. Due to the common presence of large repeat regions in both chromosomal and plasmid DNA, *de novo* assembly is much easier with longer reads. We found that plasmid kits that

produced longer fragments generated more complete assemblies. No single set of reads was able to identify all of the plasmids in both isolates, but a combination of the reads from two kits was sufficient to identify all the plasmids in both isolates. In the future we will include a dedicated plasmid step using plasmid DNA from two kits to ensure that we capture all plasmids from our avian *E. coli* isolates.

5.1.3 Summary of Chapter 4

Chapter 4 provides an in-depth analysis of the antimicrobial resistance and biofilm phenotypes of 94 *E. coli* isolated from colibacillosis outbreaks on Saskatchewan broiler farms. We measured biofilm formation in a population of *E. coli* isolated from diseased and healthy chickens to better understand the lifestyle of avian *E. coli* outside of an animal host. Because biofilm formation is often linked to increased resistance to antimicrobials, we tested each strain for phenotypic resistance to 27 different antimicrobial drugs. Systemic *E. coli* formed significantly greater biofilms in nutrient-rich media compared to cecal *E. coli* but formed weak or negligible biofilms in minimal media. In contrast, while cecal isolates overall formed weaker biofilms than systemic isolates, biofilm formation in nutrient-rich and minimal media was comparable. There were clear differences in antimicrobial resistance among systemic and cecal *E. coli*. 71% of systemic isolates were resistant to at least one drug while only 29% of cecal isolates were drug resistant. Resistance to ampicillin, tetracycline, gentamicin, and tobramycin were observed in 20-40% of systemic isolates and 10-20% of cecal isolates. We observed no correlation between antimicrobial resistance status and biofilm formation in any of the tested media. Using a crystal violet assay and biofilm growth on polystyrene pegs we identified the 21 strongest biofilm-forming isolates in each media. Of these, 11 were systemic isolates and 10 were cecal isolates. All 10 cecal isolates and one systemic isolate were susceptible to all tested antimicrobials, while the remaining 10 systemic isolates showed no pattern in their antimicrobial resistance profiles.

We observed a relationship between individual outbreaks and biofilm formation in a preferred medium. Isolates from 10/12 farms preferentially formed biofilms in a single type of media. The most common was BHI broth (7/12 outbreaks), followed by M63 broth (2/12 outbreaks), then ½ TSB (1 outbreak). There was no trend observed on two outbreaks, however broilers submitted from one of those outbreaks spent several days in a hot truck before being brought to PEX for necropsy.

5.1.4 Systemic and Cecal Isolates Likely Belong to Two Distinct Populations

Comparison of the first eight draft genomes generated from an outbreak (outbreak 1) of colibacillosis in July of 2018 suggested that systemic and cecal isolates differed greatly in several areas of their genomes. While this was a rough comparison using unpolished long read assemblies, it prompted us to investigate whether we could identify differences between the two groups. Determining the phylogroups of 96 Nanopore sequenced *E. coli* showed that 76% of cecal isolates fall into phylogroups A and B1, which are associated with non-pathogenic avian *E. coli* strains (Kim *et al.,* 2020). In contrast, 72.5% of systemic isolates are distributed among phylogroups A, D, and G. Phylogroup D contains many extraintestinal pathogenic *E. coli* (ExPEC), while phylogroup G is associated with highly virulent and drug-resistant *E. coli* isolates, including several poultry-associated lineages (Koga *et al.,* 2015; Clermont *et al.,* 2019).

The genetic differences we observed in chapter 2 prompted us to investigate whether systemic and cecal *E. coli* populations differed in other aspects. Analysis of phenotypic antimicrobial resistance in 94 *E. coli* strains revealed that the population was almost evenly split: 55% were resistant to at least 1/27 tested antibiotics, while 45% were susceptible to all tested drugs. However, when the isolates were grouped based on disease status, 71% of systemic isolates were resistant to at least 1/27 antibiotics, while only 29% of cecal isolates were resistant. If cecal *E. coli* are responsible for causing disease in broilers, we would expect to see conserved patterns of drug resistance in both systemic and cecal isolates from the same barns (Schroeder *et al.,* 2004). Instead, we found that cecal isolates were often susceptible to all tested antimicrobials, and resistant cecal isolates within a flock did not share the same AMR profiles as systemic isolates. Levels of resistance to aminoglycoside and penicillin antibiotics were 3-4 times higher in systemic isolates than in cecal isolates, though this may be partially attributable to our sample collection methods. Numerous studies have demonstrated that *E. coli* isolated from very young chicks tends to have higher levels of drug-resistance, and resistance levels decrease as the age of the bird increases (Miranda *et al.,* 2008; Chuppava *et al.,* 2019; Montoro-Dasi *et al.,* 2019). Several colibacillosis outbreaks involved sick chicks less than one week old. Because we did not collect healthy broilers at the same time as diseased birds, up to two weeks sometimes passed before healthy birds were culled and submitted.

Systemic and cecal isolates also differed with respect to biofilm formation. Biofilm formation can protect bacteria that spend extended periods of time in an unpredictable environment, such as a chicken barn, and may provide insights into the environmental niche that bacteria inhabit (MacKenzie *et al*., 2017). Biofilms can also increase tolerance to antimicrobials and disinfectants, which may make the management of colibacillosis in broilers more difficult for farmers. Over 50% of our study population produced both cellulose and curli, which are key components of biofilms. While slightly more (73% vs 55%) systemic isolates produced both components than cecal isolates, there was not a significant difference in curli and cellulose production between these two groups, and we hypothesized that most of the systemic and cecal *E. coli* would be able to form biofilms in biofilm-inducing media.

Systemic *E. coli* formed significantly stronger biofilms in rich BHI broth than in minimal M63 and ½ TSB media. In BHI broth, 72.2% of systemic isolates formed a positive biofilm, while only 28.6% formed positive biofilms in ½ TSB broth, and 23.2% formed positive biofilms in M63 broth. In contrast, over 50% of cecal isolates formed positive biofilms in both BHI (55.3%) and M63 broth (52.6%). Biofilm formation in *E. coli* is highly variable and depends on numerous genetic and environmental factors including temperature, genes for virulence and adhesion, and nutrient availability (Reisner *et al*., 2006; Skyberg *et al*., 2007). Previous studies show that *E. coli* that spend extended periods of time in a host, such as human and fecal APEC, as well as urinary pathogenic and enteroaggregative *E. coli,* are more likely to form biofilms in nutrient-rich media (Skyberg *et al*., 2007; Nielsen *et al*., 2018). Cecal *E. coli* likely persist as commensal organisms in bacterial biofilms in the intestinal niche of broilers. Systemic isolates may also rely on biofilm formation within the host to cause disease. If they persist in the intestinal niche before escaping to cause infection, they may form biofilms for protection against the host immune system and other bacteria in the gut (Desai *et al*., 2019). Similarly, pathogenic *E. coli* present in inhaled fecal dust may form a biofilm in the respiratory tract before escaping into the bloodstream to cause systemic disease (Vestby *et al*., 2020). In addition, APEC isolates are more likely to form negligible or weak biofilms in M63 media compared to avian fecal isolates (Nielsen *et al*., 2018).

Finally, we observed that systemic *E. coli* possessed an average of four plasmids, which was statistically significant compared to the average of one plasmid found in cecal isolates. While it has been suggested for years that plasmids play a role in virulence and persistence of APEC

strains, their exact role is still unknown (Johnson *et al*., 2008; Tivendale *et al*., 2009; Mellata *et al*., 2010; Magueiros *et al*., 2021).

5.1.5 Plasmid-mediated Factors Likely Contribute to Biofilm Formation and Antimicrobial Resistance

We were unable to find a correlation between antimicrobial resistance and biofilm formation in our *E. coli* population. We identified 21 strains that formed the strongest biofilms (with viable bacterial counts of over 10^7 CFU/mL) in ½ TSB, M63, and BHI broth. Of those 21, 11 were systemic isolates and 10 were cecal. All of these isolates formed strong biofilms regardless of their phenotypic AMR status, suggesting that other factors influence strong biofilm formation. The paper in chapter 3 was limited to the analysis of phenotypic data and did not explore the relationship between genetic characteristics involved in antimicrobial resistance or biofilm formation.

Antimicrobial resistance genes were almost always present on plasmids rather than the chromosome, which may explain why systemic *E. coli* were significantly more drug-resistant than cecal isolates. 41% of cecal isolates lacked any plasmid-encoded antimicrobial resistance genes, while only 10% of systemic isolates lacked any AMR genes. Consistent with the results discussed in chapter 3, various plasmid-mediated aminoglycoside, tetracycline, and beta-lactam resistance genes were found in many systemic isolates. The most common genes were *tetA, aph(3")*, *aac(3)-2d,* and *blaTEM-1B*. All of the plasmid-derived AMR genes identified in our *E. coli* isolates were present on plasmids with numerous virulence genes. This is consistent with previous analysis of APEC plasmids, which showed that AMR cassettes are common in known virulence plasmids such as pAPEC-1 and pAPEC-2 (Hall, 2007; Cummins *et al*., 2019).

Depending on the genetic background of an *E. coli* strain, carriage of antimicrobial resistance genes may hinder biofilm formation (Teh *et al*., 2014). For example, in the study by Teh *et al*., the presence of a plasmid-mediated ampicillin resistance gene significantly hindered biofilm formation in a resistant mutant *E. coli* strain compared to the wild-type strain and a strain with a different phylogenetic background (2014). Similar results were observed with the acquisition of a kanamycin resistance gene. The 96 *E. coli* analyzed in this study represent a wide variety of phylogroups. Given the diverse genetic background of Saskatchewan strains, acquisition of the

same antimicrobial resistance gene may produce different biofilm phenotypes. The presence of APEC plasmids in systemic *E. coli* from our study may explain why they formed poor biofilms in minimal media: growth of APEC strains that possessed pAPEC-1 or pAPEC-2 was significantly slower in minimal media compared to LB broth (Mellata *et al*., 2010). The cecal isolates in our study were significantly less likely to possess APEC virulence plasmids, and they formed biofilms equally well in both nutrient-rich and minimal media.

Cecal isolates formed greater biofilms in M63 media compared to systemic isolates. All the cecal isolates which formed the greatest biofilms in this media possessed chromosomal genes encoding for effectors typically found in intestinal pathogenic *E. coli*. These included genes from the locus of enterocyte effacement (*espABF, nleAB*) as well as genes for long polar fimbriae (*lpfA*). The *esp* genes in *E. coli* are involved in the translocation and adhesion of different secreted proteins and *esp*+ strains form significantly stronger biofilms than *esp*-mutant strains (Moreira *et al*., 2006). Long polar fimbriae (*lpfA*) play a role in adhesion during colonization, as well as biofilm formation (Ross *et al*., 2015; Zhou *et al*., 2020). It was previously demonstrated that some human pathogenic *E. coli* could grow well in M63 media, which may explain why our avian isolates with IPEC virulence genes formed strong biofilms in M63 broth (Nielsen *et al*., 2018).

Systemic isolates formed the strongest biofilms in BHI broth. Many of these strains possessed plasmid-encoded hemolysin genes (*hly*), which are commonly found in APEC and are significantly associated with biofilm formation (Teh *et al*., 2014). These systemic isolates also possessed a variety of chromosomal virulence genes associated with biofilm formation, such as *fyuA*, a ferric yersiniabactin uptake protein commonly associated with virulence in *Escherichia* species (Schubert *et al*., 2002). Upregulation of the *fyuA* gene significantly increases biofilm formation in low-iron environments (Hancock *et al*., 2008). E. *coli* that form strong biofilms in nutrient-rich broth such as BHI typically spend extended periods of time in the gastrointestinal tract (Nielsen *et al*., 2018). Both commensal and pathogenic *E. coli* of avian and human origin spend periods of time in the gut and would need to possess genes to facilitate colonization (Mellata, 2013; Conway and Cohen, 2015).

5.1.6 Plasmids as a Surveillance and Diagnostic Tool for AMR and Disease

Using whole genome sequencing data, we determined that systemic *E. coli* possessed a significantly greater number of plasmids than cecal *E. coli*. Plasmids play an important role in antimicrobial resistance, pathogenesis, metabolism, and survival (Flint, 1987; Sengupta and Austin, 2011; Dolejska *et al.,* 2014; Rowzwandowicz *et al.,* 2018). To better understand the differences between systemic and cecal *E. coli* isolated from Saskatchewan broiler flocks it is critical that all the genetic material from each isolate is being analyzed. A genomic DNA kit will capture some of the plasmid content of a bacterial culture, but likely not all (Nouws *et al*., 2020). We developed a dedicated plasmid capture method to determine whether we were missing plasmids using our WGS sequencing protocol. Our method identified an additional four plasmids in two *E. coli* isolates: 4957-3S1 and 4957-C3. These isolates were collected from a diseased (4957-3S1) and a healthy broiler (4957-C3) during the same colibacillosis outbreak. isolate 4957-3S1 was resistant to Ampicillin, Gentamycin, Tetracycline, and Tobramycin, while 4957-C3 was susceptible to all tested antimicrobials. Both had previously been sequenced using Nanopore and Illumina technologies, and draft assemblies from the long reads had identified three plasmids in 4957-3S1 and two plasmids in 4957-C3.

Rapid Nanopore sequencing of plasmids is a useful tool to identify and track dissemination of plasmids and is used in care settings such as hospitals (Peter *et al.,* 2020; Taxt *et al.,* 2020). Plasmids ranging from 1kb to thousands of kilobases in length can be sequenced using Nanopore technology (Wick *et al.,* 2021). However, a standard methodology for capturing and *de novo* assembling *E. coli* plasmids from genomic samples has not been developed. Most studies analyzing genome and plasmid data used a combination of Nanopore and Illumina sequencing, which is more time-consuming and costly (Gonzalez-Escalona *et al*., 2019; Taylor *et al*., 2019). Our plasmid extraction method can be completed in under 24 hours and may provide key information about antimicrobial resistance and, potentially, virulence of *E. coli* strains. Large plasmids may be a defining feature of APEC, and those plasmids may allow some strains to cause disease in other species, including humans (Jorgensen *et al.,* 2019). The presence of genes from human intestinal pathogenic *E. coli* in cecal isolates from our study population demonstrates that avian *E. coli* can carry and transmit virulence genes from human strains. The presence of additional genes for adhesion and evasion of the host immune system on virulence plasmids identified in Saskatchewan *E. coli* strains may allow some strains to behave as zoonotic pathogens, increasing the risk of extraintestinal disease in the human population. The presence of antimicrobial resistance genes on

many of these plasmids could result in difficult to treat infections in producers and/or consumers. Monitoring virulence and drug-resistance plasmids in food animals has direct implications on the health of local human populations, and this method provides a rapid and affordable way to do so.

Because nearly all the antimicrobial resistance genes identified in our *E. coli* population were present on plasmids, this plasmid capture method could easily be used to track the seasonality and spread of antimicrobial resistance trends on Saskatchewan farms.

5.1.7 Inclusion Criteria are an Important Factor in the Study of Colibacillosis and APEC

There are numerous conflicting studies discussing the presence and distribution of plasmids in APEC and commensal poultry isolates (Johnson *et al.,* 2007; Mellata *et al*., 2010; Cunha *et al*., 2012; Johnson *et al*., 2012; Mageiros *et al*., 2021). Several studies classify the possession of multiple large virulence plasmids as a defining characteristic of APEC isolates, possession of the same plasmids by cecal isolates varies greatly by study (Schouler *et al*., 2012; Ikuta *et al*., 2014; Mageiros *et al*., 2021). We found that systemic and cecal *E. coli* possessed a significantly different number of plasmids. The true number of plasmids in our isolates will have to be confirmed after a dedicated plasmid capture step, and there is a chance that sequencing additional isolates will nullify this significant difference, however, the large virulence plasmids implicated in APEC pathogenesis have likely been identified using our original WGS protocol. We suspect that using strict inclusion criteria in our study population has helped us to identify differences in these two key groups. Many studies include isolates from wide geographic locations, different presentations of disease, and numerous types of poultry (i.e., broilers, laying hens, turkeys, wild geese, and ducks). These additional confounding factors may be masking differences between APEC and commensal *E. coli*. For example, E. *coli* infections in turkeys are uncommon, and isolates recovered from turkeys possess fewer virulence genes and are less diverse than isolates recovered from chickens, suggesting that different factors may influence infection in various types of birds (Al-Kandari and Woodward, 2019). *E. coli* isolates included in our study were isolated exclusively from Saskatchewan broilers with confirmed colibacillosis and healthy broilers living in the same barns.

The differences we observed in biofilm formation and antimicrobial resistance in our *E. coli* isolates help support the hypothesis that systemic and cecal isolates belong to separate

populations. Our biofilm formation results differ somewhat from existing literature. In our study, systemic strains were more capable of forming biofilms compared to the studies by Skyberg *et al*. and Nielsen *et al*. (2007; 2018). They found that less than 50% of APEC strains formed positive biofilms in ½ TSB, M63, and BHI broth. However, their study population analyzed large, curated strains of APEC and avian commensal isolates collected from various poultry sources in multiple countries over several years (Skyberg *et al*., 2007; Nielsen *et al*., 2018). Because a significant amount of the diversity found in *E. coli* isolates is due to horizontal gene transfer, analyzing such a varied study population may be masking factors that play a role in specific environments, such as biofilm formation. Systemic *E. coli* on Saskatchewan broiler farms form strong biofilms in BHI broth. Without a comparison group from a similar environment, such as our cecal isolates, it would be very difficult to identify genetic differences in the two populations that may explain biofilm formation. Future studies exploring the pathogenesis of APEC isolates should be designed with the diversity of *E. coli* in mind.

5.2 Limitations of Work

To our knowledge, this is the first study to characterize and collect systemic *E. coli* using a standard definition of colibacillosis. Isolates analyzed in this study were recovered exclusively from broilers in a common geographic area, removing several confounding variables that exist in other studies (Mageiros *et al*., 2021). However, the 'gold-standard' method to characterize *E. coli* strains as APEC is to use a chicken or embryo lethality assay (Skyberg *et al*., 2006). We have yet not confirmed whether each isolate is pathogenic or commensal. Up to 15% of *E. coli* isolated from the intestinal contents of broilers are classified as APEC strains, which means that some of our cecal isolates may be mischaracterized (Mellata, 2013). In addition, we have no control over how quickly dead broilers are brought to PEX for necropsy. Once a broiler is dead the immune system is no longer preventing bacterial growth, which means that some of the *E. coli* strains isolates from birds with confirmed colibacillosis may not be the original pathogenic strain. To say with certainty that systemic isolates that cause disease are genetically different, we would need to confirm their lethality in chicken models.

The analysis of sequencing data presented in this thesis consists exclusively of preliminary data. High quality genome sequences will be generated by a trained bioinformatician, and those sequences will be used in an unbiased analysis of systemic and cecal isolates.

We chose to sample commensal *E. coli* from the cecal contents of healthy broilers rather than from the cloaca to avoid contamination with environmental *E. coli*. While the cecal *E. coli* provide a same-flock comparison group, we did not analyze *E. coli* from the cecal contents of broilers with confirmed colibacillosis. In order to rule out the theory that colibacillosis is caused by opportunistic strains that escape the intestinal niche, we would need to look for identical strains present in the intestines of broilers with colibacillosis. Once again, however, the amount of time that passes before a broiler is brought for necropsy could mean that pathogenic strains overwhelm and contaminate the cecal contents. We have started collecting *E. coli* strains from the cecal contents of broilers with colibacillosis, and these will be used for future analysis.

The trends observed in the current data come from a relatively small (n=96) sample of *E. coli* isolates. Over 500 isolates have been collected from Saskatchewan colibacillosis outbreaks over the duration of the study. We need to analyze additional *E. coli* genomes to confirm the trends that we saw in the current study. The original set of 96 isolates analyzed for this thesis were selected based on disease status, outbreak, colony morphology, and biofilm morphology to reduce the amount of potential clonal isolates in the set. We will need to analyze more than 58 systemic and 38 cecal isolates to determine whether some of the observed differences in these groups are statistically significant. We also need to analyze the full plasmid content of each isolate. We have confirmed that plasmids are lost due to the genomic DNA kit and size-selection steps during sequencing. While it is unlikely that we missed large plasmids containing AMR or virulence genes, this cannot be ruled out until after a plasmid capture step.

5.3 Future Directions

In order to confirm the pathogenicity of strains analyzed in this study, chicken lethality assays will need to be conducted. Once pathogenicity is established, we can begin to tease apart the role of numerous virulence genes involved in the infection process using genetic knockout experiments. To determine whether APEC plasmids play a role in systemic isolates' ability to form

significant biofilms in nutrient rich BHI media we could attempt to cure strains of their plasmids and measure their biofilm formation in the absence of, or with different combinations of, plasmids. Additionally, transforming APEC plasmids into cecal isolates that performed poorly in nutrient-rich broth would provide additional insights into how much these plasmids influence biofilm formation. Plasmid capture and sequencing for all 96 *E. coli* isolates will confirm whether there is a significant difference in the number of plasmids possessed by systemic and cecal *E. coli*. This plasmid information could be used to develop a Nanopore sequencing platform to provide rapid information about the pathogenicity and AMR status of *E. coli* isolates on Saskatchewan broiler farms. Finally, the methodology developed for the analysis and tracking of Saskatchewan colibacillosis isolates could be applied to *E. coli* infections and bacterial diseases in other livestock animals.

5.4 References

Barry, K.E., Wailan, A.M., Sheppard, A.E., Crook, D., Vegesana, K., Stoesser, N., Parikh, H.I., Sebra, R., Mathers, A.J. (2019). Don't overlook the little guy: an evaluation of the frequency of small plasmids co-conjugating with larger carbapenemase gene containing plasmids. *Plasmid* 103: 1-8. doi: 10.1016/j.plasmid.2019.03.005.

Burian, J., Guller, L., Macor, M., Kay, W.W. (1997). Small cryptic plasmids of multiplasmid, clinical *Escherichia coli*. *Plasmid*. 31(1): 2-14. doi: 10.1006/plas.1996.1273.

Chuppava, B., Keller, B., El-Wahab, A.A., Surie, C., Visscher, C. (2019). Resistance reservoirs and multi-drug resistance of commensal *Escherichia coli* from excreta and manure isolated in broiler houses with different flooring designs. *Front. Microbiol*. 10:2633. doi: 10.3389/fmicb.2019.02633.

Clermont, O., Dixit, O.V.A., Vangchhia, B., Condamine, B., Dion, S., Bridier-Nahimas, A., *et al*. (2019). Characterization and rapid identification of phylogroup G in *Escherichia coli*, a lineage with high virulence and antibiotic resistance potential. *Environ. Microbiol*. 21(8): 3107-3117. doi: 10.1111/1462-2902.14713.

Conway, T., Cohen, P.S. (2015). Commensal and pathogenic *Escherichia coli* metabolism in the gut. *Microbiol. Spectr*. 3(3). doi: 10.1128/microbiolspec.MBP-0006-2014.

Cummins, M.L., Reid, C.J., Chowdhury, P.R., Bushell, R.N., Esbert, N., Tivendale, K.A., *et al*. (2019). Whole genome sequence analysis of Australian avian pathogenic *Escherichia coli* that carry the class 1 integrase gene. *Microb. Genom*. 5(2): e000250. doi: 10.1099/mgen.0.000250.

Cunha, M.P.V., Saidenberg, A.B., Moreno, A.M., Ferreira, A.J.P., Vieira, M.A.M., Gomes, T.A.T., Knobl, T. (2017). Pandemic extra-intestinal pathogenic *Escherichia coli* (ExPEC) clonal group O6-B2-ST73 as a cause of avian colibacillosis in Brazil. *PLoS One* 12(6): e0178970. doi: 10.1371/journal.pone.0178970.

Desai, S.K., Kenney, L.J. (2019). Switching lifestyles is an *in vivo* adaptive strategy of bacterial pathogens. *Front. Cell. Infect. Microbiol.* 9:421. doi: 10.3389/fcimb.2019.00421.

Dolejska, M., Villa, L., Minoia, M., Guardabassi, L., Carattoli, A. (2014). Complete sequences of IncHI1 plasmids carrying blaCTX-M-1 and qnrS1 in equine *Escherichia coli* provide new insights into plasmid evolution. *J. Antimicrob. Chemother.* 69(9): 2388-93. doi: 10.1093/jac/dku172.

Ejrnaes, K., Stegger, M., Reisner, A., Ferry, S., Monsen, T., Holm, S., *et al.* (2011). Characteristics of *Escherichia coli* causing persistence or relapse of urinary tract infections: phylogenetic groups, virulence factors, and biofilm formation. *Virulence.* 2(6). doi: 10.4161/viru.2.6.18189.

Flint, H.J., Stewart, C.S. (1987). Antibiotic resistance patterns and plasmids of ruminal strains of *Bacteroides ruminicola* and *Bacteroides multiacidus*. *Appl. Microbiol.* 26: 450-455. doi: 10.1007/BF00253531.

Hall, R.M. (2007). Antibiotic resistance gene cluster of pAPEC-O1-R. *Antimicrob. Agents Chemother.* 51(9): 3461-3462. doi: 10.1128/AAC.00417-07.

Hancock, V., Ferrieres, L., Klemm, P. (2008). The ferric yersiniabactin uptake receptor FyuA is required for efficient biofilm formation by urinary tract infectious *Escherichia coli* in human urine. *Microb. Soc.* 154(1). doi: 10.1099/mic.0.2007/011981-0.

Johnson, T.J., Wannemuehler, Y.M., Johnson, S.J., Logue, C.M., White, D.G., Doetkott, C., Nolan, L.K. (2007). Plasmid replicon typing of commensal and pathogenic *Escherichia coli* isolates. *Appl. Environ. Microbiol.* 73(6): 1976-1983. doi: 10.1128/AEM.02171-06.

Johnson, T.J., Wannemuehler, Y., Doetkott, C., Johnson, S.J., Rosenberger, S.C., Nolan, L.K. (2008). Identification of minimal predictors of avian pathogenic *Escherichia coli* virulence for use as a rapid diagnostic tool. *J. Clin. Microbiol.* 46(12): 3987-96. doi: 10.1128/JCM.00816-08.

Johnson, T.J., Logue, C.M., Johnson, J.R., Koskowski, M.A., Sherwood, J.S., Barnes, H.J., DebRoy, C., Wannemuehler, Y.M., Obata-Yasuoka, M., Spanjaard, L., Nolan, L.K. (2012). Associations between multidrug resistance, plasmid content, and virulence potential among extraintestinal pathogenic and commensal *Escherichia coli* from humans and poultry. *Food. Pathog. And Dis.* 9(1). doi: 10.1089/fpd.2011/0961.

Jorgensen, S., Stegger, M., Kudirkiene, E., Lilje, B., Poulsen, L.L., Ronco, T., Dos Santos, T.P., Kiil, K., Bisgaard, M., *et al.* (2019). Diversity and population overlap between avian and human

Escherichia coli belonging to sequence type 95. *mSphere*. 4(1): 00333-18. doi:10.1128/mSphere.00333-18.

Kim, Y.B., Yoon, M.Y., Ha, J.S., Seo, K.W., Noh, E.B., Son, S.H., Lee, Y.J. (2020). Molecular characterization of avian pathogenic *Escherichia coli* from broiler chickens with colibacillosis. *Poul. Sci*. 99(2): 1088-1095. doi: 10.1016/j.psj.2019.10.047.

Koga, V.L., Rodrigues, G.R., Scandorieiro, S., Vespero, E.C., Oba, A., Brito, B.G., *et al*. (2015). Evaluation of the antibiotic resistance and virulence of *Escherichia coli* strains isolated from chicken carcasses in 2007 and 2003 from Parana, Brazil. *Foodborne Pathog. Dis*. 12(6). doi: 10.1089/fpd.2014.1888.

MacKenzie, K.D., Palmer, M.B., Koster, W.L., White, A.P. (2017). Examining the link between biofilm formation and the ability of pathogenic *Salmonella* strains to colonize multiple host species. *Front. Vet. Sci*. doi: 10.3389/fvets.2017.00138.

Mageiros, L., Meric, G., Bayliss, S.C., Pensar, J., Pascoe, B., Mourkas, E., Calland, J.K., *et al*. (2021). Genome evolution and the emergence of pathogenicity in avian *Escherichia coli*. *Nature Comm*. 12:765. doi: 10.1038/s41467-021-20988-w.

Mellata, M., Ameiss, K., Mo, H., Curtiss, R. (2010). Characterization of the contribution to virulence of three large plasmids of avian pathogenic *Escherichia coli* $_{X}$7122 (O78:K80:H9). *Infect Immun*. 78(4): 1528-41. doi: 10.1128/IAI.00981-09.

Mellata, M. (2013). Human and avian extraintestinal pathogenic *Escherichia coli*: infections, zoonotic risks, and antibiotic resistance trends. *Foodborn Pathog. Dis*. 10(11): 916-932. doi: 10.1089/fpd.2013.1533

Millan, A.S., Escudero, J.A., Gutierrez, B., Hidalgo, L., Garcia, N., Llagostera, M., Dominguez, L., Gonzalez-Zorn, B. (2009). Multiresistance in *Pasturella multocida* is mediated by coexistence of small plasmids. *Antimicrob. Agents Chemother*. 53(8): 3399-3404. doi: 10.1128/AAC.01522-08.

Millan, A.S., MacLean, R.C. (2017). Fitness costs of plasmids: a limit to plasmid transmission. *Microbiol. Spectr*. 5(5). doi: 10.1128/microbiolspec.MTBP-0016-2017.

Miranda, J.M., Vazquez, B.I., Barros-Velazquez, J., Cepeda, A., Franco, C.M. (2008). Evolution of resistance in poultry intestinal *Escherichia coli* during three commonly used antimicrobial therapeutic treatments in poultry. *Poul. Sci*. 87(8): 1643-1648. doi: 10.3382/ps.2007-00485.

Montoro-Dasi, L., Villagra, A., Sevilla-Navarro, S., Perez-Gracia, M.T., Vega, S., Marin, C. (2020). The dynamic of antibiotic resistance in commensal *Escherichia coli* throughout the growing period in broiler chickens: fast-growing vs. slow-growing breeds. *Poul. Sci*. 99(3): 1591-1597. doi: 10.1016/j.psj.2019.10.080.

Moreira, C.G., Palmer, K., Whiteley, M., Sircili, M.P., Trabulsi, L.R., Castro, A.F.P., Sperandio, V. (2006). Bundle-forming pili and EspA are involved in biofilm formation by enteropathogenic *Escherichia coli*. *J. Bacteriol*. 188(11): 3952-3961. doi: 10.1128/JB.00177-06.

Naves, P., del Prado, G., Huelves, L., Gracia, M., Ruiz, V., Blanco, J., Dahbi, G., *et al*. (2008). Correlation between virulence factors and *in vitro* biofilm formation by *Escherichia coli*. *Microb. Pathog*. 45(2): 86-91. doi: 10.1016/j.micpath.2008.03.003.

Nielsen, D.W., Klimavicz, J.S., Cavender, T., Wannemuehler, Y., Barbieri, N.L., Nolan, L.K., Logue, C.M. (2018). The impact of media, phylogenetic classification, and *E. coli* pathotypes on biofilm formation in extraintestinal and commensal *E. coli* from humans and animals. *Front. Microbiol*. 9(902). doi: 10.3389/fmicb.2018.00902.

Nouws, S., Bogaerts, B., Verhaegen, B., Denayer, S., Pierard, D., Marchal, K., *et al*. (2020). Impact of DNA extraction on whole genome seqeuencing analysis for characterization and relatedness of Shiga toxin-producing *Escherichia coli* isolates. *Nat. Sci. Rep*. 10: 14649. doi: 10.1038/s41598-020-71207-3.

Peter, S., Bosio, M., Gross, C., Bezdan, D., Gutierrez, J., Oberhettinger, P., *et al*. (2019). Tracking of antibiotic resistance transfer and rapid plasmid evolution in a hospital setting by Nanopore sequencing. *mSphere* 5: e00525-20. doi: 10.1128/mSphere.00525-20.

Reisner, A., Krogfelt, K.A., Klein, B.M., Zechner, E.L., Molin, S. (2006). In vitro biofilm formation of commensal and pathogenic *Escherichia coli* strains: impact of environmental and genetic factors. *J Bacteriol*. 188(10): 3572-3581. doi: 10.1128/JB.188/10.3572-3581.2006.

Ross, B.N., Rojas-Lopez, M., Cieza, R.J., McWilliams, B.D., Torres, A.G. (2015). The role of long polar fimbriae in *Escherichia coli* O104:H4 adhesion and colonization. *PLoS ONE*. 10(10): e0141845. doi: 10.1371/journal.pone.0141845.

Rozwandowicsz, M., Brouwer, M.S.M., Fischer, J., Wagenaar, J.A., Gonzalez-Zorn, B., Guerra, B., Mevius, D.J., Hordijk, J. (2018). Plasmids carrying antimicrobial resistance genes in Enterobacteriaceae. *J. Antimicrob. Chemother*. 73(5): 1121-1137. doi: 10.1093/jac/dkx488.

Schroeder, C.M., White, D.G., Meng, J. (2004). Retail meat and poultry as a reservoir of antimicrobial-resistant *Escherichia coli*. *Food Microbiol*. 21(3): 249-255. doi: 10.1016/S0740-0020(03)00074-1.

Schubert, S., Picard, B., Gouriou, S., Heesemann, J., Denamur, E. (2002). *Yersinia* high-pathogenicity island contributes to virulence in *Escherichia coli* causing extraintestinal infections. *Infect. Immun*. 70:5335-5337.

Sengupta, M., Austin, S. (2011)/. Prevalence and significance of plasmid maintentance functions in the virulence plasdmis of pathogenic bacteria. *Infect. Immunol.* 79(7): 2502-2509. doi: 10.1128/IAI.00127-11.

Skyberg, J.A., Horne, S.M., Giddings, C.W., Wooley, R.E., Gibbs, P.S., Nolan, L.K. (2003). Characterizing avian *Escherichia coli* isolates with multiplex polymerase chain reaction. *Avian Dis.* 47(4): 1441-1447.

Skyberg, J.A., Johnson, T.J., Johnson, J.R., Clabots, C., Logue, C.M., Nolan, L.K. (2006). Acquisition of avian pathogenic *Escherichia coli* plasmids by a commensal *E. coli* isolate enhances its ability to kill chicken embryos, grow in human urine, and colonize the murine kidney. *Infect. Immun.* 74(11): 6287-6292. doi: 10.1128/IAI.00363-06.

Skyberg, J., Siek, K.E., Doetkoff, C., Nolan, L.K. (2007). Biofilm formation by avian *Eschericha coli* in relation to media, source and phylogeny. *J. Appl. Microbiol.* 102(2): 548-54. doi: 10.1111/j.1365-2672.2006.03076.x

Taxt, A.M., Avershina, E., Frye, S.A., Naseer, U., Ahmad, R. (2020). Rapid identification of pathogens, antibiotic resistance genes, and plasmids in blood cultures by nanopore sequencing. *Nat. Sci. Rep.* 10: 7622. doi: 10.1038/s41598-020-64616-x.

Teh, A.H.T., Wang, Y., Dykes, G.A. (2014). The influence of antibiotic resistance gene carriage on biofilm formation by two *Escherichia coli* strains associated with urinary tract infections. *C. Journ. Microbiol.* 60: 105-111. doi: 10.1139/cjm-2013-0633.

Tivendale, K.A., Allen, J.A., Browning, G.F. (2009). Plasmid-borne virulence-associated genes have a conserved organization in virulent strains of Avian Pathogenic *Escherichia coli*. *J. Clin. Microbiol.* 47(8): 2513-2519. doi: 10.1128/JCM.00391-09.

Vestby, L.K., Gronseth, T., Simm, R., Nesse, L.L. (2020). Bacterial biofilm and its role in the pathogenesis of disease. *Antibiotics* 9(2): 59. doi: 10.3390/antibiotics9020059.

Vimr, E., Steenbergen, S., Cieslewicz, M. (1995). Biosynthesis of the polysialic acid capsule in *Escherichia coli* K1. *J. Industr. Microbiol.* 15: 352-360.

Wick, R.R., Judd, L.M., Wyres, K.L., Holt, K.E. (2021). Recovery of small plasmid sequences via Oxford Nanopore sequencing. *Microb. Genom.* 7(8): 000631. doi: 10.1099/mgen.0.000631.

Zhou, M., Ding, X., Ma, F., Xu, Y., Zhang, J., Zhu, G., Lu, Y. (2019). Long polar fimbriae contribute to pathogenic *Escherichia coli* infection to host cells. *Appl. Microbiol. Biotechnol.* 103(18): 7317-7324. doi: 10.1007/s00253-019-10014-x.

APPENDIX A. METHODS FOR BIOCHEMICAL CONFIRMATION AND PRODUCTION OF KEY BIOFILM COMPONENTS IN AVIAN *E. COLI*

A1. Biochemical Confirmation of *E. coli* Isolates

E. coli are gram-negative, lactose-fermenting bacilli that produce characteristic pink colonies on MacConkey agar (**Fig A1**) (Bolton *et al*., 1996).

Figure A1. *E. coli* colonies grown on MacConkey agar have a characteristic pink colour and often turn the media cloudy due to the presence of precipitated bile salts.

Because both *Klebsiella* and *Enterobacter* species are gram-negative, typically lactose-fermenting bacilli, we also perform citrate and indole biochemical tests to further characterize presumed-*E. coli* isolates. *E. coli* cannot use sodium citrate as a sole carbon source, while many *Klebsiella* and *Enterobacter* species can (Vaughn *et al*., 1950). Likewise, *E. coli* can hydrolyze tryptophan in tryptone broth to produce indole, while most *Klebsiella* and *Enterobacter* species do not (Hicks and Ryan, 1976).

One colony of presumed *E. coli* was streaked on a Simmons Citrate Agar plate to test for citrate utilization (Van Hofwegen *et al*., 2016). Plates were incubated at 37°C for 18 hours to look for visual colour changes. *E. coli* strains are typically citrate negative and do not change the colour of the agar, meaning that they cannot use citrate as a sole carbon source (Vaughn *et al*., 1950). Next, we checked for the ability of strains to produce indole through the catalyzation of a Tryptone deamination reaction (Niemi *et al*., 2003). In this test, one *E. coli* colony was inoculated

in 2 mL of 2% tryptone broth and incubated at 37°C for 18 hours. After 18h, 3-4 drops of Kovac's reagent (Sigma-Aaldrich, 4-(Dimethylamino) benzaldehyde solution) were added and the tubes were observed for 2 minutes for the presence of a pink ring, indicating a positive test. If an isolate was citrate (-) and indole (+), a single colony was inoculated in 5 mL of LB broth and incubated at 37°C for 18 hours with shaking. All bacterial cultures were stored in 50% glycerol at -80°C for long-term preservation.

As of October 2021, a total of 496 *E. coli* have been collected from 18 Saskatchewan colibacillosis outbreaks. Systemic isolates have been recovered from 18 outbreaks and cecal isolates from the same flocks have been recovered from 14/18 outbreaks. A set of 96 strains were selected for both Nanopore and Illumina sequencing based on their outbreak, disease status, cellulose and curli production, and colony morphology. All 96 have been sequenced on the Nanopore MinION. 10/96 strains failed Illumina quality control and were not sequenced. These will be repeated at a later date. All 96 strains have been tested for their phenotypic resistance to 27 different antibiotics (see chapter 3).

A2. Biofilm Testing: Cellulose and Curli Tests

E. coli isolates were streaked on LB agar from a 50% glycerol freezer stock and incubated overnight at 37°C. A single colony was subsequently inoculated in 5 mL of LB broth overnight at 37°C with shaking. 4 uL of *E. coli* culture was spotted onto Congo Red Agar (20 μg/mL of Congo Red dye in 1% tryptone, 0.5% yeast extract, 1.5% agar; pH 7.2-7.4) and Calcofluor White Agar (200 mg/L fluorescent brightener #28 (Sigma-Aldrich) in 1% tryptone, 0.5% yeast extract, 1.5% agar) to visualize curli and cellulose production (Solano *et al*., 2002). Plates were incubated at 28°C for 5 days before identifying rdar morphotype and cellulose (**Fig A2**). Isolates were selected for sequencing based on their colony morphology, rdar morphotype, and cellulose production.

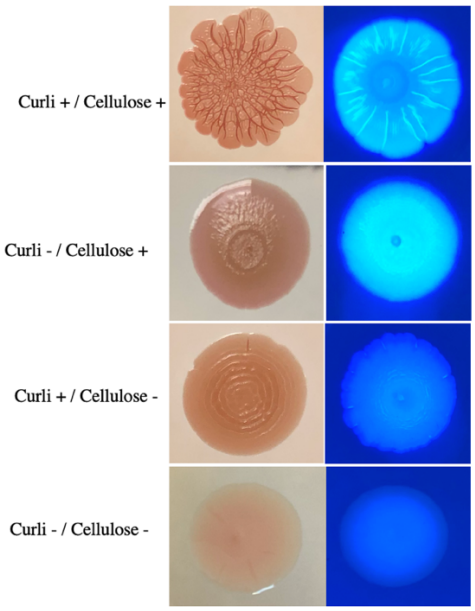

Figure A2. *E. coli* rdar morphotypes (left) and positive and negative cellulose production (right) in avian *E. coli* isolated from Saskatchewan broilers.

A3. References.

Bolton, F.J., Crozier, L., Williamson, J.K. (1996). Isolation of *Escherichia coli* O157 from raw meat products. *Lett. Appl. Microbiol.* 23(5): 317-321. doi: 10.1111/j.1472-765X.1996.tb00198.x.

Hicks, M.J., Ryan, K.J. (1976). Simplified scheme for identification of prompt lactose-fermenting members of the *Enterobacteriaceae*. *J. Clin. Microbiol.* 4(6): 511-514.

Niemi, R.M., Mentu, J., Siitonen, A., Niemela, S.I. (2003). Confirmation of *Escherichia coli* and its distinction from *Klebsiella* species by gas and indole formation at 44 and 45°C. *J. Appl. Microbiol.* 95: 1242-1249. doi: 10.1046/j.1365-2672.2003.02125.x.

Solano, C., Garcia, B., Valle, J., Berasain, C., Ghigo, J., GAmazo, C., Lasa, I. (2002). Genetic analysis of Salmonella enteritidis biofilm formation: critical role of cellulose. *Mol. Microbiol.* 43(3): 793-808. doi: 10.1046/j.1365-2958.2002.02802.x.

Vaughn, R.H., Osborne, J.T., Wedding, G.T., Tabachnick, J. (1950). *J. Bacteriol.* 60(2): 119-127.

Van Hofwegen, D.J., Hovde, C.J., Minnich, S.A. (2016). Rapid evolution of citrate utilization by *Escherichia coli* by direct selection requires *citT* and *dctA*. *J. Bacteriol.* 198(7): 1022-1034. doi: 10.1128/JB.000831-1

APPENDIX B

Table B1. A summary of the sample names, DNA concentrations, and reads generated from the Illumina HiSeq 400 run. Samples that failed QC and did not proceed to sequencing are highlighted in red. These will be repeated at a later date. A total of 86 samples were successfully sequenced and generated between 9 and 23 million reads per sample.

Sample Name	Illumina Barcode	i7 Index Sequence	Input DNA (ng/uL)	HiSeq Lane	Reads Generated (Million)
3862-H1	A1	TTACCGAC	7.1	1	18.2
4957-3S1	A2	AGTGACCT	7.9	1	19.3
23315-H1	A3	TCGGATTC	0.4	--	--
4957-C1	A4	CAAGGTAC	7.8	1	23.6
6041-C2	A5	TCCTCATG	7.2	1	17.3
23315-C4	A6	GTCAGTCA	8.3	1	15.6
6245-2L1	A7	CGAATACG	2.6	1	16.0
6245-1S2	A8	TCTAGGAG	6.8	1	20.0
6245-1H1	A9	GCGAACTA	7.8	1	14.2
6245-3L1	A10	CGTATCTC	14.1	1	14.1
9226-1L2	A11	GTACACCT	8.8	1	16.4
9226-1S1	A12	CGGCATTA	6.4	1	16.5
9226-2S1	B1	TCGTCTGA	6.2	1	18.5
9226-2L1	B2	AGCCTATC	7.6	1	15.2
9226-3H1	B3	CTGTACCA	2.4	1	20.5
9226-3L1	B4	AGACCTTG	2.5	1	15.3
6245-C4	B5	AGGATAGC	45.6	1	13.1
6245-3H2	B6	CCTTCCAT	54.4	1	15.7
0012-1H1	B7	GTCCTTGA	34.2	1	13.1
0012-2L1	B8	TGCGTAAC	84.0	1	14.2
0012-3S1	B9	CACAGACT	88.0	1	14.9
0012-C1	B10	TTACGTGC	52.2	1	14.4
0012-C5	B11	CCAAGGTT	18.5	1	13.3
0012-C7	B12	CACGCAAT	23.0	1	13.4
9619-1H2	C1	TTCCAGGT	28.8	1	14.5
9619-2L1	C2	TCATCTCC	25.0	1	13.2
9619-3S1	C3	GAGAGTAC	10.7	1	13.4
9619-C1	C4	GTCGTTAC	34.2	1	14.0
9619-C6	C5	GGAGGAAT	9.3	1	14.4
9619-C8	C6	AGGAACAC	22.2	1	13.4
0205-1L3	C7	CAGTGCTT	10.6	1	16.5
0205-2L2	C8	CTTGCTAG	15.1	1	14.3
23315-C8	C9	TGGAAGCA	2.2	2	15.2

23315-C5	C10	AGCTAAGC	8.0	2	12.5
23315-C9	C11	GAACGGTT	12.1	2	16.9
23315-L3	C12	GGAATGTC	6.0	2	15.9
3862-C9	D1	TACGGTCT	23.8	2	14.7
3862-S2	D2	CCAGTATC	24.4	2	19.0
3862-C11	D3	TCTACGCA	9.2	2	18.5
3862-C8	D4	GTAACCGA	8.3	2	17.0
3826-S1	D5	GACGTCAT	8.3	2	16.4
3862-L1	D6	CTTACAGC	7.6	2	11.5
3862-C2	D7	TCCATTGC	4.4	2	14.7
3862-C6	D8	AGCGAGAT	18.4	2	13.4
9413-2H1	D9	CAATAGCC	15.9	2	13.6
9413-C4	D10	AAGACACC	4.1	2	17.2
9413-C1	D11	CCAGTTGA	5.2	2	13.2
9413-3S1	D12	TGGTGAAG	10.6	2	13.1
9413-1H1	E1	AAGACGT	1.0	--	--
9413-2S2	E2	TTGCGAGA	1.3	--	--
9413-1S2	E3	GCAATTCC	1.4	--	--
6041-1H1	E4	GAATCCGT	0.6	--	--
6041-3S1	E5	CCGCTTAA	0.6	--	--
6041-C6	E6	TACCTGCA	0.7	2	14.5
6041-C9	E7	GTCGATTG	0.5	--	--
6041-2L1	E8	TATGGCAC	0.6	--	--
6245-1L1	E9	CTCGAACA	18.8	2	14.2
6245-C6	E10	CAACTCCA	30.4	2	17.5
6245-2H3	E11	GTCATCGT	51.4	2	16.0
6245-C1	E12	GGACATCA	21.8	2	12.1
6245-C2	F1	CAGGTTCA	33.0	2	18.1
6245-3S1	F2	GAACGAAG	21.4	2	12.0
4957-C5	F3	CTCAGAAG	70.8	2	15.1
4957-2L1	F4	CATGAGCA	19.6	2	17.9
4957-C3	F5	GACGAACT	5.6	3	19.4
4957-1H1	F6	AGACGCTA	1.5	3	17.5
4957-2L3	F7	ATAACGCC	4.0	3	12.9
4957-1H2	F8	GAATCACC	2.2	3	9.8
4957-3S3	F9	GGCAAGTT	3.3	3	15.7
4957-C6	F10	GATCTTGC	1.5	3	11.6
0205-C9	F11	CAATGCGA	1.8	3	11.9
0205-3H1	F12	GGTGTACA	2.9	3	14.5
0205-C3	G1	TAGGAGCT	4.2	3	18.0
0205-C7	G2	CGAATTGC	1.3	3	12.7
0205-3S1	G3	GTCCTAAG	3.7	3	17.0
2402-1H1	G4	CTTAGGAC	2.2	3	15.8
2402-2L1	G5	TCCACGTT	7.5	3	17.2
2402-3S1	G6	CAACACAG	2.2	3	13.8

2402-C2	G7	GCCTTAAC	1.2	3	10.5
2402-C3	G8	GTAAGGTG	1.6	3	17.1
2402-C4	G9	AGCTACCA	0.9	--	--
7578-2H3	G10	CTTCACTG	2.8	3	14.5
7578-1L2	G11	GGTTGAAC	1.2	3	16.2
7578-2L2	G12	GATAGCCA	2.5	3	17.5
7578-2H2	H1	TACTCCAG	2.7	3	16.0
7578-1H1	H2	GGAAGAGA	1.2	3	19.1
23315-S2	H3	GCGTTAGA	5.1	3	16.6
9619-3L1	H4	ATCTGACC	1.6	3	20.6
9619-2S1	H5	AACCAGAG	1.9	3	19.8
9226-C1	H6	GTACCACA	1.5	3	17.9
9226-C2	H7	GGTATAGG	3.8	3	18.2
9226-C5	H8	CGAGAGAA	2.2	3	18.4
9226-2H2	H9	CAGCATAC	2.2	3	22.7
6041-1L2	H10	CTCGACTT	3.6	3	22.3
6041-3L1	H11	CTTCGGTT	0.6	--	--
6245-C5	H12	CCACAACA	6.5	3	19.9

Table B2. The mean $\log_{10}$ density (CFU) of starting inoculum from planktonic and biofilm cells before disinfectant challenge.

S.No	Strain ID	Starting cells (Mean log 10 ±SD)	
		Planktonic cells	Biofilm cells
1	9226-B3H1	7.68 ± 0.35	8.49 ± 0.63
2	4957-B2L3	7.57 ± 0.65	6.84 ± 0.21
3	6245-B2H3	7.57 ± 0.31	7.90 ± 0.55
4	9619-H12	7.45 ± 0.30	6.86 ± 0.05
5	6041-B3L1	7.48 ± 0.29	8.18 ± 0.46
6	9619-L31	7.45 ± 0.41	8.47 ± 0.32
7	9413-B1S2	7.24 ± 0.54	7.11 ± 0.29
8	6245-C2	7.33 ± 0.42	8.04 ± 0.45
9	4957-C5	7.32 ± 0.41	7.07 ± 0.37
10	6245-C5	7.56 ± 0.36	7.85 ± 0.54
11	0012-C1	7.21 ± 0.17	7.86 ± 0.59
12	6041-C2	7.45 ± 0.30	8.22 ± 0.08

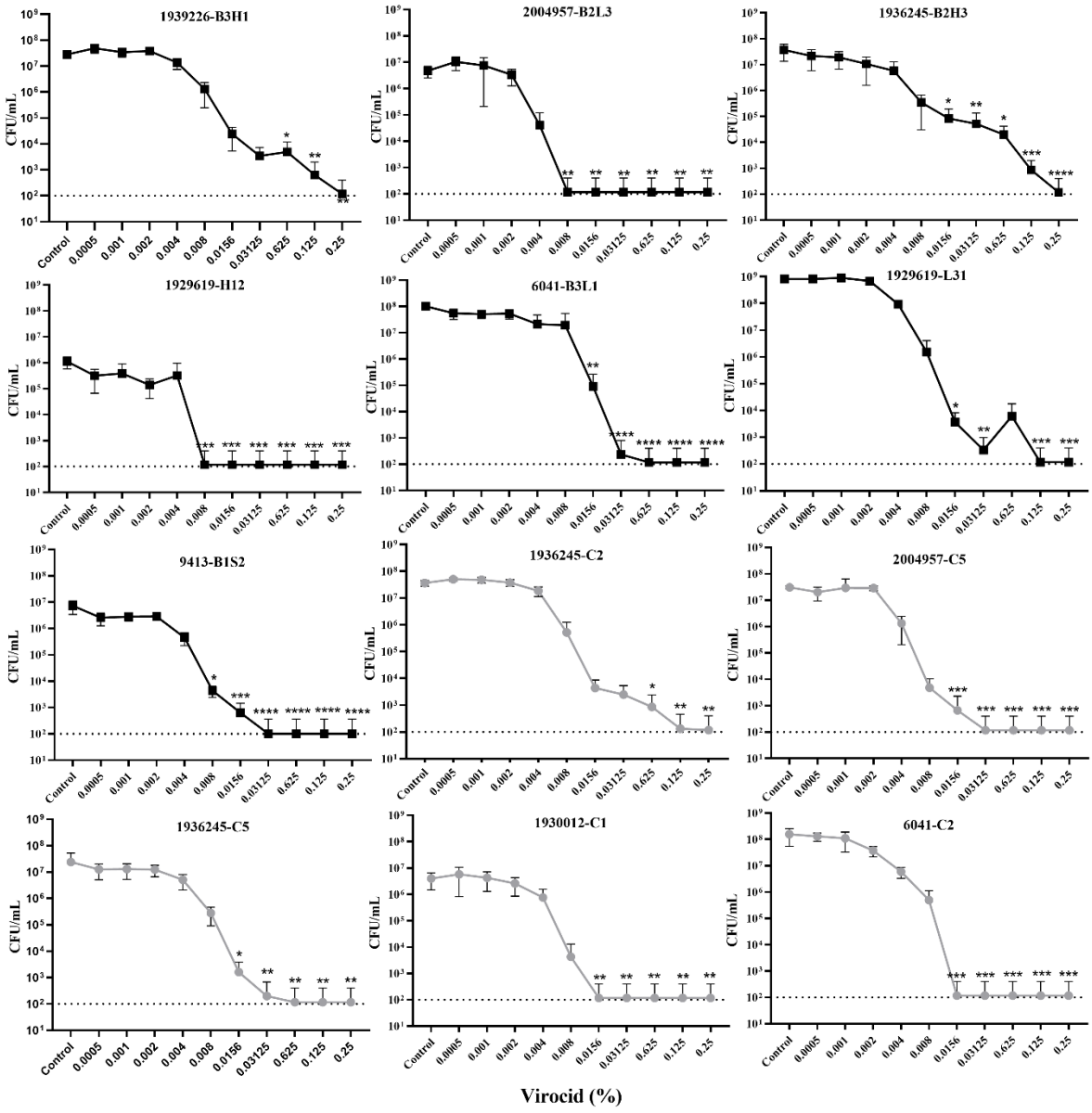

Figure B1. Determination of viable cells after Virocid treatment on systemic (Black line) and cecal (Grey line) *E. coli* biofilms. Viable cells from each peg (n = 6) were enumerated after biofilm growth for 24 h and following Virocid exposure for 30 min. Symbols on the graph represent the mean ± SD from three independent experiments. Statistical significance is represented as follows: *P < 0.05; **P < 0.005; ***P < 0.0005; ****P < 0.0001.

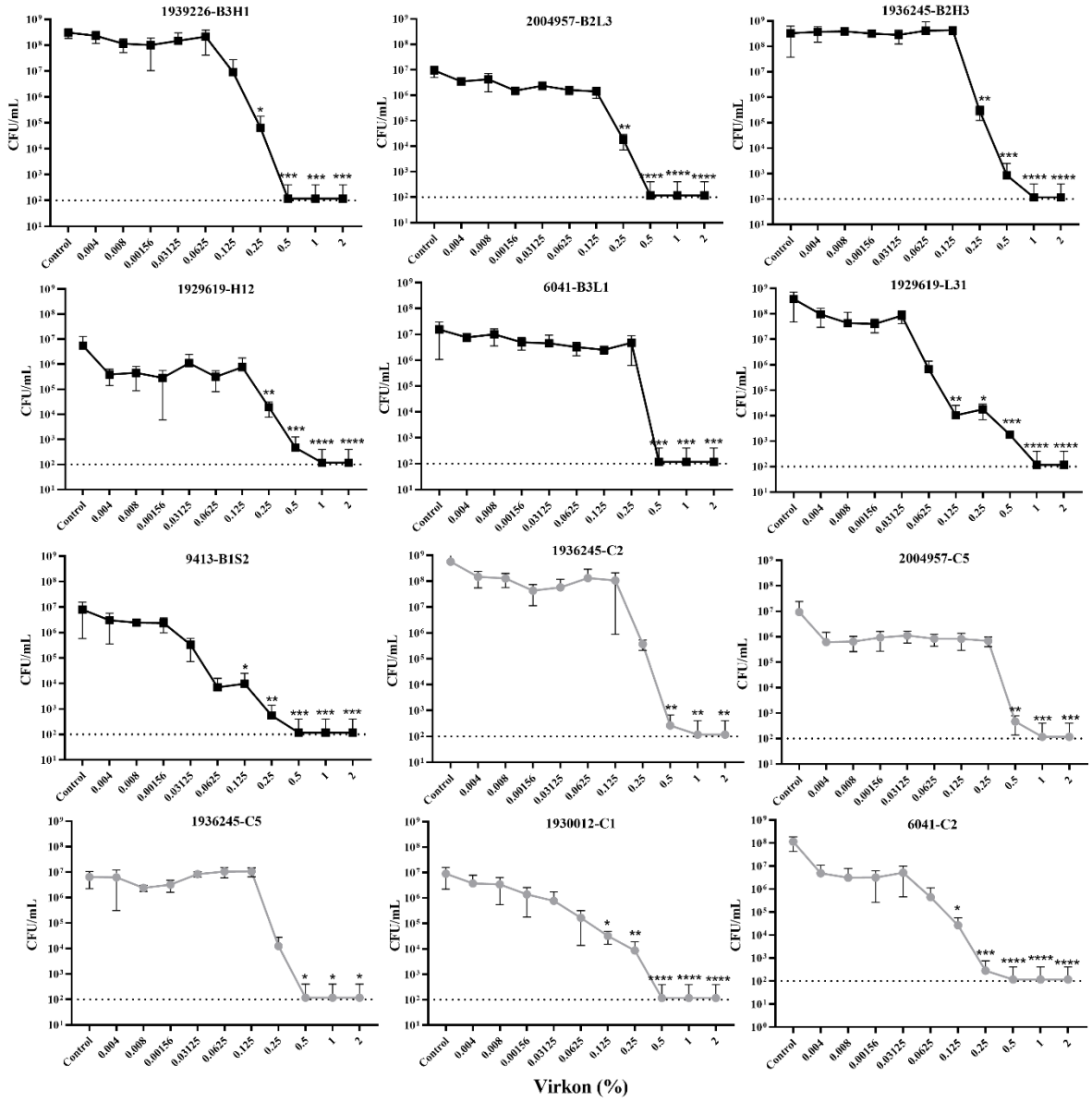

Figure B2. Determination of viable cells after Virkon treatment on systemic (Black line) and cecal (Grey line) *E. coli* biofilms. Viable cells from each peg (n = 6) were enumerated after biofilm growth for 24 h and following Virkon exposure for 30 min. Symbols on the graph represent the mean ± SD from three independent experiments. Statistical significance is represented as follows: *P < 0.05; **P < 0.005; ***P < 0.0005; ****P < 0.0001.

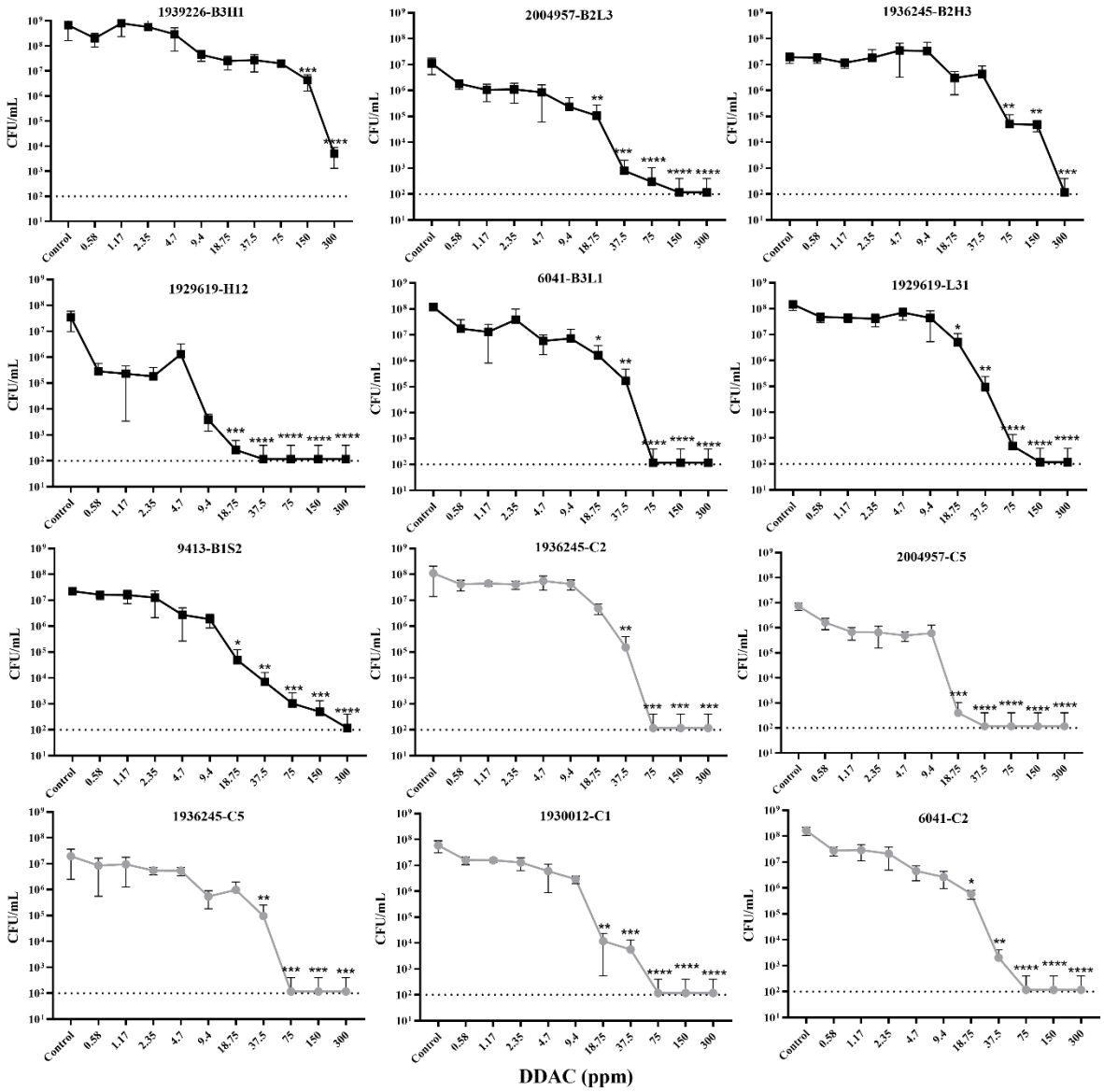

Figure B3. Determination of viable cells after DDAC treatment on systemic (Black line) and cecal (Grey line) *E. coli* biofilms. Viable cells from each peg (n = 6) were enumerated after biofilm growth for 24 h and following DDAC exposure for 30 min. Symbols on the graph represent the mean ± SD from three independent experiments. Statistical significance is represented as follows: *P < 0.05; **P < 0.005; ***P < 0.0005; ****P < 0.0001.

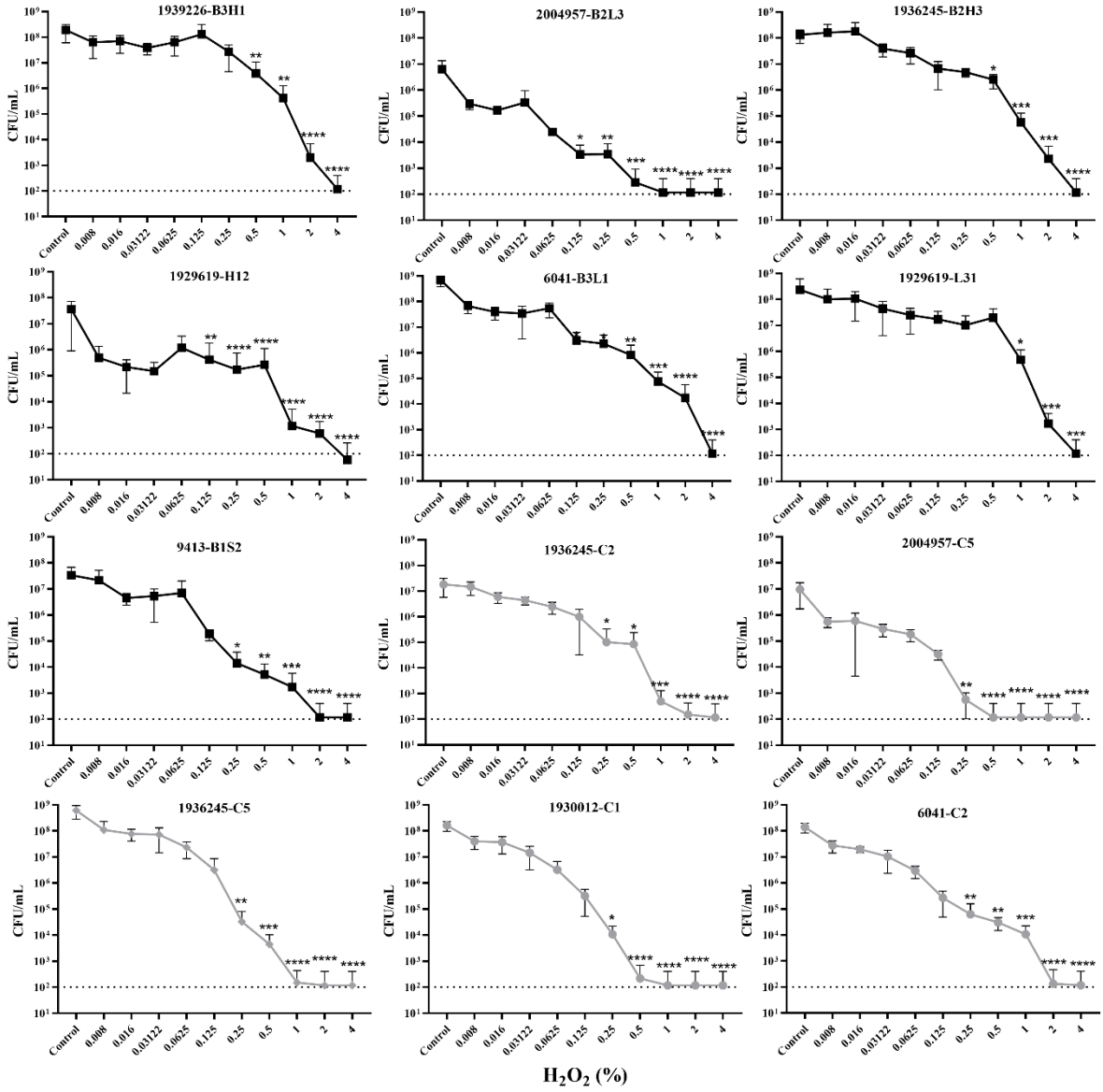

Figure B4. Determination of viable cells after H_2O_2 treatment on systemic (Black line) and cecal (Grey line) *E. coli* biofilms. Viable cells from each peg (n = 6) were enumerated after biofilm growth for 24 h and following H_2O_2 exposure for 30 min. Symbols on the graph represent the mean ± SD from three independent experiments. Statistical significance is represented as follows: *P < 0.05; **P < 0.005; ***P < 0.0005; ****P < 0.0001.

Table B3. Summary of disinfectants and their concentrations used in the study

Disinfectants	**Virocid**	**Virkon™**	**(DDAC)**	**H_2O_2**
Manufacturer/Supplier	CID LINES N.V., Ieper, Belgium.	Antec International LTD., Suffolk, UK.	KleenGlow, PACE Chemicals LTD., BC, Canada.	Plant Life Products, BC, Canada.
Group of Ingredients	Two different quaternary ammonia, modern aldehyde, alcohol, pineoil and other ingredients	Pentapotassium bis(peroxymonosulphate) bis(sulphate), Sodium Dodecylbenzene Sulfonate, Butanedioic acid, 2-hydroxy-sulphamic acid, Potassium hydrogen sulphate, Sodium chloride, Dipotassium peroxodisulphate, Dipotassium disulphate, Dipentene.	Isopropyl Alcohol, Didecyldimethyl Ammonium Chloride (DDAC) and Ethanol	H_2O_2 and Water
Active Ingredients	Alkyl dimethyl benzyl ammonium chloride (17.06%); Didecyl dimethyl ammonium chloride (7.8%); Glutaraldehyde (10.8%) and other ingredients (64.415%).	Potassium monopersulfate (21.4%)	DDAC (7.5%)	H_2O_2 (29%) technical grade
pH	6.5	5.0 – 6.5	6.0 – 7.0	4.7
Recommended Conc.	0.25%	1%	300 ppm	Not Available
Recommended contact time	10-15 min	10 min	Not Available	Not Available
Concentrations and exposure time used in this study				
Concentration used on Planktonic cells for 18 h[#]	0.0125% – 0.000025%	1% – 0.002%	30ppm – 0.06ppm	0.5% – 0.001%
Concentration used on Planktonic or biofilm cells for 30 min	0.25% – 0.0005%	2% – 0.004%	300ppm – 0.6ppm	4% – 0.008%

[#] MIC and MBC of disinfectants were determined by following CLSI guidelines (CLSI, 2020).