DISSERTATION

# INFLUENCE OF MANAGEMENT PRACTICES ON VIRULENCE FACTORS, ANTIMICROBIAL RESISTANCE GENES AND HEAVY METAL RESISTANCE GENES IN BROILER CHICKEN PRODUCTION

Submitted by

Reed Richard Woyda

Graduate Degree Program in Cell and Molecular Biology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Spring 2023

Doctoral Committee

Advisor: Zaid Abdo

Adelumola Oladeinde
Josh Daniels
Dan Sloan
Mark Stenglien

ABSTRACT

INFLUENCE OF MANAGEMENT PRACTICES ON VIRULENCE FACTORS, ANTIMICROBIAL RESISTANCE GENES AND HEAVY METAL RESISTANCE GENES IN BROILER CHICKEN PRODUCTION

The main bacterial species associated with food-borne illness in humans are *Escherichia coli*, *Salmonella* species and *Campylobacter* species. The ability of a bacterial strain to survive the food-production pipeline and to mount an infection and cause disease in humans is dependent on an array of genetic factors. The presence of specific virulence factors will influence the severity of disease while antimicrobial resistance genes affect the choice and efficacy of treatment. Management practices in poultry production aim at reducing the incidence of poultry and human bacterial pathogens and, in general, at maintaining a healthy flock and a healthy global population. However, the influence of management practices, in a post-antibiotic era, on pathogenic bacterial species, and in particular the selective pressures imposed on genetic factors such as antimicrobial and metal resistance and virulence factors, are understudied.

In Chapter 2, we provide a robust bacterial genomic analysis pipeline which is used for subsequent analysis in the following chapters. Chapter 3 provides an understanding of the current antimicrobial resistance and virulence factors present in chicken production and human clinical settings. This work found these host sources harbored different antimicrobial resistance genes and virulence factors that can classify them into phylogroups and host origin. In Chapter 4, through characterization of *Campylobacter* species isolated from broiler litter, we determined the reused litter environment selected for *Campylobacter* species lacking virulence factors aiding in

colonization of chicken and human hosts. In Chapter 5, we determined the practice of adding copper sulfate to drinking water, commonly used for growth promotion or sanitization, may have selected for, and provided a reservoir for, *Salmonella* strains harboring plasmid-borne copper resistance genes.

Overall, this work provides a computational pipeline for the high-throughput analysis of bacterial genomes and provides insights into selective pressures imposed on pathogenic bacterial species by modern-day management practices.

# ACKNOWLEDGMENTS

I would like to express my gratitude, and thanks to my partner for her endless support and love. To my dog, Huxley, for entertaining breaks while writing this dissertation. To my mother and father for endlessly supporting my educational path and for my mother's bountiful treat-filled packages sent from home. I would also like to acknowledge my advisor and mentor, Dr. Zaid Abdo and Dr. Ade Oladeinde, for the superb guidance, advise and direction they have provided me over the last four years.

Finally, I would like the thank all the Abdo Lab members, present and past, and the Cell and Molecular Biology program students and faculty for their support. I am grateful to all the lab personnel at the USDA ARS for their support throughout my time collaborating with them.

# DEDICATION

I would like to dedicate this dissertation to my grandmother and grandfather who have instilled in me a passion for lifelong learning and for providing a path towards this dissertation.

TABLE OF CONTENTS

# CHAPTER 1: INTRODUCTION

## 1.1 Chicken production and foodborne illness

Consumption of chicken meat is highly associated with foodborne illness in the United States (Dewey-Mattia *et al.*, 2018). *Escherichia coli* (*E. coli*), *Campylobacter* species and *Salmonella* species are leading causes of bacterial-derived food-borne illness in humans (Epps *et al.*, 2013; Dewey-Mattia *et al.*, 2018). Severity of disease and efficacy of treatment for these pathogens depends on many factors such as: 1) the presence of one or more antimicrobial resistance genes (ARGs) conferring resistance to antimicrobials critically important to human medicine (World Health Organization, 2019) and 2) presence, and combinations of, virulence factors (VFs) which affect a strains' ability to infect and cause disease in a given host. Management practices, specifically those within the broiler house, are generally in place to produce safe chicken meat for consumer consumption. However, these practices (e.g. administration of antibiotics, feed/water additives, acidifier treatment to litter, temperature and humidity control of the broiler house, half house brooding, litter reuse, house cleaning procedures, etc.) influence the bacterial populations not only within broilers, but within the litter as well (Pope and Cherry, 2000; Cressman *et al.*, 2010; Roberts *et al.*, 2013; Scott *et al.*, 2018; Fan *et al.*, 2022; Oladeinde *et al.*, 2023). Thus, management practices impose selective pressures on individual bacterial species and therefore selective pressures are also acting at the individual gene level i.e., at the level of ARGs and VFs as well as genes regulating environmental factors such as heavy metals. Understanding the impact that chicken production management practices have on foodborne

pathogens, as well as their repertoire of ARGs, VFs and heavy metal resistance genes, is crucial to protecting both the health of humans and food-producing animals

This research aims to evaluate pathogenic bacteria present in broiler chicken production within the United States and to characterize the genetic diversity of bacteria isolated from peanut-hull-based litter during consecutive litter reuse. Additionally, this research aims to understand how management practices and environmental conditions influence both pathogenic bacterial populations as well as the genetic material harbored by individual strains e.g., antimicrobial resistance genes, virulence factors and heavy metal resistance genes).

## 1.2 Poultry litter as a selective agent and a medium for monitoring pathogenic bacteria

Poultry litter is the material which is placed on the floors of broiler houses, used to raise chickens for chicken meat, and is an initial mixture of a bedding medium which accumulates chicken feces, feathers, excess feed, bodily secretions, and water. In the United States, it is a common management practice in broiler chicken production to raise multiple flocks of birds on the same litter. However, Canadian (*Code of practice for the care and handling of hatching eggs, breeders, chickens, and turkeys.*, 2016) and European agencies (ADAS, 2016) both advocate for fresh litter for each flock. These restrictions are, in part, due to food-borne pathogens such as *Salmonella*, *Campylobacter* and *E. coli* being detected in litter (Mellata *et al.*, 2003; Line and Bailey, 2006; Dziva and Stevens, 2008; Vaz *et al.*, 2017; Valeris-Chacin *et al.*, 2021). In opposition to the regional recommendations in Canada and Europe are findings of pathogens being reduced or maintained at constant levels with consecutive litter reuse (Roll, Dai Prá and Roll, 2011; Brooks *et al.*, 2016). Additionally, litter reuse is economically beneficial to farmers,

as fresh litter does not need to be purchased for each new flock, it can be used as a fertilizer, and can even supplement cattle feed in some cases (Jay Daniel, K.C. Olson, 2005; Yamak *et al.*, 2015).

Common litter mediums in the United States include wood shavings and rice hulls and the choice is impacted by cost and regional availability (Diarra *et al.*, 2021). For example, Georgia, USA, is a leading producer of both broiler chicken meat and peanuts which has resulted in many farmers in this region choosing to use the residual peanut hulls as litter bedding material (Georgia Peanut Commission, 2021; USDA, National Agricultural Statistics Service, 2021). Studies on the microbial ecology of chicken litter under a re-utilization regime are limited to historically common litter mediums such as wood shavings and rice hulls (Kelley *et al.*, 1995; Roll, Dai Prá and Roll, 2011; Chinivasagam, Tran and Blackall, 2012; Wei *et al.*, 2013; Muniz MV, MSc *et al.*, 2014; Smith *et al.*, 2016; Rauber Würfel *et al.*, 2019; Valeris-Chacin *et al.*, 2021, 2022). In general, these studies aimed to determine if the usage of specific litter types correlate with changes in bird weight, feed-conversion ratio, morbidity, mortality or in pathogenic bacterial populations. However, no research has been performed evaluating these factors within broiler chicken production utilizing peanut hull-based litter. This research aims to understand the pathogenic bacterial populations present within peanut hull-based litter and to determine the influence of consecutive litter reuse on pathogenic bacterial species as well as their genetic contents.

## 1.3 Foodborne pathogens in poultry production

Poultry litter has been extensively shown to harbor many microorganisms including chicken commensal microbiota (i.e. bacteria which reside within or on the host and do not cause harm),

environmental bacteria, bacteria harboring antimicrobial resistance, and pathogenic strains capable of food-borne illness in humans (Schefferle, 1965; Kelley *et al.*, 1995, 1998; Hofacre *et al.*, 2000; Terzich *et al.*, 2000; Cressman *et al.*, 2010; Brooks *et al.*, 2016; Valeris-Chacin *et al.*, 2021). While the fresh litter contains its own initial bacterial populations, composed mainly of harmless environmental bacteria, litter reuse has been shown to accumulate an abundance of bacteria from the chicken intestinal tract (Cressman *et al.*, 2010; Brooks *et al.*, 2016). *Escherichia coli, Salmonella* and *Campylobacter* are all considered commensal bacteria within a chicken host (Humphrey, 2006; Kaesbohrer *et al.*, 2012; de Perio *et al.*, 2013). While *Salmonella* and *Campylobacter* are both unlikely to cause clinical disease in chickens, *E. coli*, specifically a group designated avian pathogenic *E. coli* (APEC), are capable of causing avian colibacillosis; a major endemic disease within the poultry industry (Dziva and Stevens, 2008). In addition to being the causative agent of collibacilosis in chickens, *E. coli* are a leading cause of foodborne disease outbreaks in the United States and highly associated with the consumption of chicken meat (Lynch *et al.*, 2006; Dewey-Mattia *et al.*, 2018). Different strains of these bacterial pathogens may also be introduced into a broiler house by various modes (Golden, Rothrock and Mishra, 2021). As these particular bacteria are considered commensal, strains may be introduced into a broiler house when a new flock of birds arrives from the hatchery. Both *Salmonella* and *Campylobacter* have been detected in transport cages of young chickens (Zancan *et al.*, 2000; Berrang, Northcutt and Cason, 2004). Berrang and colleagues also determined that *Campylobacter* detected from fecal samples in transport cages were capable of spreading to consecutive flocks (Berrang, Northcutt and Cason, 2004). Alternative to introduction of new strains from incoming birds, research has shown bacteria such as *Campylobacter* can persist through broiler house cleaning procedures and therefore persist within subsequent flocks (Eberle

*et al.*, 2012). Therefore, it is crucial to monitor for pathogenic bacterial species in consecutive flocks as well as along the production pipeline i.e., from the hatchery, through growth in broiler houses and through downstream processing.

## 1.4 Antimicrobials and antibiotic alternatives in food-producing animals

Antibiotics have been utilized in animal agriculture since the 1940s and 1950s, but their role has changed dramatically over time (Kirchhelle, 2018). Antibiotics became widely used after the discovery of penicillin and the subsequent discovery of streptomycin in 1943 (Woodruff, 2014) and were widely used in food-producing animals to cure infection and to promote growth. However, soon after the introduction of antibiotics came reports of bacteria capable of surviving antibiotic treatment, this was termed antibiotic resistance (Rammelkamp and Maxon, 1942). By 2017, the US National Antimicrobial Resistance Monitoring System reported that 14% of all tested *Salmonella* isolates from retail chicken meat were resistance to more than 3 classes of antibiotics, termed multidrug resistant (MDR) (U.S. Department of Health and Human Services, 2019). The risk of resistance to medicines critically important to human medicine is still a threat today. There was a 12% increase in the total amount of medically important antimicrobials used in chicken production from 2020 to 2021 (measured in kg of active ingredient), even though there has been a dramatic decrease over the last 5 years (69% decline from 2016 to 2021) (United States Food and Drug Administration, Center for Veterinary Medicine, 2022). Similarly, a decline has been observed for not medically important antimicrobials (42% decline from 2016 to 2021). However, even though a dramatic decline has been observed in the usage of antibiotics in food-producing animals, the prevalence of antibiotic resistant bacteria, both within the hospital

setting and those isolated from retail meat and food-producing animals is high. In 2011, 94% of *Salmonella* infections in the United States were estimated to be foodborne (Scallan *et al.*, 2011) and, within a similar timeframe (2002-2015), 58% were resistant to at least one antimicrobial (U.S. Department of Health and Human Services, 2019). From 2019 to 2020 there has been a 15% increase in resistant infections starting during hospitalization, in part due to the COVID-19 pandemic (CDC, 2022). Additionally, the COVID-19 pandemic has delayed reporting of antimicrobial resistance threats, in particular for drug-resistant nontyphoidal *Salmonella* and drug-resistant *Campylobacter* (CDC, 2022). Thus, while the FDA reports for humans indicate relatively unchanged levels of resistance for *Salmonella* and only a 5% increase for ciprofloxacin-resistant *Campylobacter,* the strain on the healthcare system from the COVID-19 pandemic has resulted in an increase in healthcare-associated infections and supply chain difficulties (Strathdee, Davies and Marcelin, 2020; FDA, 2022). Taken together, antibiotic resistance genes, while experiencing less selective pressure due to a decline in the usage of antimicrobials, are persisting through food production and subsequently proving detrimental to treatment of infections in humans.

Antibiotic alternatives such as probiotics, prebiotics and heavy metals are in use today to combat the global threat of antimicrobial resistance while attempting to gain the same benefits which antibiotics provided e.g., prevention of disease and growth promotion. However, similar to that of antimicrobials, rigorous research needs to be performed to understand the long-term impact of these alternatives. For example, a study reported that a probiotic strain of *Bacillus subtilis* isolated from a commercial probiotic feed was found to harbored both the *vanA* (contributes to glycopeptide resistance) and *tetW* (contributes to tetracycline resistance) genes (Wannaprasat *et al.*, 2009). These antimicrobial resistance genes could prove detrimental if

introduced into pathogenic species through horizontal gene transfer. Additionally, heavy metals such as copper are frequently added to poultry diets as an alternative to an antibiotic growth promoter or added to waterlines for sanitation purposes (Scott *et al.*, 2018). Multiple studies have indicated the co-occurrence of antibiotic resistance genes and metal resistance genes, highlighting the potential risk associated with the use of heavy metals and their associated risk to human health (Li, Xia and Zhang, 2017; Bombaywala, Purohit and Dafale, 2021; Li *et al.*, 2022). As heavy metals may act as selective agents in the selection and propagation of ARGs, reducing antibiotic usage within poultry production may not be sufficient to reduce their prevalence in foodborne pathogens (Baker-Austin *et al.*, 2006; Seiler and Berendonk, 2012).

## 1.5 Mechanisms of acquiring resistance

Antibiotic resistance may be natural or acquired. Natural resistance includes resistance which is always expressed within a species or that which is induced when a species is exposed to an antimicrobial (Cox and Wright, 2013). For example, gram-positive bacteria have a thick but permeable outer layer which leads to an intrinsic resistance to some antibiotics (Randall *et al.*, 2013). Alternatively, gram-negative bacteria have an additional outer membrane consisting of tightly packed fatty acids resulting in low permeability, and therefore increased resistance, to various antibiotics (Delcour, 2009). Acquired resistance occurs through various mechanisms which allow the sharing, or uptake, of genetic material from the same or, in some cases, different species. Mechanisms of acquired resistance are grouped under the term horizontal gene transfer (HGT) and consist of: 1) transformation (uptake of foreign DNA from the environment), 2) transposition (movement of DNA between chromosomal DNA to plasmids), and 3) conjugation (DNA transfer from a donor to a recipient bacteria). Many studies have shown that HGT plays an

instrumental role in the dissemination of antimicrobial resistance (Colavecchio *et al.*, 2017). It has been demonstrated *in vitro* that antimicrobial resistance genes, carried on an IncK2 plasmid, can be transferred between a *Salmonella enterica* serovar Heidelberg strain, isolated from a chicken carcass, and *E. coli*. (Oladeinde *et al.*, 2019). While this study was performed *in vitro*, it highlights the ability of *Salmonella* Heidelberg to obtain antimicrobial resistance from *E. coli*, a common bacterium of the chicken gastrointestinal tract.

## 1.6 Research aims and objectives

Broadly, this research aims to further understand food-borne pathogens within broiler chicken production in the United States. This was done through: 1) an in-depth study of the antimicrobial resistance genes present in *E. coli* isolates from both poultry production and the human clinical setting and 2) through a longitudinal sampling of poultry litter at a commercial scale to determine changes in bacterial pathogen populations and their genetic contents.

To perform an in-depth genomic characterization of a microbial population, a high-throughput, robust bioinformatics tool was needed. Accordingly, this research encompasses the creation of an analysis pipeline, Reads2Resistome (Chapter 2). Reads2Resistome generates a comprehensive description of bacterial genomes and their antimicrobial resistance genes and virulence factors.

To understand the current genetic factors contributing to virulence and antimicrobial resistance, which are present in both chicken production and human clinical settings (i.e. those obtained from humans in hospital settings with *E. coli* infections) within the United States, we compared the genetic relatedness of *E. coli* isolates from these settings on a large scale (Chapter 3). Specifically, this aimed to answer the following questions: 1) how do isolates from these two

populations differ with respect to their antimicrobial resistance genes and the drug classes to which they confer resistance? 2) How do these isolates differ in their presence of virulence factors? 3) Based on the isolates' antimicrobial resistance and virulence factor profiles, can their origin (i.e., human clinical or chicken production) be determined?

Lastly, we characterized the genetic diversity of the pathogenic species *Campylobacter* (Chapter 4) and *Salmonella* (Chapter 5) in peanut hull-based litter. This includes understanding their prevalence and persistence over multiple flock cycles, as well as characterizing the genetic changes with respect to antimicrobial resistance and virulence. We aimed to answer the following research questions: 1) are these pathogenic species present in peanut-hull based litter and, if so, how does their prevalence change over time? 2) Do these pathogens harbor antimicrobial or heavy metal resistance? 3) Do these pathogens carry virulence factors which may aid in their colonization of chicken or human tissues? 4) Is the prevalence of these pathogens, or the genetic factors they harbor, associated with management practices (e.g. litter reuse)?

Overall, this dissertation explores the various types of genetic elements within commensal and pathogenic species isolated from broiler chicken litter and lays the foundation for future studies in these species and beyond.

## 1.7 Significance

The outcomes of this research will benefit broiler chicken farmers, poultry litter purchasers (ie. for fertilizer or animal feed), chicken meat consumers, and academics in the areas of microbial population genetics, microbiology and pathology. This study aimed to contribute to the body of knowledge on alternative litter mediums, specifically the presence and persistence of pathogenic bacteria which may reside in them, to better inform farmers on management practices. Our

research improves our understanding of the pathogenic microbial ecology of an understudied litter medium, peanut hulls, and highlights environmental factors which influence pathogen populations. As litter may be reused and later applied as a fertilizer or used as feed, our research addressed the potential microbial composition within reused litter to better inform downstream management practices. The methodologies employed to compare and contrast the genetic contents of the identified pathogenic isolates is sufficiently general and available for other researchers to employ to study other bacterial populations. Additionally, we made the bioinformatics tool, Reads2Resistome, free and open-source, available to researchers and the public to enable high-throughput analysis of bacterial sequencing data and generation of antimicrobial resistance and virulence factor data. Lastly, our research provides in-depth knowledge on antimicrobial resistance genes, metal resistance genes and virulence factors present in pathogenic species in broiler chicken production under management practices including “No Antibiotics Ever” and peanut hull-based litter.

## 1.8 Limitations and scope

In furthering the collective understanding of the microbial ecology and population genetics of broiler chicken production pathogens, the main objectives have a defined scope. First, our studies focus on the main causative agents of food-borne illness in the United States: *E. coli, Salmonella* species and *Campylobacter* species. Second, in the comparison of *E. coli* isolates across the United States, we limited our sampling populations to a subset of the poultry industry, broiler chicken production, and within humans, the clinical setting. These two populations aid in characterizing *E. coli* isolates only within the specific type of poultry we are interested in, and exclude other food-producing animals, such as turkeys and duck. Third, our longitudinal study,

aimed at characterizing the genetic diversity of pathogenic species in broiler chicken production, focused on two main pathogenic species, *Salmonella* and *Campylobacter* and litter sampling that occurred over 3 litter reuses. Sampling over multiple flock cycles is a key aspect of this research as it elucidates how pathogenic bacterial populations change over time with litter reuse. We used next-generation sequencing to compare isolates' genetic contents with confidence. Our sequencing methodology was limited to individual bacterial isolates; bacterial community studies are beyond the scope of this research.

## 1.9 References

ADAS (2016) *Comparison of the regulatory framework and key practices in the poultry meat supply chain in the EU and USA: a study for the Association of Poultry Processors and Poultry Trade in the EU Countries.* ADAS, Ltd. Available at: https://britishpoultry.org.uk/identity-cms/wp-content/uploads/2018/05/2016-ADAS-EU-US-comparison.pdf.

Ahumada-Santos, Y.P. *et al.* (2020) 'Association of phylogenetic distribution and presence of integrons with multidrug resistance in Escherichia coli clinical isolates from children with diarrhoea', *Journal of Infection and Public Health*, 13(5), pp. 767–772. Available at: https://doi.org/10.1016/J.JIPH.2019.11.019.

Alboukadel Kassambara, Fabian Mundt (2020) 'factoextra: Extract and Visualize the Results of Multivariate Data Analyses'. Available at: https://CRAN.R-project.org/package=factoextra.

Alcock, B.P. *et al.* (2020) 'CARD 2020: Antibiotic resistome surveillance with the comprehensive antibiotic resistance database', *Nucleic Acids Research*, 48(D1), pp. D517–D525. Available at: https://doi.org/10.1093/nar/gkz935.

Andy Liaw, Matthew Wiener (2002) 'Classification and Regression by Randomforest'. Available at: http://CRAN.R-project.org/doc/Rnews/.

Antipov, D. *et al.* (2016) 'HybridSPAdes: An algorithm for hybrid assembly of short and long reads', *Bioinformatics*, 32(7), pp. 1009–1015. Available at: https://doi.org/10.1093/bioinformatics/btv688.

Bagel, S. *et al.* (1999) ‘Impact of gyrA and parC mutations on quinolone resistance, doubling time, and supercoiling degree of Escherichia coli’, *Antimicrobial Agents and Chemotherapy*, 43(4), pp. 868–875. Available at: https://doi.org/10.1128/AAC.43.4.868/ASSET/F739B9CE-E8B5-442A-B569-4072FF1489AA/ASSETS/GRAPHIC/AC0490515001.JPEG.

Baker-Austin, C. *et al.* (2006) ‘Co-selection of antibiotic and metal resistance’, *Trends in Microbiology*, 14(4), pp. 176–182. Available at: https://doi.org/10.1016/j.tim.2006.02.006.

Berrang, M.E., Northcutt, J.K. and Cason, J.A. (2004) ‘Recovery of Campylobacter from Broiler Feces During Extended Storage of Transport Cages’, *Poultry Science*, 83(7), pp. 1213–1217. Available at: https://doi.org/10.1093/ps/83.7.1213.

Bessonov, K. *et al.* (2021) ‘ECTyper: in silico Escherichia coli serotype and species prediction from raw and assembled whole-genome sequence data’, *Microbial genomics*, 7(12). Available at: https://doi.org/10.1099/MGEN.0.000728.

Biswas, K., Sharma, P. and Joshi, S.R. (2019) ‘Co-occurrence of antimicrobial resistance and virulence determinants in enterococci isolated from traditionally fermented fish products’, *Journal of Global Antimicrobial Resistance*, 17, pp. 79–83. Available at: https://doi.org/10.1016/j.jgar.2018.11.012.

Black, R.E. *et al.* (2010) ‘Global, regional, and national causes of child mortality in 2008: a systematic analysis’, *The Lancet*, 375(9730), pp. 1969–1987. Available at: https://doi.org/10.1016/S0140-6736(10)60549-1.

Bolger, A.M., Lohse, M. and Usadel, B. (2014) ‘Trimmomatic: A flexible trimmer for Illumina sequence data’, *Bioinformatics*, 30(15), pp. 2114–2120. Available at: https://doi.org/10.1093/bioinformatics/btu170.

Bombaywala, S., Purohit, H.J. and Dafale, N.A. (2021) ‘Mobility of antibiotic resistance and its co-occurrence with metal resistance in pathogens under oxidative stress’, *Journal of Environmental Management*, 297, p. 113315. Available at: https://doi.org/10.1016/j.jenvman.2021.113315.

Bortolaia, V. *et al.* (2020) ‘ResFinder 4.0 for predictions of phenotypes from genotypes’, *The Journal of antimicrobial chemotherapy*, 75(12), pp. 3491–3500. Available at: https://doi.org/10.1093/JAC/DKAA345.

Boyle, D.P. and Zembower, T.R. (2015) ‘Epidemiology and Management of Emerging Drug-Resistant Gram-Negative Bacteria: Extended-Spectrum β-Lactamases and Beyond’, *Urologic Clinics*, 42(4), pp. 493–505. Available at: https://doi.org/10.1016/J.UCL.2015.05.005.

Breiman, L. (2001) ‘Random Forests’, *Machine Learning*, 45(1), pp. 5–32. Available at: https://doi.org/10.1023/A:1010933404324.

Brooks, J.P. *et al.* (2016) ‘Cultivation and qPCR Detection of Pathogenic and Antibiotic-Resistant Bacterial Establishment in Naive Broiler Houses’, *Journal of Environmental Quality*, 45(3), pp. 958–966. Available at: https://doi.org/10.2134/jeq2015.09.0492.

Campitelli, E. (2022) ‘ggnewscale: Multiple Fill and Color Scales in ggplot2’. Zenodo. Available at: https://doi.org/10.5281/ZENODO.2543762.

Caniça, M. *et al.* (2019) ‘Antibiotic resistance in foodborne bacteria’, *Trends in Food Science and Technology*, 84(August 2018), pp. 41–44. Available at: https://doi.org/10.1016/j.tifs.2018.08.001.

Carattoli, A. *et al.* (2014) ‘In Silico detection and typing of plasmids using plasmidfinder and plasmid multilocus sequence typing’, *Antimicrobial Agents and Chemotherapy*, 58(7), pp. 3895–3903. Available at: https://doi.org/10.1128/AAC.02412-14.

CDC (2019) ‘Surveillance for Foodborne Disease Outbreaks United States , 2014 : Annual Report’, pp. 1–16.

CDC (2022) *COVID-19: U.S. Impact on Antimicrobial Resistance. 2022 Special Report)*. Centers for Disease Control and Prevention. Available at: https://www.cdc.gov/drugresistance/pdf/covid19-impact-report-508.pdf.

Centers for Disease Control and Prevention (2019) ‘Antibiotic resistance threats in the United States’, *Centers for Disease Control and Prevention*, pp. 1–113. Available at: https://doi.org/CS239559-B.

Chinivasagam, H.N., Tran, T. and Blackall, P.J. (2012) ‘Impact of the Australian litter re-use practice on Salmonella in the broiler farming environment’, *Food Research International*, 45(2), pp. 891–896. Available at: https://doi.org/10.1016/j.foodres.2011.06.014.

Clermont, O. *et al.* (2013) ‘The Clermont Escherichia coli phylo-typing method revisited: Improvement of specificity and detection of new phylo-groups’, *Environmental Microbiology Reports*, 5(1), pp. 58–65. Available at: https://doi.org/10.1111/1758-2229.12019.

Clermont, O. *et al.* (2019) ‘Characterization and rapid identification of phylogroup G in Escherichia coli, a lineage with high virulence and antibiotic resistance potential’, *Environmental Microbiology*, 21(8), pp. 3107–3117. Available at: https://doi.org/10.1111/1462-2920.14713.

*Code of practice for the care and handling of hatching eggs, breeders, chickens, and turkeys.* (2016). Council National Farm Animal Care Council. Available at: https://www.nfacc.ca/poultry-code-of-practice.

Colavecchio, A. *et al.* (2017) ‘Bacteriophages contribute to the spread of antibiotic resistance genes among foodborne pathogens of the Enterobacteriaceae family - A review’, *Frontiers in Microbiology*. Available at: https://doi.org/10.3389/fmicb.2017.01108.

Comery, R. *et al.* (2013) ‘Identification of potentially diarrheagenic atypical enteropathogenic Escherichia coli strains present in Canadian food animals at slaughter and in retail meats’, *Applied and Environmental Microbiology*, 79(12), pp. 3892–3896. Available at: https://doi.org/10.1128/AEM.00182-13/SUPPL_FILE/ZAM999104460SO1.PDF.

Cortés, P. *et al.* (2010) ‘Isolation and Characterization of Potentially Pathogenic Antimicrobial-Resistant Escherichia coli Strains from Chicken and Pig Farms in Spain’, *Applied and Environmental Microbiology*, 76(9), pp. 2799–2805. Available at: https://doi.org/10.1128/AEM.02421-09.

Cox, G. and Wright, G.D. (2013) ‘Intrinsic antibiotic resistance: Mechanisms, origins, challenges and solutions’, *International Journal of Medical Microbiology*, 303(6–7), pp. 287–292. Available at: https://doi.org/10.1016/j.ijmm.2013.02.009.

Cressman, M.D. *et al.* (2010) ‘Interrelations between the microbiotas in the litter and in the intestines of commercial broiler chickens’, *Applied and Environmental Microbiology*, 76(19), pp. 6572–6582. Available at: https://doi.org/10.1128/AEM.00180-10/ASSET/ABCF04F7-716D-4DFE-ABA9-AD65D0826667/ASSETS/GRAPHIC/ZAM999101364005B.JPEG.

Croxen, M.A. *et al.* (2013) ‘Recent advances in understanding enteric pathogenic Escherichia coli’, *Clinical Microbiology Reviews*, 26(4), pp. 822–880. Available at: https://doi.org/10.1128/CMR.00022-13/ASSET/1535F48F-03D1-4FCC-A2D1-996296B3EB69/ASSETS/GRAPHIC/ZCM9990924420008.JPEG.

De Coster, W. *et al.* (2018) ‘NanoPack: Visualizing and processing long-read sequencing data’, *Bioinformatics*, 34(15), pp. 2666–2669. Available at: https://doi.org/10.1093/bioinformatics/bty149.

De Maio, N. *et al.* (2019) ‘Comparison of long-read sequencing technologies in the hybrid assembly of complex bacterial genomes’, *Microbial Genomics*, 5(9). Available at: https://doi.org/10.1099/mgen.0.000294.

Deborah Chen, H. and Frankel, G. (2005) ‘Enteropathogenic Escherichia coli: unravelling pathogenesis’, *FEMS Microbiology Reviews*, 29(1), pp. 83–98. Available at: https://doi.org/10.1016/J.FEMSRE.2004.07.002.

Del Mar Tavío, M. *et al.* (1999) ‘Mechanisms involved in the development of resistance to fluoroquinolones in Escherichia coli isolates’, *Journal of Antimicrobial Chemotherapy*, 44(6), pp. 735–742. Available at: https://doi.org/10.1093/JAC/44.6.735.

Delcour, A.H. (2009) ‘Outer membrane permeability and antibiotic resistance’, *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1794(5), pp. 808–816. Available at: https://doi.org/10.1016/j.bbapap.2008.11.005.

Dewey-Mattia, D. *et al.* (2018) ‘Surveillance for Foodborne Disease Outbreaks - United States, 2009-2015’, *MMWR Surveillance Summaries*, 67(10), pp. 1–11. Available at: https://doi.org/10.15585/MMWR.SS6710A1.

Dho-Moulin, M. and Morris Fairbrother, J. (1999) ‘Avian pathogenic Escherichia coli (APEC)’, *Veterinary Research, BioMed Central*, 30(3). Available at: https://hal.archives-ouvertes.fr/hal-00902571.

DI Tommaso, P. *et al.* (2017) ‘Nextflow enables reproducible computational workflows’, *Nature Biotechnology*. Nature Publishing Group. Available at: https://doi.org/10.1038/nbt.3820.

Diarra, S. *et al.* (2021) ‘Alternative Bedding Materials for Poultry: Availability, Efficacy, and Major Constraints’, *Frontiers in Veterinary Science*, 8, p. 899. Available at: https://doi.org/10.3389/FVETS.2021.669504/BIBTEX.

Dziva, F. and Stevens, M.P. (2008) ‘Colibacillosis in poultry: unravelling the molecular basis of virulence of avian pathogenic *Escherichia coli* in their natural hosts’, *Avian Pathology*, 37(4), pp. 355–366. Available at: https://doi.org/10.1080/03079450802216652.

Eberle, K.N. *et al.* (2012) ‘A One Year Study of Newly Constructed Broiler Houses for the Prevalence of Campylobacter’, *International Journal of Poultry Science*, 12(1), pp. 29–36. Available at: https://doi.org/10.3923/ijps.2013.29.36.

Epps, S. *et al.* (2013) ‘Foodborne Campylobacter: Infections, Metabolism, Pathogenesis and Reservoirs’, *International Journal of Environmental Research and Public Health*, 10(12), pp. 6292–6304. Available at: https://doi.org/10.3390/ijerph10126292.

Ewels, P. *et al.* (2016) ‘MultiQC: Summarize analysis results for multiple tools and samples in a single report’, *Bioinformatics*, 32(19), pp. 3047–3048. Available at: https://doi.org/10.1093/bioinformatics/btw354.

Ewels, P.A. *et al.* (2019) ‘Nf-core: Community curated bioinformatics pipelines’, *bioRxiv*, p. 610741. Available at: https://doi.org/10.1101/610741.

Ewers, C. *et al.* (2007) ‘Avian pathogenic, uropathogenic, and newborn meningitis-causing Escherichia coli: How closely related are they?’, *International Journal of Medical Microbiology*, 297(3), pp. 163–176. Available at: https://doi.org/10.1016/J.IJMM.2007.01.003.

Ewers, C. *et al.* (2009) ‘Intestine and environment of the chicken as reservoirs for extraintestinal pathogenic Escherichia coli strains with zoonotic potential’, *Applied and Environmental Microbiology*, 75(1), pp. 184–192. Available at: https://doi.org/10.1128/AEM.01324-08/SUPPL_FILE/TABLE_S3_AEM_01234_08.DOC.

Fan, Y. *et al.* (2022) ‘The Use of Disinfectant in Barn Cleaning Alters Microbial Composition and Increases Carriage of Campylobacter jejuni in Broiler Chickens’, *Applied and Environmental Microbiology*, 88(10), p. e0029522. Available at: https://doi.org/10.1128/AEM.00295-22.

FDA (2018) ‘Summary Report on Antimicrobials Sold or Distributed for Use in Food-Producing Animals’, *Center for Veterinary Medicine*, (May), pp. 1–14.

FDA (2022) *2019 NARMS Update*. National Antimicrobial Resistance Monitoring System (NARMS), U.S. Department of Health and Human Services (Food and Drug Administration (FDA), Centers for Disease Control and Prevention (CDC)), the U.S. Department of Agriculture (USDA) (Food Safety and Inspection Service (FSIS), Animal and Plant Health Inspection Service, and Agricultural Research Service). Available at: https://www.fda.gov/animal-veterinary/national-antimicrobial-resistance-monitoring-system/2019-narms-update-integrated-report-summary.

Feldgarden, M. *et al.* (2019) 'Validating the AMRFINder tool and resistance gene database by using antimicrobial resistance genotype-phenotype correlations in a collection of isolates', *Antimicrobial Agents and Chemotherapy*, 63(11). Available at: https://doi.org/10.1128/AAC.00483-19.

Fenske, G.J. and Scaria, J. (2021) 'Analysis of 56,348 Genomes Identifies the Relationship between Antibiotic and Metal Resistance and the Spread of Multidrug-Resistant Non-Typhoidal Salmonella', *Microorganisms 2021, Vol. 9, Page 1468*, 9(7), p. 1468. Available at: https://doi.org/10.3390/MICROORGANISMS9071468.

Ferri, M. *et al.* (2017) 'Antimicrobial resistance: A global emerging threat to public health systems', *Critical Reviews in Food Science and Nutrition*, 57(13), pp. 2857–2876. Available at: https://doi.org/10.1080/10408398.2015.1077192.

Gao, Q. *et al.* (2012) *Roles of iron acquisition systems in virulence of extraintestinal pathogenic Escherichia coli: salmochelin and aerobactin contribute more to virulence than heme in a chicken infection model*. Available at: https://doi.org/10.1186/1471-2180-12-143.

García, M.R. and Cabo, M.L. (2018) ‘Optimization of E. coli inactivation by benzalkonium chloride reveals the importance of quantifying the inoculum effect on chemical disinfection’, *Frontiers in Microbiology*, 9(JUN), p. 1259. Available at: https://doi.org/10.3389/FMICB.2018.01259/FULL.

Georgia Peanut Commission (2021) *Georgia Peanuts Fact Sheet*. Available at: https://gapeanuts.com/wp-content/uploads/2022/02/2021-Annual-Report.pdf.

Girgis, H.S., Hottes, A.K. and Tavazoie, S. (2009) ‘Genetic Architecture of Intrinsic Antibiotic Susceptibility’, *PLOS ONE*, 4(5), p. e5629. Available at: https://doi.org/10.1371/JOURNAL.PONE.0005629.

Golden, C.E., Rothrock, M.J. and Mishra, A. (2021) ‘Mapping foodborne pathogen contamination throughout the conventional and alternative poultry supply chains’, *Poultry Science*, 100(7), p. 101157. Available at: https://doi.org/10.1016/j.psj.2021.101157.

Guabiraba, R. and Schouler, C. (2015) ‘Avian colibacillosis: Still many black holes’, *FEMS Microbiology Letters*, 362(15), p. 118. Available at: https://doi.org/10.1093/femsle/fnv118.

Gupta, S.K. *et al.* (2014) ‘ARG-annot, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes’, *Antimicrobial Agents and Chemotherapy*, 58(1), pp. 212–220. Available at: https://doi.org/10.1128/AAC.01310-13.

Gurevich, A. *et al.* (2013) ‘QUAST: Quality assessment tool for genome assemblies’, *Bioinformatics*, 29(8), pp. 1072–1075. Available at: https://doi.org/10.1093/bioinformatics/btt086.

Hadley, W. (2016) *Ggplot2*. New York, NY: Springer Science+Business Media, LLC.

He, Y. *et al.* (2020) 'PRAP: Pan Resistome analysis pipeline', *BMC Bioinformatics*, 21(1), p. 20. Available at: https://doi.org/10.1186/s12859-019-3335-y.

Hofacre, C.L. *et al.* (2000) 'Presence of Fluoroquinolone-Resistant Coliforms in Poultry Litter', *Avian Diseases*, 44(4), p. 963. Available at: https://doi.org/10.2307/1593073.

Humphrey, T. (2006) 'Are happy chickens safer chickens? Poultry welfare and disease susceptibility 1', *British Poultry Science*, 47(4), pp. 379–391. Available at: https://doi.org/10.1080/00071660600829084.

Hunt, M. *et al.* (2014) 'A comprehensive evaluation of assembly scaffolding tools', *Genome Biology*, 15(3), p. R42. Available at: https://doi.org/10.1186/gb-2014-15-3-r42.

Ingle, D.J. *et al.* (2011) 'Biofilm formation by and thermal niche and virulence characteristics of Escherichia spp.', *Applied and Environmental Microbiology*, 77(8), pp. 2695–2700. Available at: https://doi.org/10.1128/AEM.02401-10.

Jakobsen, L. *et al.* (2011) 'Microarray-based detection of extended virulence and antimicrobial resistance gene profiles in phylogroup B2 Escherichia coli of human, meat and animal origin', *Journal of Medical Microbiology*, 60(10), pp. 1502–1511. Available at: https://doi.org/10.1099/jmm.0.033993-0.

Jakobsen, L. *et al.* (2012) 'Is Escherichia coli urinary tract infection a zoonosis? Proof of direct link with production animals and meat', *European Journal of Clinical Microbiology and*

*Infectious Diseases*, 31(6), pp. 1121–1129. Available at: https://doi.org/10.1007/S10096-011-1417-5/FIGURES/2.

Jay Daniel, K.C. Olson (2005) *Feeding Poultry Litter to Beef Cattle*. University of Missouri-Columbia, Division of Animal Sciences. Available at: https://extension.missouri.edu/media/wysiwyg/Extensiondata/Pub/pdf/agguides/ansci/g02077.pdf.

Johnning, A. *et al.* (2015) 'Resistance Mutations in gyrA and parC are Common in Escherichia Communities of both Fluoroquinolone-Polluted and Uncontaminated Aquatic Environments', *Frontiers in Microbiology*, 6(DEC), p. 1355. Available at: https://doi.org/10.3389/FMICB.2015.01355.

Johnson, T.J. *et al.* (2006) 'DNA sequence of a ColV plasmid and prevalence of selected plasmid-encoded virulence genes among avian Escherichia coli strains', *Journal of Bacteriology*, 188(2), pp. 745–758. Available at: https://doi.org/10.1128/JB.188.2.745-758.2006.

Johnson, T.J. *et al.* (2007) 'The genome sequence of avian pathogenic Escherichia coli strain O1:K1:H7 shares strong similarities with human extraintestinal pathogenic E. coli genomes', *Journal of Bacteriology*, 189(8), pp. 3228–3236. Available at: https://doi.org/10.1128/JB.01726-06/ASSET/85203383-A758-4288-B983-30303F75676B/ASSETS/GRAPHIC/ZJB0080766960003.JPEG.

Johnson, T.J. *et al.* (2008) 'Comparison of extraintestinal pathogenic Escherichia coli strains from human and avian sources reveals a mixed subset representing potential zoonotic pathogens', *Applied and Environmental Microbiology*, 74(22), pp. 7043–7050. Available at:

https://doi.org/10.1128/AEM.01395-08/ASSET/35350ED5-B35D-41B0-A1E2-D880DD80BCFF/ASSETS/GRAPHIC/ZAM0220893520003.JPEG.

Kaesbohrer, A. *et al.* (2012) ‘Emerging Antimicrobial Resistance in Commensal Escherichia coli with Public Health Relevance: Emerging Antimicrobial Resistance in Commensal E. coli’, *Zoonoses and Public Health*, 59, pp. 158–165. Available at: https://doi.org/10.1111/j.1863-2378.2011.01451.x.

Kaper, J.B., Nataro, J.P. and Mobley, H.L.T. (2004) ‘Pathogenic Escherichia coli’, *Nature Reviews Microbiology 2004 2:2*, 2(2), pp. 123–140. Available at: https://doi.org/10.1038/nrmicro818.

Katouli, M. (2010) ‘Population structure of gut Escherichia coli and its role in development of extra-intestinal infections’, *Iranian Journal of Microbiology*, 2(2), p. 59.

Kelley, T.R. *et al.* (1995) ‘Bacterial Pathogens and Indicators in Poultry Litter during Re-Utilization’, *Journal of Applied Poultry Research*, 4(4), pp. 366–373. Available at: https://doi.org/10.1093/japr/4.4.366.

Kelley, T.R. *et al.* (1998) ‘Antibiotic resistance of bacterial litter isolates’, *Poultry Science*, 77(2), pp. 243–247. Available at: https://doi.org/10.1093/ps/77.2.243.

Khairy, R.M. *et al.* (2019) ‘Phylogenic classification and virulence genes profiles of uropathogenic E. coli and diarrhegenic E. coli strains isolated from community acquired infections’, *PLOS ONE*, 14(9), p. e0222441. Available at: https://doi.org/10.1371/JOURNAL.PONE.0222441.

Kirchhelle, C. (2018) ‘Pharming animals: a global history of antibiotics in food production (1935–2017)’, *Palgrave Communications*, 4(1), p. 96. Available at: https://doi.org/10.1057/s41599-018-0152-2.

Kurtzer, G.M., Sochat, V. and Bauer, M.W. (2017) ‘Singularity: Scientific containers for mobility of compute’, *PLoS ONE*. Edited by A. Gursoy, 12(5), p. e0177459. Available at: https://doi.org/10.1371/journal.pone.0177459.

Lakin, S.M. *et al.* (2017) ‘MEGARes: An antimicrobial resistance database for high throughput sequencing’, *Nucleic Acids Research*, 45(D1), pp. D574–D580. Available at: https://doi.org/10.1093/nar/gkw1009.

Li, G. *et al.* (2012) ‘Proteome response of an extraintestinal pathogenic Escherichia coli strain with zoonotic potential to human and chicken sera’, *Journal of Proteomics*, 75(15), pp. 4853–4862. Available at: https://doi.org/10.1016/J.JPROT.2012.05.044.

Li, H. (2016) ‘Minimap and miniasm: Fast mapping and de novo assembly for noisy long sequences’, *Bioinformatics*, 32(14), pp. 2103–2110. Available at: https://doi.org/10.1093/bioinformatics/btw152.

Li, L.-G., Xia, Y. and Zhang, T. (2017) ‘Co-occurrence of antibiotic and metal resistance genes revealed in complete genome collection’, *The ISME Journal*, 11(3), pp. 651–662. Available at: https://doi.org/10.1038/ismej.2016.155.

Li, X. *et al.* (2022) ‘Metagenomic evidence for co-occurrence of antibiotic, biocide and metal resistance genes in pigs’, *Environment International*, 158, p. 106899. Available at: https://doi.org/10.1016/j.envint.2021.106899.

Line, J.E. and Bailey, J.S. (2006) ‘Effect of On-Farm Litter Acidification Treatments on Campylobacter and Salmonella Populations in Commercial Broiler Houses in Northeast Georgia’, *Poultry Science*, 85(9), pp. 1529–1534. Available at: https://doi.org/10.1093/ps/85.9.1529.

Liu, B. *et al.* (2019) ‘VFDB 2019: A comparative pathogenomic platform with an interactive web interface’, *Nucleic Acids Research*, 47(D1), pp. D687–D692. Available at: https://doi.org/10.1093/nar/gky1080.

Lu, S. *et al.* (2016) ‘Insights into the evolution of pathogenicity of Escherichia coli from genomic analysis of intestinal E. coli of Marmota himalayana in Qinghai–Tibet plateau of China’, *Emerging Microbes & Infections*, 5(12), p. e122. Available at: https://doi.org/10.1038/EMI.2016.122.

Lutteropp, S., Kozlov, A.M. and Stamatakis, A. (2019) ‘A Fast and Memory-Efficient Implementation of the Transfer Bootstrap’, *bioRxiv* [Preprint]. Available at: https://doi.org/10.1101/734848.

Lynch, M. *et al.* (2006) ‘Surveillance for Foodborne-Disease Outbreaks — United States, 1998–2002’, *MMWR Surveillance Summaries*, 55(SS10), pp. 1–34.

Mainil, J. (2013) ‘Escherichia coli virulence factors’, *Veterinary Immunology and Immunopathology*, 152(1–2), pp. 2–12. Available at: https://doi.org/10.1016/J.VETIMM.2012.09.032.

Mao, B.H. *et al.* (2012) ‘Identification of Escherichia coli Genes Associated with Urinary Tract Infections’, *Journal of Clinical Microbiology*, 50(2), p. 449. Available at: https://doi.org/10.1128/JCM.00640-11.

Marijke Van Moerbeke (2019) ‘IntClust: integration of multiple data sets with clustering techniques. Version 0.1.0.’ Available at: https://CRAN.R-project.org/package=IntClust.

Mellata, M. *et al.* (2003) ‘Role of Virulence Factors in Resistance of Avian Pathogenic *Escherichia coli* to Serum and in Pathogenicity’, *Infection and Immunity*, 71(1), pp. 536–540. Available at: https://doi.org/10.1128/IAI.71.1.536-540.2003.

Mellata, M. (2013) ‘Human and avian extraintestinal pathogenic escherichia coli: Infections, zoonotic risks, and antibiotic resistance trends’, *Foodborne Pathogens and Disease*. Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA. Available at: https://doi.org/10.1089/fpd.2013.1533.

Miller, D.L. *et al.* (2019) ‘Distance Sampling in *R*’, *Journal of Statistical Software*, 89(1). Available at: https://doi.org/10.18637/jss.v089.i01.

Mitchell, N.M. *et al.* (2015) ‘Zoonotic potential of Escherichia coli isolates from retail chicken meat products and eggs’, *Applied and environmental microbiology*, 81(3), pp. 1177–1187. Available at: https://doi.org/10.1128/AEM.03524-14.

Moulin-Schouleur, M. *et al.* (2006) ‘Common virulence factors and genetic relationships between O18:K1:H7 Escherichia coli isolates of human and avian origin’, *Journal of Clinical Microbiology*, 44(10), pp. 3484–3492. Available at:

https://doi.org/10.1128/JCM.00548-06/ASSET/A4A5E7AD-BABB-4FEA-B194-89BEA5862078/ASSETS/GRAPHIC/ZJM0100667760003.JPEG.

Muniz MV, MSc, E. *et al.* (2014) ‘Presence of Salmonella spp. in reused broiler litter Presencia de Salmonella spp. en cama reutilizada de pollos de engorde/Presença de Salmonella spp. em cama reutilizada de frangos de corte’, *Revista Colombiana de Ciencias Pecuarias*, 27(1), pp. 12–17.

Nurk, S. *et al.* (2013) ‘Assembling genomes and mini-metagenomes from highly chimeric reads’, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 7821 LNBI, pp. 158–170. Available at: https://doi.org/10.1007/978-3-642-37195-0_13.

Nyachuba, D.G. (2010) ‘Foodborne illness: is it on the rise?’, *Nutrition Reviews*, 68(5), pp. 257–269. Available at: https://doi.org/10.1111/J.1753-4887.2010.00286.X.

Oladeinde, A. *et al.* (2019) ‘Horizontal Gene Transfer and Acquired Antibiotic Resistance in Salmonella enterica Serovar Heidelberg following In Vitro Incubation in Broiler Ceca’, *Applied and Environmental Microbiology*, 85(22). Available at: https://doi.org/10.1128/aem.01903-19.

Oladeinde, A. *et al.* (2022) ‘Litter Commensal Bacteria Can Limit the Horizontal Gene Transfer of Antimicrobial Resistance to Salmonella in Chickens’, *Applied and Environmental Microbiology* [Preprint]. Available at: https://doi.org/10.1128/aem.02517-21.

Oladeinde, A. *et al.* (2023) ‘Management and environmental factors influence the prevalence and abundance of food-borne pathogens and commensal bacteria in peanut hull-based broiler litter’, *Poultry Science*, p. 102313. Available at: https://doi.org/10.1016/j.psj.2022.102313.

Pan, Y. *et al.* (2020) ‘Coexistence of Antibiotic Resistance Genes and Virulence Factors Deciphered by Large-Scale Complete Genome Analysis’, *mSystems*, 5(3). Available at: https://doi.org/10.1128/MSYSTEMS.00821-19/SUPPL_FILE/MSYSTEMS.00821-19-S0001.DOC.

Panth, Y. (2019) ‘Colibacillosis in poultry: A review’, *Journal of Agriculture and Natural Resources*, 2(1), pp. 301–311. Available at: https://doi.org/10.3126/JANR.V2I1.26094.

Panunzi, L.G. (2020) ‘sraX: A Novel Comprehensive Resistome Analysis Tool’, *Frontiers in Microbiology*, 11, p. 52. Available at: https://doi.org/10.3389/fmicb.2020.00052.

de Perio, M.A. *et al.* (2013) ‘*Campylobacter* Infection in Poultry-Processing Workers, Virginia, USA, 2008–2011’, *Emerging Infectious Diseases*, 19(2), pp. 286–288. Available at: https://doi.org/10.3201/eid1902.121147.

Pope, M.J. and Cherry, T.E. (2000) ‘An evaluation of the presence of pathogens on broilers raised on poultry litter treatment®-treated litter’, *Poultry Science*, 79(9), pp. 1351–1355. Available at: https://doi.org/10.1093/ps/79.9.1351.

Price, L.B. *et al.* (2007) ‘The Persistence of Fluoroquinolone-Resistant Campylobacter in Poultry Production’, *Environmental Health Perspectives*, 115(7), p. 1035. Available at: https://doi.org/10.1289/EHP.10050.

R Core Team (2021) ‘R: A language and environment for statistical computing.’ R Foundation for Statistical Computing, Vienna, Austria. Available at: https://www.R-project.org/.

Rammelkamp, C.H. and Maxon, T. (1942) ‘Resistance of Staphylococcus aureus to the Action of Penicillin.’, *Experimental Biology and Medicine*, 51(3), pp. 386–389. Available at: https://doi.org/10.3181/00379727-51-13986.

Ramos, S. *et al.* (2020) ‘Escherichia coli as Commensal and Pathogenic Bacteria among Food-Producing Animals: Health Implications of Extended Spectrum β-Lactamase (ESBL) Production’, *Animals : an Open Access Journal from MDPI*, 10(12), pp. 1–15. Available at: https://doi.org/10.3390/ANI10122239.

Randall, C.P. *et al.* (2013) ‘The Target of Daptomycin Is Absent from Escherichia coli and Other Gram-Negative Pathogens’, *Antimicrobial Agents and Chemotherapy*, 57(1), pp. 637–639. Available at: https://doi.org/10.1128/AAC.02005-12.

Rauber Würfel, S.D.F. *et al.* (2019) ‘Population Dynamics of Thermotolerant Campylobacter in Broilers Reared on Reused Litter’, *Foodborne Pathogens and Disease*, 16(11), pp. 738–743. Available at: https://doi.org/10.1089/fpd.2019.2645.

Roberts, B.N. *et al.* (2013) ‘Spatial and temporal analysis of microbial populations in production broiler house litter in the southeastern United States’, *Journal of Applied Poultry Research*, 22(4), pp. 759–770. Available at: https://doi.org/10.3382/japr.2012-00688.

Roll, V.F.B., Dai Prá, M.A. and Roll, A.P. (2011) ‘Research on salmonella in broiler litter reused for up to 14 consecutive flocks’, *Poultry Science*, 90(10), pp. 2257–2262. Available at: https://doi.org/10.3382/ps.2011-01583.

RStudio Team (2016) ‘RStudio: Integrated Development Environment for R’. Boston, MA: RStudio, Inc.

Russo, T.A. and Johnson, J.R. (2003) ‘Medical and economic impact of extraintestinal infections due to Escherichia coli: focus on an increasingly important endemic problem’, *Microbes and Infection*, 5(5), pp. 449–456. Available at: https://doi.org/10.1016/S1286-4579(03)00049-2.

Sáenz, Y. *et al.* (2003) ‘Mutations in gyrA and parC genes in nalidixic acid-resistant Escherichia coli strains from food products, humans and animals’, *Journal of Antimicrobial Chemotherapy*, 51(4), pp. 1001–1005. Available at: https://doi.org/10.1093/JAC/DKG168.

Scallan, E. *et al.* (2011) ‘Foodborne illness acquired in the United States–major pathogens’, *Emerging infectious diseases*, 17(1), pp. 7–15. Available at: https://doi.org/10.3201/EID1701.P11101.

Schefferle, H.E. (1965) ‘The Microbiology of Built Up Poultry Litter’, *Journal of Applied Bacteriology*, 28(3), pp. 403–411. Available at: https://doi.org/10.1111/j.1365-2672.1965.tb02170.x.

Scott, A. *et al.* (2018) ‘Effect of different levels of copper nanoparticles and copper sulphate on performance, metabolism and blood biochemical profiles in broiler chicken’, *Journal of Animal Physiology and Animal Nutrition*, 102(1), pp. e364–e373. Available at: https://doi.org/10.1111/jpn.12754.

Seemann, T. (2014) ‘Prokka: Rapid prokaryotic genome annotation’, *Bioinformatics*, 30(14), pp. 2068–2069. Available at: https://doi.org/10.1093/bioinformatics/btu153.

Seemann, T. (no date) ‘ABRICATE’, *Github* [Preprint]. Available at: https://github.com/tseemann/abricate.

Seiler, C. and Berendonk, T.U. (2012) ‘Heavy metal driven co-selection of antibiotic resistance in soil and water bodies impacted by agriculture and aquaculture’, *Frontiers in Microbiology*, 3. Available at: https://doi.org/10.3389/fmicb.2012.00399.

Seo, K.W. and Lee, Y.J. (2018) ‘Prevalence and Characterization of β-Lactamases Genes and Class 1 Integrons in Multidrug-Resistant Escherichia coli Isolates from Chicken Meat in Korea’, *https://home.liebertpub.com/mdr*, 24(10), pp. 1599–1606. Available at: https://doi.org/10.1089/MDR.2018.0019.

Skyberg, J.A. *et al.* (2006) ‘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’, *Infection and Immunity*, 74(11), pp. 6287–6292. Available at: https://doi.org/10.1128/IAI.00363-06/ASSET/1107A68C-A398-4DD0-BB6E-52A326A03724/ASSETS/GRAPHIC/ZII0110662750002.JPEG.

Slater, S.L. *et al.* (2018) ‘The type III secretion system of pathogenic escherichia coli’, *Current Topics in Microbiology and Immunology*, 416, pp. 51–72. Available at: https://doi.org/10.1007/82_2018_116/TABLES/1.

Smati, M. *et al.* (2013) ‘Real-time PCR for quantitative analysis of human commensal Escherichia coli populations reveals a high frequency of subdominant phylogroups’, *Applied and Environmental Microbiology*, 79(16), pp. 5005–5012. Available at: https://doi.org/10.1128/AEM.01423-13/SUPPL_FILE/ZAM999104614SO1.PDF.

Smith, S. *et al.* (2016) ‘The impact of environmental conditions on *Campylobacter jejuni* survival in broiler faeces and litter’, *Infection Ecology & Epidemiology*, 6(1), p. 31685. Available at: https://doi.org/10.3402/iee.v6.31685.

Sorsa, L.J. *et al.* (2003) ‘Characterization of an iroBCDEN gene cluster on a transmissible plasmid of uropathogenic Escherichia coli: Evidence for horizontal transfer of a chromosomal virulence factor’, *Infection and Immunity*, 71(6), pp. 3285–3293. Available at: https://doi.org/10.1128/IAI.71.6.3285-3293.2003.

Stamatakis, A. (2006) ‘RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models’, *Bioinformatics*, 22(21), pp. 2688–2690. Available at: https://doi.org/10.1093/bioinformatics/btl446.

Starikova, E.V. *et al.* (2020) ‘Phigaro: High-throughput prophage sequence annotation’, *Bioinformatics*, 36(12), pp. 3882–3884. Available at: https://doi.org/10.1093/bioinformatics/btaa250.

Strathdee, S.A., Davies, S.C. and Marcelin, J.R. (2020) ‘Confronting antimicrobial resistance beyond the COVID-19 pandemic and the 2020 US election’, *The Lancet*, 396(10257), pp. 1050–1053. Available at: https://doi.org/10.1016/S0140-6736(20)32063-8.

Su, M., Satola, S.W. and Read, T.D. (2019) ‘Genome-based prediction of bacterial antibiotic resistance’, *Journal of Clinical Microbiology*. American Society for Microbiology. Available at: https://doi.org/10.1128/JCM.01405-18.

Terzich, M. *et al.* (2000) ‘Survey of Pathogens in Poultry Litter in the United States’, *Journal of Applied Poultry Research*, 9(3), pp. 287–291. Available at: https://doi.org/10.1093/JAPR/9.3.287.

Tivendale, K.A. *et al.* (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’, *Infection and Immunity*, 78(8), pp. 3412–3419. Available at: https://doi.org/10.1128/IAI.00347-10/ASSET/D143C0CA-4259-4FFE-A5D1-120CDAA88CFA/ASSETS/GRAPHIC/ZII9990987270004.JPEG.

Torsten Seemann (2015) ‘Snippy, Rapid haploid variant calling and core genome alignment’, *GitHub repository* [Preprint]. Available at: https://github.com/tseemann/snippy.

Tseytin, I. *et al.* (2018) ‘The role of EscD in supporting EscC polymerization in the type III secretion system of enteropathogenic Escherichia coli’, *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1860(2), pp. 384–395. Available at: https://doi.org/10.1016/J.BBAMEM.2017.10.001.

United States Food and Drug Administration, Center for Veterinary Medicine (2022) ‘2021 Summary Report On Antimicrobials Sold or Distributed for Use in Food-Producing Animals’, pp. 1–49.

U.S. Department of Health and Human Services (2019) *2016-2017 NARMS Integrated Summary*. Rockville, MD: Food and Drug Administration (FDA).

USDA, National Agricultural Statistics Service (2021) *Poultry - Production and Value 2020 Summary*. ISSN: 1949-1573. Available at: https://www.nass.usda.gov/Publications/Todays_Reports/reports/plva0421.pdf.

Valeris-Chacin, R. *et al.* (2021) 'Association of Broiler Litter Microbiome Composition and Campylobacter Isolation', *Frontiers in Veterinary Science*, 8, p. 495. Available at: https://doi.org/10.3389/fvets.2021.654927.

Valeris-Chacin, R. *et al.* (2022) 'Longitudinal Changes in Campylobacter and the Litter Microbiome throughout the Broiler Production Cycle', *Applied and Environmental Microbiology* [Preprint]. Edited by D. Ercolini. Available at: https://doi.org/10.1128/AEM.00667-22.

Van Hoek, A.H.A.M. *et al.* (2011) 'Acquired antibiotic resistance genes: An overview', *Frontiers in Microbiology*, 2(SEP), p. 203. Available at: https://doi.org/10.3389/FMICB.2011.00203/BIBTEX.

Vaser, R. *et al.* (2017) 'Fast and accurate de novo genome assembly from long uncorrected reads', *Genome Research*, 27(5), pp. 737–746. Available at: https://doi.org/10.1101/gr.214270.116.

Vaz, C.S.L. *et al.* (2017) 'Interventions to reduce the bacterial load in recycled broiler litter', *Poultry Science*, 96(8), pp. 2587–2594. Available at: https://doi.org/10.3382/ps/pex063.

Wang, M. *et al.* (2022) 'VRprofile2: detection of antibiotic resistance-associated mobilome in bacterial pathogens', *Nucleic Acids Research*, 50. Available at: https://doi.org/10.1093/nar/gkac321.

Wang, S. *et al.* (2016) ‘Escherichia coli type III secretion system 2 ATPase EivC Is involved in the motility and virulence of avian pathogenic Escherichia coli’, *Frontiers in Microbiology*, 7(AUG), p. 1387. Available at: https://doi.org/10.3389/FMICB.2016.01387/BIBTEX.

Wannaprasat, W. *et al.* (2009) ‘Quality analysis of commercial probiotic products for food animals’, *The Southeast Asian Journal of Tropical Medicine and Public Health*, 40(5), pp. 1103–1112.

Wei, S. *et al.* (2013) ‘Abundance of pathogens in the gut and litter of broiler chickens as affected by bacitracin and litter management’, *Veterinary Microbiology*, 166(3–4), pp. 595–601. Available at: https://doi.org/10.1016/j.vetmic.2013.06.006.

Wick, R.R. *et al.* (2015) ‘Bandage: Interactive visualization of de novo genome assemblies’, *Bioinformatics*, 31(20), pp. 3350–3352. Available at: https://doi.org/10.1093/bioinformatics/btv383.

Wick, R.R. *et al.* (2017) ‘Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads’, *PLoS Computational Biology*, 13(6). Available at: https://doi.org/10.1371/journal.pcbi.1005595.

Woodruff, H.B. (2014) ‘Selman A. Waksman, Winner of the 1952 Nobel Prize for Physiology or Medicine’, *Applied and Environmental Microbiology*, 80(1), pp. 2–8. Available at: https://doi.org/10.1128/AEM.01143-13.

World Health Organization, G.W.H. (2019) *WHO | WHO list of Critically Important Antimicrobials (CIA)*. Available at: http://www.who.int/foodborne_disease/resistance/cia/en/#.UiMEZ7zmSDA.mendeley.

Yamak, U.S. *et al.* (2015) ‘Farklı Gelişme Düzeyindeki Etlik Piliçlerde Altlığın Yeniden Kullanılmasının Broiler Performansı, Foot-Pad Dermatitis ve Altlık Kalitesine Etkileri’, *Kafkas Universitesi Veteriner Fakultesi Dergisi* [Preprint]. Available at: https://doi.org/10.9775/kvfd.2015.13957.

Yang, S.C. *et al.* (2017) ‘Current pathogenic Escherichia coli foodborne outbreak cases and therapy development’, *Archives of Microbiology 2017 199:6*, 199(6), pp. 811–825. Available at: https://doi.org/10.1007/S00203-017-1393-Y.

Yoshida, C.E. *et al.* (2016) ‘The salmonella in silico typing resource (SISTR): An open web-accessible tool for rapidly typing and subtyping draft salmonella genome assemblies’, *PLoS ONE*, 11(1). Available at: https://doi.org/10.1371/journal.pone.0147101.

Yu, G. (2020) ‘Using ggtree to Visualize Data on Tree-Like Structures’, *Current Protocols in Bioinformatics*, 69(1). Available at: https://doi.org/10.1002/cpbi.96.

Zancan, F.T. *et al.* (2000) ‘Salmonella spp investigation in transport boxes of day-old birds’, *Brazilian Journal of Microbiology*, 31(3), pp. 229–232. Available at: https://doi.org/10.1590/S1517-83822000000300016.

# CHAPTER 2: READS2RESISTOME: AN ADAPTABLE AND HIGH-THROUGHPUT WHOLE-GENOME SEQUENCING PIPELINE FOR BACTERIAL RESISTOME CHARACTERIZATION[1]

## 2.1 Summary

Accurately characterizing the resistome is crucial for prevention and mitigation of emerging antibiotic resistance threats to animal and human health. We define the resistome to be the collection of antimicrobial resistance (AMR) genes and their precursors in microorganisms. These genes may be found embedded in a bacterial chromosome or plasmid. Reads2Resistome is a tool which allows researchers to assemble and annotate bacterial genomes using long or short read sequencing technologies, or both in a hybrid approach, with focus on the resistome characterization. Using a massively parallel analysis pipeline, Reads2Resistome performs assembly and annotation with the goal of producing an accurate and comprehensive description of a bacterial genome and AMR and virulence genes' content. Reads2Resistome is freely available as an open-source package under the MIT license, and can be downloaded via GitHub (https://github.com/BioRRW/Reads2Resistome). Key features of the Reads2Resistome pipeline include quality control of input sequencing reads, genome assembly, genome annotation and AMR gene, virulence gene, and prophage characterization. Based on the case study we observed

1 This work has been previously published on a preprint server: Woyda, R., Oladeinde, A., Abdo, Z., 2020. Reads2Resistome: An adaptable and high-throughput whole-genome sequencing pipeline for bacterial resistome characterization. bioRxiv 2020.05.18.102715. https://doi.org/10.1101/2020.05.18.102715. This work has been previously published in a peer reviewed journal: Woyda, R., Oladeinde, A., Abdo, Z., 2023. Chicken Production and Human Clinical Escherichia coli Isolates Differ in Their Carriage of Antimicrobial Resistance and Virulence Factors. Appl. Environ. Microbiol. e01167-22. https://doi.org/10.1128/aem.01167-22

hybrid assembly, although the most time-intensive assembly method, produces a highly contiguous genome assembly with robust gene annotation, prophage identification, and resistome characterization compared to short read alone and long read alone assemblies.

## 2.2 Introduction

The Centers for Disease Control (CDC) estimates that, per year, at least two million people will become infected by a drug-resistant bacteria and at least 23,000 will die as a result of such infections (Centers for Disease Control and Prevention, 2019). The primary method for determining antimicrobial resistance (AMR) in clinical laboratories is culture-based antimicrobial susceptibility testing (AST). However, declining cost of next-generation sequencing technologies has enabled increased sequencing depth and more accurate identification of AMR elements within bacterial genomes (Su, Satola and Read, 2019). Additionally, PacBio and Oxford Nanopore MinION sequencing technologies can provide sequencing reads longer than 10kb, improving our ability to generate complete and accurate genome assemblies (De Maio *et al.*, 2019). When identifying AMR genes, a researcher might use a dozen or more individual tools with each relying on their own databases and dependencies. Here we present Reads2Resistome, a streamlined pipeline for quality control, assembly and AMR and virulence characterization for bacterial genomes, which utilizes short sequencing reads, long sequencing reads or both, in a hybrid assembly approach. We illustrate the utility of Reads2Resistome to assemble, annotate, characterize AMR genes, virulence genes and prophage sequences of two bacterial isolates recovered from the ceca of 2-week-old broiler chickens.

## 2.3 Methods and Implementation

Reads2Resistome is scripted using Nextflow (DI Tommaso *et al.*, 2017), a parallel Domain Specific Languages (DSL) workflow framework, and is integrated with Singularity (Kurtzer, Sochat and Bauer, 2017), an open-source container platform with focus towards high-performance computing (HPC) workloads. The Reads2Resistome pipeline includes three main steps: quality control, assembly, and annotation of assembled bacterial contigs. Reads2Resistome takes long read sequences, short read sequences or both as input and performs quality control of short reads using Trimmomatic (Bolger, Lohse and Usadel, 2014) and long read quality visualization using NanoPlot (De Coster *et al.*, 2018). Both quality-controlled short reads and long reads are then assembled using Unicycler (Wick *et al.*, 2017). Unicycler generates an assembly graph using short reads, then uses long reads to simplify the graph to generate accurate assemblies. In the event the input consists of short reads only, Unicycler employs SPAdes (Nurk *et al.*, 2013; Antipov *et al.*, 2016) for assembly and subsequently polishes the resulting graph by bridging contigs. Long read-only assembly is performed using miniasm (Li, 2016) and Racon (Vaser *et al.*, 2017) employed through Unicycler. Annotation is performed with Prokka (Seemann, 2014) using one of the provided custom databases (described in the Resistome characterization section below), which are pre-built from collections of specific bacterial species and subtypes, or using the Prokka default database. Resistome characterization is performed using ABRICATE (Seemann, no date). Nextflow implementation using Singularity provides version control over the various open-source tools ensuring reproducible results. Reads2Resistome output contains the following for each input isolate: visualization of both raw and quality-controlled reads; assembled contigs with a corresponding assembly graph along with an assembly quality assessment; gene and resistome annotation files; genome alignment files in

BAM format; and optional serovar predictions. All documentation and pipeline usage are publicly available at https://github.com/BioRRW/Reads2Resistome.

### 2.3.1 Streamline high-throughput analysis

Reads2Resistome is designed for high-throughput bacterial sequence input and performs quality control, genome assembly and subsequent genome and AMR and virulence gene annotation in a parallel, high-throughput manner (**Figure 2.1**). Reads2Resistome is able to accommodate input of different species within the same run and can perform species-specific genome assembly quality control and gene annotation. A comma-separated values (CSV) file, generated by the user, enables input of multiple different isolates regardless of the isolate identity. The user can also specify a pipeline-provided database for genome annotation or can choose to use the default database utilized by Prokka. For genome assembly quality assessment, done by QUAST, the user can optionally add a user-provided reference genome for additional reference-specific metrics. Pipeline outputs for quality control and genome quality assessment are aggregated by MultiQC into a HTML report (Ewels *et al.*, 2016). In addition to quality control, assembly and annotation, genome alignments are generated for further comparison. For *Salmonella spp*. optional serovar prediction is performed using SISTR (Yoshida *et al.*, 2016).

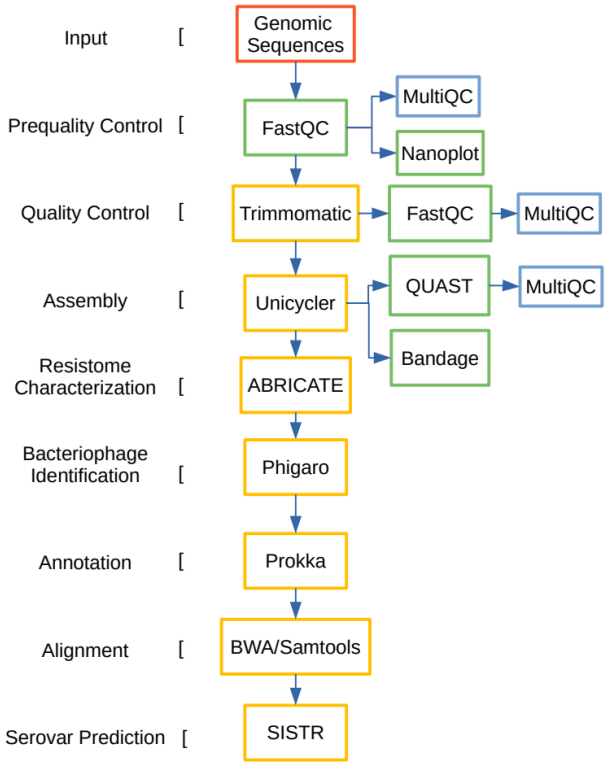



**Figure 2.1.** Key processes in Reads2Resistome pipeline.

### 2.3.2 Adaptable to cutting-edge sequencing technologies

Inclusion of long read sequences into bacteria assembly aids in resolving repeat regions of genomes and contributes to genome completeness (Hunt *et al.*, 2014). Reads2Resistome is designed to be adaptable and flexible in that it can accommodate assembly in three different approaches: long read-only assembly, short read-only assembly, and hybrid assembly. In each assembly approach, Unicycler is used to generate genome assemblies and assembly graphs which are visualized with Bandage (**Table S2.1**) (Wick *et al.*, 2015).

### 2.3.3 Resistome characterization

Reads2Resistome characterizes resistome content using ABRICATE and Phigaro. ABRICATE uses the assembled contigs to screen for AMR and virulence genes from various databases; ARG-ANNOT antibiotic resistance gene database (Gupta *et al.*, 2014) the Comprehensive Antibiotic Resistance Database (CARD) (Alcock *et al.*, 2020), MEGARes Antimicrobial Database for High-Throughput Sequencing (Lakin *et al.*, 2017) NCBI AMRFinderPlus (Feldgarden *et al.*, 2019) PlasmidFinder (Carattoli *et al.*, 2014), ResFinder (Bortolaia *et al.*, 2020) and VirulenceFinder database (Liu *et al.*, 2019). ABRICATE compiles results into a single report containing hits from each database and Reads2Resistome provides an output file for each isolate. Phigaro uses the assembled contigs to detect putative taxonomic annotations and the output is collected and reported by Reads2Resistome for each isolate (Starikova *et al.*, 2020).

## 2.4 Case study

Using Reads2Resistome we assembled and characterized the AMR genes, virulence genes and prophage sequences of genomes associated with two bacterial isolates recovered from the ceca of 2-week-old broiler chickens; SH-IC: *Salmonella enterica* serovar Heidelberg (*S*. Heidelberg) and EC-IC: *Escherichia coli* (**Table 2.1**). Illumina, PacBio and Oxford Nanopore MinION sequences were used to evaluate and compare the three assembly methods available through the pipeline; short read-only, long read-only and hybrid assembly.

**Table 2.1.** Summary of isolate information.

| Bacterial strains | Strain ID | SRA accession no. | Sequencing platform | Coverage* |
|---|---|---|---|---|
| *Escherichia coli* | EC-IC | SRR11808523 | Illumina | 56.95 |
| | | SRR11808522 | MinION | 35.35 |
| | | SRR11808521 | PacBio | 28.08 |
| *Salmonella* Heidelberg | SH-IC | SRR11808520 | Illumina | 20.25 |
| | | SRR11808519 | MinION | 16.69 |
| | | SRR11808518 | PacBio | 33.14 |
| * Coverage estimated from total quality-controlled bases divided by the genome size (*E. coli*: 4800000bp , *S.* Heidelberg: 4600000bp) | | | | |

Genome assembly and annotation metrics were compiled from QUAST and Prokka outputs. Hybrid assembly of both EC-IC and SH-IC using MinION long reads and Illumina short reads gave the fewest contigs, longest total length and highest number of annotated genes as compared to long read assembly using MinION. Hybrid assembly of both isolates using PacBio reads resulted in fewer contigs but comparable total length to that of the MinION hybrid assembly. Genome contiguity was best obtained by hybrid assembly and can be visualized with Bandage-generated graphs (**Table S2.1**). While hybrid and long read-only assemblies are comparable with respect to number of contigs and genome length, the long read-only assembly greatly lacked in annotated genomic features and resistome elements.

Annotated genes and features across all assembly methods for both isolates were considerably reduced under the long read-only assembly, while both short read and hybrid methods resulted in comparable numbers of annotated genes. We suspect this is due to relative lower quality of long reads as compared to Illumina short reads. This is mirrored in AMR and virulence gene characterization and prophage identification. While both short read and hybrid assembly methods for both isolates resulted in comparable identified resistome elements and prophage sequences, long read-only assembly identified elements were significantly reduced (**Table S2.3,S2.4**).

**Table 2.2.** Summary of evaluation for assembled isolates under various assembly conditions.

| | | Short Read | Hybrid | | Long Read | |
|---|---|---|---|---|---|---|
| | | Illumina | MinION | PacBio | MinION | PacBio |
| Isolate | Assembly Metrics and Annotated Features | | | | | |
| EC-IC | No. contigs | 220 | 11 | 10 | 16 | 6 |
| | Largest Contig (bp) | 143515 | 3902334 | 3890483 | 4876904 | 3256237 |
| | Total Length (bp) | 5082305 | 5292549 | 5286973 | 5266949 | 5421221 |
| | N50 (bp) | 42032 | 3902334 | 3890483 | 4876904 | 3256237 |
| | L50 | 36 | 1 | 1 | 1 | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | | | |
| | GC (%) | 50.54 | 50.38 | 50.36 | 50.05 | 48.86 |
| | tRNAs | 80 | 90 | 90 | 32 | 12 |
| | CRISPRs | 1 | 1 | 1 | 0 | 0 |
| | Predicted CDS | 4765 | 5041 | 5004 | 4826 | 4837 |
| | Annotated Genes | 2277 | 2297 | 2295 | 430 | 185 |
| SH-IC | No. contigs | 57 | 19 | 8 | 90 | 5 |
| | Largest Contig (bp) | 460444 | 2046586 | 4750196 | 64272 | 2033197 |
| | Total Length (bp) | 4844513 | 4869998 | 4899506 | 1448734 | 5035134 |
| | N50 (bp) | 213070 | 1175028 | 4750196 | 16758 | 1502365 |
| | L50 | 9 | 2 | 1 | 28 | 2 |
| | GC (%) | 52.1 | 52.08 | 52.1 | 50.44 | 50.52 |
| | tRNAs | 77 | 76 | 82 | 2 | 30 |
| | CRISPRs | 3 | 3 | 3 | 25 | 1 |
| | Predicted CDS | 4554 | 4573 | 4581 | 1387 | 5976 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | Annotated Genes | 2037 | 2044 | 2043 | 87 | 213 |

The pipeline was run using the following commands for short read-only, hybrid, and long read-only assemblies, respectively:

| |
|---|
| $ nextflow R2R-0.0.2.nf --assembly nonhybrid--input containers/data/input_nonhybrid.csv --output temp/output -w temp/work --threads 64 -with-report --name R2R_Nonhybrid_Assembly |
| $ nextflow R2R-0.0.2.nf --input containers/data/input_hybrid.csv --output temp/output -w temp/work --threads 64 -with-report --name R2R_Hybrid_Assembly |
| $ nextflow R2R-0.0.2.nf --assembly longread --input containers/data/input_longread.csv --output temp/output -w temp/work --threads 64 -with-report --name R2R_Long-Read_Assembly |

Each command was executed independently on a Linux server with 128 compute cores and 504GB of memory. Resources allocated and run-time in **Table S2.2** were obtained from the "report.html" which is generated using the '-with-report' option.

## 2.5 Discussion

Reads2Resistome provides a streamlined high-throughput analysis pipeline for the assembly, genome annotation and resistome characterization for bacterial sequencing reads. The pipeline can perform three methods of assembly: short read-only, long read-only, or a hybrid method utilizing both short and long reads. The user can generate an input CSV file containing multiple

isolate samples from different species, all of which can be fed into the Reads2Resistome pipeline under a user-specified assembly method. Pipeline output is generated for quality control, assembly and annotation for each isolate. The pipeline is executable on both Mac and Linux operating systems and is well-suited for institutions and organizations which maintain, or have access to high performance computing (HPC) for the analysis of "big data."

Results from our case study indicated that a highly contiguous genome assembly with robust gene annotation, prophage identification, and resistome characterization is best obtained under a hybrid assembly approach. While hybrid assembly is the most time-intensive assembly method, it produces the most complete annotated genomes in our case study. Long read-only assembly is able to produce a respectable genome length with high contiguity but falls short when annotating genomic features.

The ability to input multiple samples via the input CSV allows users to analyze hundreds of samples regardless of their identity. This prevents tedious bash scripting and collecting of various output files and ensures consistent and correct naming of all output files. Pipelines such as bacass from nf-core (Ewels *et al.*, 2019) provide genome assembly under short, long and hybrid approaches but do not offer resistome annotation. PRAP (He *et al.*, 2020) and sraX (Panunzi, 2020) offer robust resistome analysis but do not provide genome assembly within the pipeline. Therefore, Reads2Resistome is unique in providing a robust pipeline for genome assembly, genome annotation and resistome characterization.

## 2.6 Limitations

Reads2Resistome is designed to be deployed on Linux servers with at least 16GB of RAM and at least 16 compute cores. Running Reads2Resistome on a personal machine will allow full

completion of the pipeline but time requirements will be daunting using less than 16 compute cores. Prokka gene annotation custom databases consist of *Escherichia coli*, *Campylobacter*, *Salmonella*, *Enterococcus* and *Staphylococcus*; all other input samples will use the default Prokka database. Currently serovar prediction is only provided for *Salmonella spp*. using SISTR. Additionally, obtaining high-quality genome assemblies along with accurate resistome characterization is dependent on the quality and depth of sequencing obtained for the input isolates.

## 2.8 REFERENCES

Alcock, B.P., Raphenya, A.R., Lau, T.T.Y., Tsang, K.K., Bouchard, M., Edalatmand, A., Huynh, W., Nguyen, A.L.V., Cheng, A.A., Liu, S., Min, S.Y., Miroshnichenko, A., Tran, H.K., Werfalli, R.E., Nasir, J.A., Oloni, M., Speicher, D.J., Florescu, A., Singh, B., Faltyn, M., Hernandez-Koutoucheva, A., Sharma, A.N., Bordeleau, E., Pawlowski, A.C., Zubyk, H.L., Dooley, D., Griffiths, E., Maguire, F., Winsor, G.L., Beiko, R.G., Brinkman, F.S.L., Hsiao, W.W.L., Domselaar, G.V., McArthur, A.G., 2020. CARD 2020: Antibiotic resistome surveillance with the comprehensive antibiotic resistance database. Nucleic Acids Res. 48, D517–D525. https://doi.org/10.1093/nar/gkz935

Antipov, D., Korobeynikov, A., McLean, J.S., Pevzner, P.A., 2016. HybridSPAdes: An algorithm for hybrid assembly of short and long reads. Bioinformatics 32, 1009–1015. https://doi.org/10.1093/bioinformatics/btv688

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bortolaia, V., Kaas, R.S., Ruppe, E., Roberts, M.C., Schwarz, S., Cattoir, V., Philippon, A., Allesoe, R.L., Rebelo, A.R., Florensa, A.F., Fagelhauer, L., Chakraborty, T., Neumann, B., Werner, G., Bender, J.K., Stingl, K., Nguyen, M., Coppens, J., Xavier, B.B., Malhotra-Kumar, S., Westh, H., Pinholt, M., Anjum, M.F., Duggett, N.A., Kempf, I., Nykäsenoja, S., Olkkola, S., Wieczorek, K., Amaro, A., Clemente, L., Mossong, J., Losch, S., Ragimbeau, C., Lund, O., Aarestrup, F.M., 2020. ResFinder 4.0 for predictions of phenotypes from genotypes. J. Antimicrob. Chemother. 75, 3491–3500. https://doi.org/10.1093/JAC/DKAA345

Carattoli, A., Zankari, E., Garciá-Fernández, A., Larsen, M.V., Lund, O., Villa, L., Aarestrup, F.M., Hasman, H., 2014. In Silico detection and typing of plasmids using plasmidfinder and plasmid multilocus sequence typing. Antimicrob. Agents Chemother. 58, 3895–3903. https://doi.org/10.1128/AAC.02412-14

Centers for Disease Control and Prevention, 2019. Antibiotic resistance threats in the United States. Cent. Dis. Control Prev. 1–113. https://doi.org/CS239559-B

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, 2666–2669. https://doi.org/10.1093/bioinformatics/bty149

De Maio, N., Shaw, L.P., Hubbard, A., George, S., Sanderson, N.D., Swann, J., Wick, R., Oun, M.A., Stubberfield, E., Hoosdally, S.J., Crook, D.W., Peto, T.E.A., Sheppard, A.E., Bailey, M.J., Read, D.S., Anjum, M.F., Sarah Walker, A., Stoesser, N., 2019. Comparison of long-read sequencing technologies in the hybrid assembly of complex bacterial genomes. Microb. Genomics 5. https://doi.org/10.1099/mgen.0.000294

DI Tommaso, P., Chatzou, M., Floden, E.W., Barja, P.P., Palumbo, E., Notredame, C., 2017. Nextflow enables reproducible computational workflows. Nat. Biotechnol. https://doi.org/10.1038/nbt.3820

Ewels, P., Magnusson, M., Lundin, S., Käller, M., 2016. MultiQC: Summarize analysis results for multiple tools and samples in a single report. Bioinformatics 32, 3047–3048. https://doi.org/10.1093/bioinformatics/btw354

Ewels, P.A., Peltzer, A., Fillinger, S., Alneberg, J., Patel, H., Wilm, A., Garcia, M.U., Di Tommaso, P., Nahnsen, S., 2019. Nf-core: Community curated bioinformatics pipelines. bioRxiv 610741. https://doi.org/10.1101/610741

Feldgarden, M., Brover, V., Haft, D.H., Prasad, A.B., Slotta, D.J., Tolstoy, I., Tyson, G.H., Zhao, S., Hsu, C.H., McDermott, P.F., Tadesse, D.A., Morales, C., Simmons, M., Tillman, G., Wasilenko, J., Folster, J.P., Klimke, W., 2019. Validating the AMRFINder tool and resistance gene database by using antimicrobial resistance genotype-phenotype correlations in a collection of isolates. Antimicrob. Agents Chemother. 63. https://doi.org/10.1128/AAC.00483-19

Gupta, S.K., Padmanabhan, B.R., Diene, S.M., Lopez-Rojas, R., Kempf, M., Landraud, L., Rolain, J.M., 2014. ARG-annot, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. Antimicrob. Agents Chemother. 58, 212–220. https://doi.org/10.1128/AAC.01310-13

He, Y., Zhou, X., Chen, Z., Deng, X., Gehring, A., Ou, H., Zhang, L., Shi, X., 2020. PRAP: Pan Resistome analysis pipeline. BMC Bioinformatics 21, 20. https://doi.org/10.1186/s12859-019-3335-y

Hunt, M., Newbold, C., Berriman, M., Otto, T.D., 2014. A comprehensive evaluation of assembly scaffolding tools. Genome Biol. 15, R42. https://doi.org/10.1186/gb-2014-15-3-r42

Kurtzer, G.M., Sochat, V., Bauer, M.W., 2017. Singularity: Scientific containers for mobility of compute. PLoS ONE 12, e0177459. https://doi.org/10.1371/journal.pone.0177459

Lakin, S.M., Dean, C., Noyes, N.R., Dettenwanger, A., Ross, A.S., Doster, E., Rovira, P., Abdo, Z., Jones, K.L., Ruiz, J., Belk, K.E., Morley, P.S., Boucher, C., 2017. MEGARes: An

antimicrobial resistance database for high throughput sequencing. Nucleic Acids Res. 45, D574–D580. https://doi.org/10.1093/nar/gkw1009

Li, H., 2016. Minimap and miniasm: Fast mapping and de novo assembly for noisy long sequences. Bioinformatics 32, 2103–2110. https://doi.org/10.1093/bioinformatics/btw152

Liu, B., Zheng, D., Jin, Q., Chen, L., Yang, J., 2019. VFDB 2019: A comparative pathogenomic platform with an interactive web interface. Nucleic Acids Res. 47, D687–D692. https://doi.org/10.1093/nar/gky1080

Nurk, S., Bankevich, A., Antipov, D., Gurevich, A., Korobeynikov, A., Lapidus, A., Prjibelsky, A., Pyshkin, A., Sirotkin, A., Sirotkin, Y., Stepanauskas, R., McLean, J., Lasken, R., Clingenpeel, S.R., Woyke, T., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2013. Assembling genomes and mini-metagenomes from highly chimeric reads. Lect. Notes Comput. Sci. Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinforma. 7821 LNBI, 158–170. https://doi.org/10.1007/978-3-642-37195-0_13

Panunzi, L.G., 2020. sraX: A Novel Comprehensive Resistome Analysis Tool. Front. Microbiol. 11, 52. https://doi.org/10.3389/fmicb.2020.00052

Seemann, T., 2014. Prokka: Rapid prokaryotic genome annotation. Bioinformatics 30, 2068–2069. https://doi.org/10.1093/bioinformatics/btu153

Seemann, T., n.d. ABRICATE. Github.

Starikova, E.V., Tikhonova, P.O., Prianichnikov, N.A., Rands, C.M., Zdobnov, E.M., Ilina, E.N., Govorun, V.M., 2020. Phigaro: High-throughput prophage sequence annotation. Bioinformatics 36, 3882–3884. https://doi.org/10.1093/bioinformatics/btaa250

Su, M., Satola, S.W., Read, T.D., 2019. Genome-based prediction of bacterial antibiotic resistance. J. Clin. Microbiol. https://doi.org/10.1128/JCM.01405-18

Vaser, R., Sović, I., Nagarajan, N., Šikić, M., 2017. Fast and accurate de novo genome assembly from long uncorrected reads. Genome Res. 27, 737–746. https://doi.org/10.1101/gr.214270.116

Wick, R.R., Judd, L.M., Gorrie, C.L., Holt, K.E., 2017. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput. Biol. 13. https://doi.org/10.1371/journal.pcbi.1005595

Wick, R.R., Schultz, M.B., Zobel, J., Holt, K.E., 2015. Bandage: Interactive visualization of de novo genome assemblies. Bioinformatics 31, 3350–3352. https://doi.org/10.1093/bioinformatics/btv383

Yoshida, C.E., Kruczkiewicz, P., Laing, C.R., Lingohr, E.J., Gannon, V.P.J., Nash, J.H.E., Taboada, E.N., 2016. The salmonella in silico typing resource (SISTR): An open web-accessible tool for rapidly typing and subtyping draft salmonella genome assemblies. PLoS ONE 11. https://doi.org/10.1371/journal.pone.0147101

# CHAPTER 3: CHICKEN PRODUCTION AND HUMAN CLINICAL ESCHERICHIA COLI ISOLATES DIFFER IN THEIR CARRIAGE OF ANTIMICROBIAL RESISTANCE AND VIRULENCE FACTORS[2]

## 3.1 Summary

Contamination of food animal products by *Escherichia coli* (*E. coli*) is a leading cause of food-borne disease outbreaks, hospitalizations, and deaths in humans. Chicken is the most consumed meat both in the United States and across the globe according to the United States Department of Agriculture. Although *E. coli* is a ubiquitous commensal bacterium of the gut of humans and animals, its ability to acquire antimicrobial resistance (AMR) genes and virulence factors (VF) can lead to the emergence of pathogenic strains that are resistant to critically important antibiotics. Thus, it is important to identify the genetic factors that contribute to the virulence and AMR of *E. coli.* In this study, we performed in-depth genomic evaluation of AMR genes and VF of *E. coli* genomes available through the National Antimicrobial Resistance Monitoring System GenomeTrackr database. Our objective was to determine the genetic relatedness of chicken production isolates and human clinical isolates. To achieve this aim, we first developed a massively parallel analytical pipeline (Reads2Resistome) to accurately characterize the resistome of each *E. coli* genome including the AMR and VF harbored. We used Random Forests and hierarchical clustering to show that AMR genes and VF are sufficient to classify isolates into

2 This work has been previously published in a peer reviewed journal: Woyda, R., Oladeinde, A., Abdo, Z., 2023. Chicken Production and Human Clinical Escherichia coli Isolates Differ in Their Carriage of Antimicrobial Resistance and Virulence Factors. Appl. Environ. Microbiol. e01167-22. https://doi.org/10.1128/aem.01167-22

different pathogenic phylogroups and host origin. We found that the presence of key type III secretion system and AMR genes differentiated human clinical isolates from chicken production isolates. These results further improve our understanding of the interconnected role AMR and VF play in shaping the evolution of pathogenic *E. coli* strains.

## 3.2 Importance

Pathogenic *Escherichia coli* causes disease in both humans and food-producing animals. *E. coli* pathogenesis is dependent on a repertoire of virulence factors and antimicrobial resistance genes. Food-borne outbreaks are highly associated with the consumption of under cooked and contaminated food products. This association highlights the need to understand the genetic factors that make *E. coli* virulent and pathogenic in humans and poultry. This research shows that *E. coli* isolates originating from human clinical settings and chicken production harbor different antimicrobial resistance genes and virulence factors that can be used to classify them into phylogroups and host origins. In addition, to aid in the repeatability and reproducibility of the results presented in this study, we have made a public repository of the Reads2Resistome pipeline and have provided the accession numbers associated with the *E. coli* genomes analyzed.

## 3.3 Introduction

*Escherichia coli* are ubiquitous commensal bacteria in the gut of both humans and food-producing animals and rarely cause disease but may acquire antimicrobial resistance (AMR) genes and virulence factors (VF) resulting in increased pathogenicity (Boyle and Zembower, 2015). Pathogenic *E. coli* has consistently ranked in the top 5 causative agents of disease outbreaks, outbreak-associated illnesses, and hospitalizations in the United States (Lynch *et al.*, 2006; Dewey-Mattia *et al.*, 2018; CDC, 2019) and is responsible for billions of dollars of annual

health care associated costs in the United States (Russo and Johnson, 2003; Nyachuba, 2010). Pathogenic E. coli infections in humans may result in acute to severe diarrhea or dysentery, urinary tract infections, and meningitis (Kaper, Nataro and Mobley, 2004; Croxen *et al.*, 2013; Ramos *et al.*, 2020). E. coli pathogenesis is dependent on the repertoire of VF which enable the bacterium to evade host defenses, adhere to host surfaces and successfully invade, and replicate in, host tissues. For example, VF such as toxins, iron-acquisition systems and fimbriae play integral roles in the pathogenicity of extra-intestinal E. coli strains that cause urinary tract infections in humans (Mainil, 2013). These uropathogenic E. coli strains are able to colonize human mucosal surfaces due to surface adhesion VF including the P, F and Type 1 fimbriae encoded by *pap, sfa* and *fim* genes (Khairy *et al.*, 2019).

Outbreaks in human populations caused by pathogenic *E. coli* are attributed to the consumption of under cooked and contaminated foods including meats and vegetables (Yang *et al.*, 2017). *E. coli* infections in food-producing animals usually results in diarrhea but can include acute fatal septicemia, airsacculitis, pericarditis and perihepatitis (Panth, 2019). In poultry, avian pathogenic *E. coli* infections, termed avian colibacillosis, can result in economic loss due to cost of treatment as well as from culling of flocks (Ewers *et al.*, 2009). Colibacillosis is a leading cause of morbidity and mortality in poultry, noted by a decrease in feed conversion ratio, egg production and decreased hatching rates (Dho-Moulin and Morris Fairbrother, 1999; Mellata, 2013; Guabiraba and Schouler, 2015). The severity of the disease depends on the VF repertoire of the strain including genes encoding iron acquisition and transport systems (Johnson *et al.*, 2006).

Extraintestinal pathogenic *E. coli* strains that cause urinary tract infections, neonatal meningitis, and sepsis have been shown to have similar VF as *E. coli* isolates from meat and

animal sources (Mitchell *et al.*, 2015). Likewise, some avian pathogenic *E. coli* and extraintestinal pathogenic *E. coli* isolates have been reported to share similar VF and belong to the same multilocus sequence type (MLST) and phylogroup (Moulin-Schouleur *et al.*, 2006; Ewers *et al.*, 2007; Johnson *et al.*, 2007, 2008). Specifically, a microarray-based study of *E. coli* isolates from both human and animal sources in Denmark identified 66-87 genes, including both virulence factors and antimicrobial resistance genes, which were present in both human urinary tract infection isolates, and isolates obtained from poultry and pig products (Jakobsen *et al.*, 2011). Genes such as the iron acquisition genes *iutA* and *iroN*, as well as fimbrial, *papA* encoding the Pap fimbrial major subunit, and adehsion genes, *papGII* encoding Pap adhesion, were detected in both urinary tract infection and poultry related isolates (Jakobsen *et al.*, 2011). *In vivo* and *in vitro* experimental studies have shown that extraintestinal pathogenic *E. coli* recovered from humans can cause disease in avian hosts and similarly, avian pathogenic *E. coli* isolates recovered from avian host can cause disease in mammalian models (Skyberg *et al.*, 2006; Tivendale *et al.*, 2010; Jakobsen *et al.*, 2012; Li *et al.*, 2012; Comery *et al.*, 2013).

Infections with pathogenic *E. coli* are commonly treated with antibiotics, however increasing levels of AMR impose difficulties in selecting effective treatment options (Caniça *et al.*, 2019). The spread and increased prevalence of AMR has been linked to both the overuse and misuse of antibiotics in the human clinical setting as well as food animal production (Ferri *et al.*, 2017). Consumer opinion has influenced a reduction of antimicrobial use in food animal production to mitigate the spread of AMR. However, recent studies have indicated that AMR can still persist in food animal production even after the removal or stoppage of antibiotics (Fenske and Scaria, 2021; Oladeinde *et al.*, 2022). The co-mobilization and co-acquisition of AMR and VF through horizontal gene transfer can result in highly pathogenic *E. coli* strains. Recent studies

have demonstrated a correlation and close genetic linkage of AMR genes and VF in pathogenic bacteria strains (Biswas, Sharma and Joshi, 2019; Pan *et al.*, 2020). Yu Pan et al. (Pan *et al.*, 2020) performed a comprehensive analysis of over 9,000 bacterial genomes from multiple species and hosts and observed the coexistence of AMR genes and VF from human-associated pathogens. Therefore, the aim of this study is to evaluate and compare the current distribution of AMR genes and VF present in *E. coli* isolates from chicken production and human clinical settings in the United States.

To do so, we determined the AMR genes and VF present in nearly 800 *E. coli* genomes collected from chicken production and the human clinical settings. Data for all isolates were obtained from the National Antimicrobial Resistance Monitoring System (NARMS). We used the World Health Organizations (WHO) most recent classification of antimicrobials important to human health to determine the extent to which resistance to these antimicrobials is prevalent in the chicken production and human clinical settings. We hypothesized that resistance to antimicrobials of critical and high importance to human health would be present in both human and chicken derived isolates. Additionally, we hypothesized that *E. coli* isolates from human clinical settings and chicken production will differ in their carriage of AMR genes and VF. We found that *E. coli* isolates obtained from chicken production and human clinical settings harbored AMR genes predicted to confer resistance to many antimicrobials classified as highly and critically important to human health by the WHO. Furthermore, we were able to classify all *E. coli* isolates into their respective phylogroups and host origins using random forest classification and hierarchical clustering of the AMR genes and VF found in their genomes.

## 3.4 Results

### 3.4.1 *E. coli* isolate selection and genome assembly

To evaluate and compare the current distribution of AMR genes and VF present in *E. coli* isolates from chicken production and the human clinical settings in the United States, we took advantage of the U.S. National Antimicrobial Resistance Monitoring System (NARMS) Genome Trackr database (available through NCBI https://www.ncbi.nlm.nih.gov/pathogens). We chose NARMS because it monitors enteric bacteria and pathogens across 20 states to determine their resistance to antimicrobials used in veterinary and human medicine. NARMS is a collaboration between agencies within the U.S. Department of Health and Human Services (HHS) (FDA and CDC) and the U.S. Department of Agriculture (USDA) (Food Safety and Inspection Service, Animal and Plant Health Inspection Service, and the Agricultural Research Service). Each USDA agency taking part in NARMS tests bacterial samples at various stages of food animal production. The NARMS database holds bacterial genome sequences that originate from 15 distinct human, animal and food sources and are categorized into four NARMS samples sources: human clinical isolates; food animal production isolates from cecal samples at slaughter; samples routinely collected at inspected establishments as part of FSIS verification testing; and raw meats (chicken, ground turkey, ground beef, and pork chops) collected at retail outlets in 20 states (23 sites). We selected 837 chicken production isolates and 874 human clinical isolates, with collection dates from January 1$^{st}$ 2018 until March 30$^{th}$ 2020, using “Species”, “Host” and “Location” search filters (see materials and methods for a detailed description). *E. coli* isolate data were either available as Illumina FASTQ reads or as FASTA assemblies.

One hundred and thirty FASTA assemblies were obtained for human clinical isolates, whereas all chicken production isolates and remaining human clinical isolates had available FASTQ reads. The FASTA assemblies were obtained through the isolate's associated BioProject, and quality control was done using QUAST (Gurevich *et al.*, 2013). The FASTQ reads for each isolate were first normalized to 30X coverage using Trinity's insilico_read_normalization.pl script (https://github.com/trinityrnaseq/trinityrnaseq/wiki). After quality control, the average genome coverage was 35.8X and 31.7X for human clinical and chicken production isolates, respectively (**Figure 3.1A**). Three hundred and eighty-five chicken production isolates and 535 human clinical isolates were excluded from further analysis because they did not meet our preset quality control metrics based on genome coverage greater than or equal to 20X and assembled genome quality (N50 >= 10,000bp, **Figure 3.1B**). The remaining *E. coli* isolates from chicken production and human clinical settings were confirmed to be *E. coli* through MLST (**Table S3.1**, https://github.com/tseemann/mlst). This resulted in our study population consisting of 452 (57%) chicken production isolates and 339 (43%) human clinical isolates (**Figure 3.2A and B**). Isolates originated from a diverse set of sources (**Table S3.1**) and locations within the United States **(Figure 3.2C**).

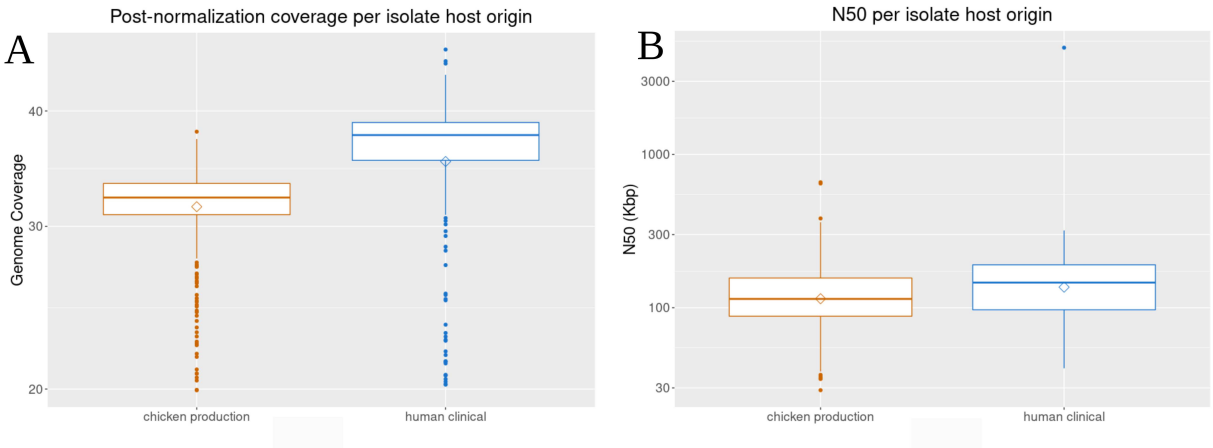



**Figure 3.1.** Average coverage, post-normalization, and average N50, post-assembly, of human clinical and poultry production isolates. (A) Isolate FASTQ reads were downloaded from the

NCBI and normalized to 30× using Trinity's insilico_read_normalization.pl script. (B) Isolates with available FASTQ reads were assembled and annotated using Reads2Resistome, including N50 calculation by QUAST. FASTA assemblies were annotated using the additional script provided by Reads2Resistome for post-assembly annotation. A diamond shape indicates a mean value, and the horizontal lines within a box indicate the mean. The upper and lower sides of the box correspond to the 25th and 75th percentiles. Plots were generated in R v4.0.4 using ggplot2 v3.3.3.

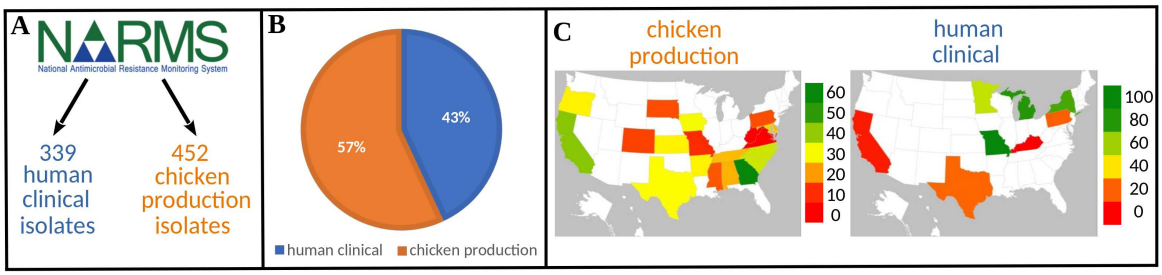



**Figure 3.2.** Graphical description of human clinical and poultry production isolates. FASTQ reads or FASTA assemblies for each E. coli isolate was downloaded from the NARMS GenomeTrackr for chicken production and human clinical isolates (A). Isolates were selected from a submission date of 01/01/2018 – 03/30/2020 using the filtering criteria described in the Materials and Methods. After quality control and filtering, 452 chicken production isolates and 339 human clinical isolates remained (B). Isolates geographic location across the United States (C).

### 3.4.2 AMR gene distribution not explained by isolation year or location

We conducted principal component analysis (PCA) to determine if the presence or absence of identified AMR genes corresponded to the state origin of the isolates. Overlapping 95% confidence intervals were observed for all isolates in all states **(Figure S3.1).** We also conducted PCA analysis on the presence/absence matrix of these AMR genes to assess impact of time of sampling. Separation of isolates based on their associated isolation year was also not observed (**Figure S3.2**). Accordingly, year and location will not be factored in further analysis.

### 3.4.3 Determining the phylogeny of E. coli isolates phylogroups

To determine if there is an evolutionary based separation between chicken production and human clinical isolates, we performed a whole genome alignment to the *E. coli* K12-MG1655 reference genome (accession U00096). Single-nucleotide polymorphisms (SNPs) with respect to the reference genome were used to generate a phylogenetic tree. In addition, we used ClermonTyping (Clermont *et al.*, 2019) to classify the isolates into phylogroups. *E. coli* isolates within the phylogenetic tree clustered into single clades and chicken production isolates were less divergent from the *E. coli* K12 reference genome than human clinical isolates (**Figure 3.1**). *E. coli* isolates were phylotyped into one of the phylogroups A, B1, B2, C, clade I, D, E, ‘E or clade I’, F, and G (**Table 3.1**). The majority of the *E. coli* genomes in our dataset belonged to phylogroups A, B1, and B2. Phylogroups B1 and B2 were the most represented groups with each accounting for ~23% of the total *E. coli* isolates. Phylogroup determination for one isolate was unsuccessful (identified as unknown) but the isolate clustered with phylogroup A isolates on the SNP-based tree.

Phylogroup classification is useful for the rapid and easy identification of potentially virulent and AMR isolates (Clermont *et al.*, 2019). Based on the Clermont quadruplex *in silico* method, our study isolates represented all seven main phylogroups (A, B1, B2, C, D, E and F), cryptic clade I, and the newly identified phylogroup G. One isolate was classified as unknown. We found significant differences in the proportion of times phylogroups A, B1, B2, E, and G were seen with respect to their host-origin. Phylogroups A, B1 and G were found at higher frequencies in chicken production isolates while B2 and E were higher in human clinical isolates (**Table 3.1**). Eighty-one percent of chicken production isolates belonged to phylogroup B1 whereas 85% of human clinical isolates were classified as B2. The most divergent subclade within B2 consisted of closely related human clinical isolates, many of which originated from the same BioProject (PRJNA489090) which included many isolates from the human infant gut. Although the B2 phylogroup was dominated by human clinical isolates, there were a few chicken isolates nested within the B2 clades. Phylogroup E was dominated by closely related human clinical isolates (81%) and was the largest cluster of human clinical isolates within the phylogenetic tree. Isolates identified as belonging to the *Escherichia* cryptic clade I phylogroup (NCBI genome accession numbers: SRR9984913, SRR10687720, SRR9852843, SRR9875260, SRR9875260, SRR10267809, SRR10687983, SRR9852805) (Clermont *et al.*, 2013) clustered together within the phylogenetic tree and were all of chicken origin. No human clinical *E. coli* isolates were identified as belonging to clade I even though clade I has been reported to have 2 to 3% prevalence in human *E. coli* isolates (Ingle *et al.*, 2011; Smati *et al.*, 2013; Ahumada-Santos *et al.*, 2020). Phylogroup G was dominated by chicken production isolates (95%).

The analysis above shows that in most cases the phylogroup classification supported the phylogenetic distribution within the SNP-based tree, however, we found several instances where

human clinical and chicken production isolates were closely related. Likewise, we found cases where the SNP-based classification contradicted the phylogroup classification. For instance, human clinical strains mainly belong to phylogroup A (Ingle *et al.*, 2011; Smati *et al.*, 2013; Ahumada-Santos *et al.*, 2020), however, we identified 85.58% of our chicken production isolates belong to this phylogroup. Taken together, our isolate collection serves as a good representation and distribution of the currently described *E. coli* phylogroups, with a few exceptions.

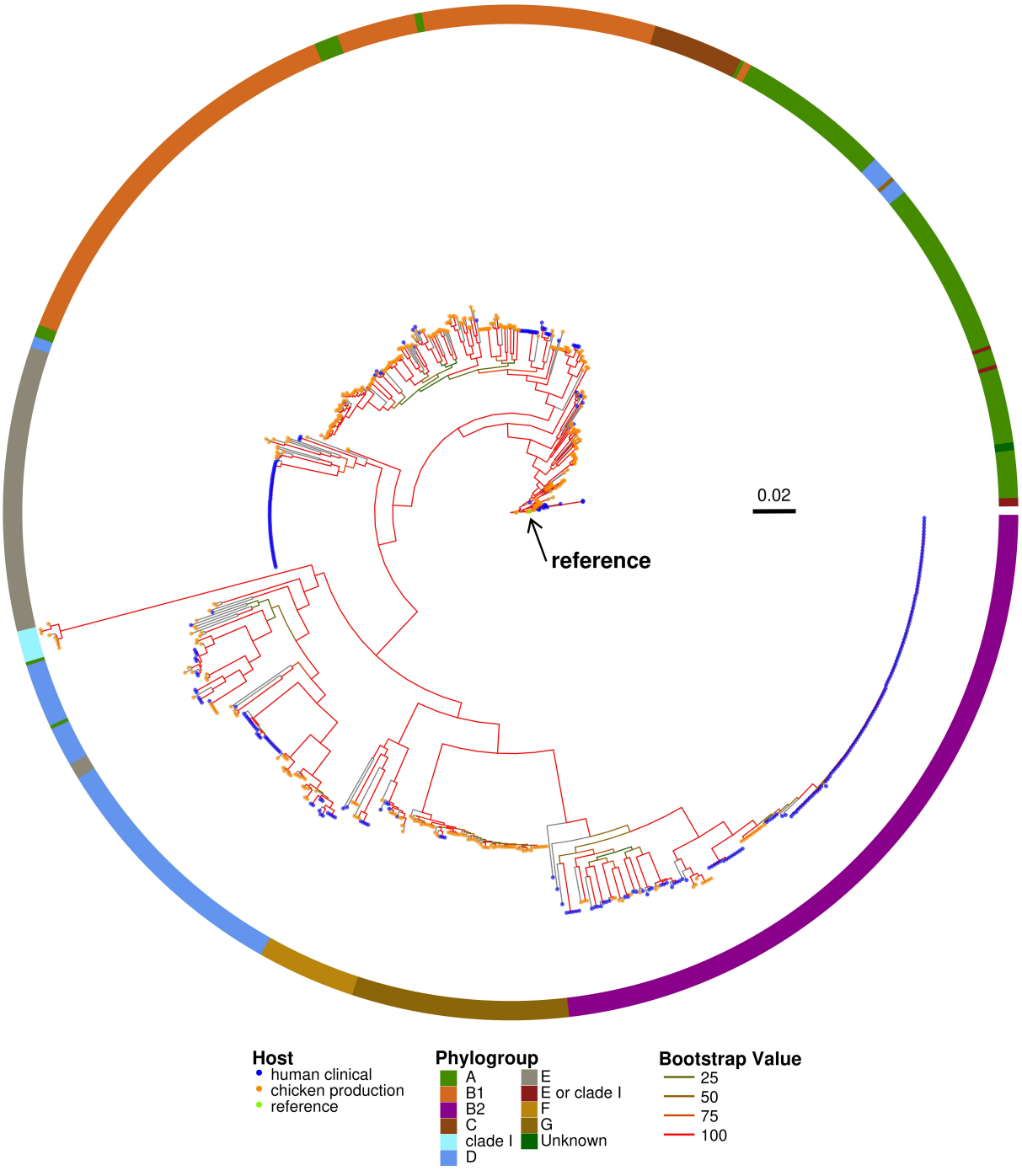

0.02
reference
Host
human clinical
chicken production
reference
Phylogroup
A
B1
B2
C
clade I
D
E
E or clade I
F
G
Unknown
Bootstrap Value
25
50
75
100

**Figure 3.3.** Whole genome SNP-based maximum-likelihood phylogenetic tree. SNPs found between the *E. coli* K12-MG1655 (accession U00096) reference genome and each of the 791 *E. coli* isolates was used for alignment using Snippy v4.6.0 (Snippy 2018). The resulting whole genome SNP alignment was used to construct maximum-likelihood phylogenetic trees under a GTR+GAMMA model using RAxML with 100 bootstraps (RAxMLHPC-PTHREADS) v8.2.12. Strength of nodal support indicated by branch color: red(100), orange(50>x≤75), darkorange(25>x≤50), green(0>x≤25). (**Ring**) Phylogroups identified using ClermonTyping v1.4.0. Phylogroups labeled by color: A(green), B1(orange), B2(magenta), C(brown), clade I(aquamarine), D(blue), E(gray), E or clade I(maroon), F(gold), G(darkgold), Unknown(darkgreen).

**Table 3.1.** Distribution of phylogroups in chicken production and human clinical isolates. Phylogroup for each isolate was determined using ClermonTyping v1.4.0 using default settings for each isolate genome assembly. Phylogroups composed of different hosts are indicated with (*), as determined by maximum likelihood ratio test with a Benjamini-Hochberg adjusted p value < 0.05.

| **phylogroup** | **number of isolates (% total) (n = 791)** | **chicken production (% total) (n = 452)** | **human clinical (% total) (n = 339)** |
|---|---|---|---|
| A | 132 (16.62%) | 109 (85.58%)***** | 23 (17.42%)***** |
| B1 | 183 (23.05%) | 148 (80.87%)***** | 35 (19.13%)***** |
| B2 | 184 (23.17%) | 28 (15.21%)***** | 156 (84.78%)***** |
| C | 23 (2.90%) | 18 (78.26%) | 5 (21.74%) |
| Clade I | 8 (1.01%) | 8 (100%) | 0 (0%) |
| D | 103 (12.97%) | 56 (54.37%) | 47 (45.63%) |
| E | 75 (9.45%) | 14 (18.67%)***** | 61 (81.33%)***** |
| E or Clade I | 4 (0.05%) | 4 (100%) | 0 (0%) |
| F | 26 (3.27%) | 15 (57.69%) | 11 (42.31%) |
| G | 55 (6.93%) | 52 (94.55%)***** | 3 (5.45%)***** |
| Unknown | 1 (0.13%) | 0 (0%) | 1 (100%) |

### 3.4.4 Chicken production and human clinical isolates harbor a diverse set of AMR genes and mutations

We used the ResFinder (Bortolaia *et al.*, 2020) to identify acquired AMR genes in each *E. coli* genome. Additionally, we used R2R and the Comprehensive Antimicrobial Resistance Database (CARD) via the Resistance Gene Identifier (RGI) (Alcock *et al.*, 2020) (see Materials and Methods for details) to identify all AMR genes, including those encoded on the bacterial chromosome and gene mutations that can result in resistant genotypes. We found 42 acquired AMR genes predicted to confer resistance to 21 drug classes (**Table 3.2**). Fifty-seven percent of the isolates carried at least 1 acquired AMR gene while one human clinical isolate carried 21 acquired AMR genes. Chicken production isolates were more likely to harbor acquired AMR genes (65.7%) than human clinical isolates (46.3%) (adjusted $p << 0$). Chicken production *E. coli* isolates carried more acquired AMR genes per isolate (Range = 0 to 8, mean = 1.97 per isolate) compared to human clinical isolates (Range = 0 to 21, mean = 1.87 per isolate). Fosfomycin (*fosA7*) and peptide antibiotics (*mcr-9*) were significantly higher in chicken production isolates while aminoglycosides *(aac(6')-Ib-cr*) and fluoroquinolones (*qrnB19)* were significantly higher in human clinical isolates (**Figure 3.4A, Table S3.3**).

**Table 3.2.** Acquired antimicrobial resistance genes identified within chicken production and human clinical isolate FASTA assemblies using ResFinder. ResFinder output hits for each gene were filtered (>=80% identity to database reference query). Corresponding resistance-conferring drug class for each gene was identified using The Comprehensive Antibiotic Resistance Database (CARD).

| Gene | Drug Class | Gene | Drug Class |
|---|---|---|---|
| *aac(3)-IIa* | aminoglycoside antibiotic | *dfrA1* | diaminopyrimidine antibiotic |
| *aac(3)-IId* | aminoglycoside antibiotic | *dfrA12* | diaminopyrimidine antibiotic |
| *aac(3)-IV* | aminoglycoside antibiotic | *dfrA14* | diaminopyrimidine antibiotic |
| *aac(6')-Ib-cr* | fluoroquinolone antibiotic, aminoglycoside antibiotic | *dfrA17* | diaminopyrimidine antibiotic |
| *aadA1* | aminoglycoside antibiotic | *erm(B)* | lincosamide antibiotic, macrolide antibiotic, streptogramin antibiotic |
| *aadA2* | aminoglycoside antibiotic | *floR* | phenicol antibiotic |
| *aadA5* | aminoglycoside antibiotic | *formA* | disinfecting agents |
| *aph(3'')-Ib* | aminoglycoside antibiotic | *fosA7* | fosfomycin |
| *aph(3')-Ia* | aminoglycoside antibiotic | *mcr-9* | peptide antibiotic |
| *aph(3')-IIa* | aminoglycoside antibiotic | *mdf(A)* | rhodamine, tetracycline antibiotic, benzalkonium chloride |
| *aph(3')-III* | aminoglycoside antibiotic | *mph(A)* | macrolide antibiotic |
| *aph(4)-Ia* | aminoglycoside antibiotic | *qacE* | disinfecting agents and intercalating dyes |

| | | | |
|---|---|---|---|
| *aph(6)-Ic* | aminoglycoside antibiotic | *qnrB19* | fluoroquinolone antibiotic |
| *aph(6)-Id* | aminoglycoside antibiotic | *rmtB* | aminoglycoside antibiotic |
| *blaCMY-2* | penam, carbapenem, cephalosporin, cephamycin | *sitABCD* | disinfecting agents |
| *BlaCTX-M-14* | cephalosporin | *sul1* | sulfonamide antibiotic |
| *BlaCTX-M-15* | penam, cephalosporin | *sul2* | sulfonamide antibiotic |
| *BlaOXA-1* | penam, carbapenem, cephalosporin | *sul3* | sulfonamide antibiotic |
| *blaOXA-244* | penam, carbapenem, cephalosporin | *tet(A)* | tetracycline antibiotic |
| *blaTEM-1B* | monobactam, penam, penem, cephalosporin | *tet(B)* | tetracycline antibiotic |
| *catA1* | phenicol antibiotic | *tet(C)* | tetracycline antibiotic |

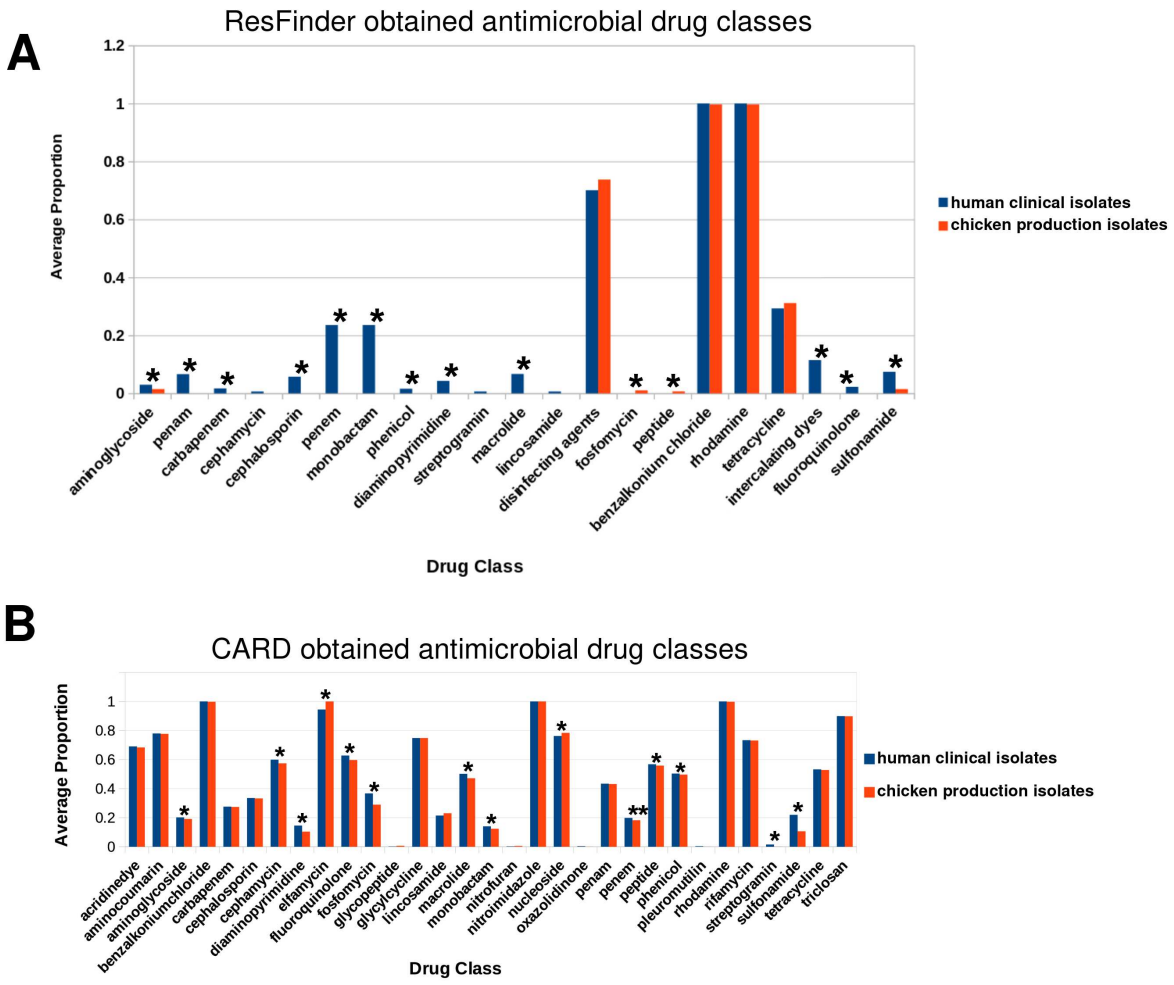



**Figure 3.4.** Average proportions of identified antimicrobial resistance genes in human clinical and poultry production isolates. (**A**) Acquired antimicrobial resistance genes identified within chicken production and human clinical isolate FASTA assemblies using ResFinder. ResFinder output hits for each gene were filtered (>=80% identity to database reference query). (B) Antimicrobial resistance drug classes identified within chicken production and human clinical isolate FASTA assemblies using Resistance Gene Identifier (RGI) and ABRICATE employed via Reads2Resistome. AMR database hits were filtered (>=95% identity to database reference query) and the highest hit from each database was retained. Genes conferring resistance to drug classes were enumerated for each isolate and a proportion was calculated using the total number of genes in the study population conferring resistance to a given drug class. Drug classes significantly differing between chicken production isolates and human clinical isolates were determined by Wilcoxon rank-sum test and (*) indicates a Benjamini-Hochberg adjusted p value < 0.05.

Using the CARD database we found 208 AMR genes, including all but 3 genes (*aac(3')-IIa, sitABCD* and *formA)* identified by ResFinder (**Table S3.1**). A majority of AMR genes were shared between human clinical isolates and chicken production isolates (**Figure 3.5A**). The AMR genes were predicted to confer resistance to 31 drug classes including cephalosporins, fluoroquinolones, penams and tetracyclines (**Figure 3.4B**). We found AMR genes for 31 drug classes in human clinical isolates, while 28 of the 31 drug classes were found in chicken production isolates. AMR genes absent from chicken production isolates were predicted to confer resistance to streptogramin (*ermB* and *msrA*), oxazolidinone (*msrA*) and pleuromutilin (*msrA* and *taeA*). We found significant differences in antimicrobial drug classes between chicken production and human clinical isolates for 14 of the 31 drug classes (adjusted $p < 0.05$, **Figure 3.4B, Table S3.4**). Genes predicted to confer resistance to diaminopyrimidine, phenicol, sulfonamide, penem, aminoglycoside, fluoroquinolone, macrolide, and peptide antibiotics were more prevalent within human clinical isolates (adjusted $p << 0$, **Figure 3.4B**). Genes predicted to confer cephamycin resistance – a highly important antimicrobial group for human medicine (World Health Organization, 2019) – were more prevalent in human clinical isolates (adjusted $p << 0$, **Figure 3.4B**) compared to chicken isolates. Genes predicted to confer resistance to elfamycin and nucleoside antibiotics were more prevalent within the chicken production isolates compared to human isolates (adjusted $p << 0$, **Figure 3.4B**). The frequency of genes predicted to confer resistance to tetracycline was similar (adjusted $p = 0.27$, **Figure 3.4B**) between host origin.

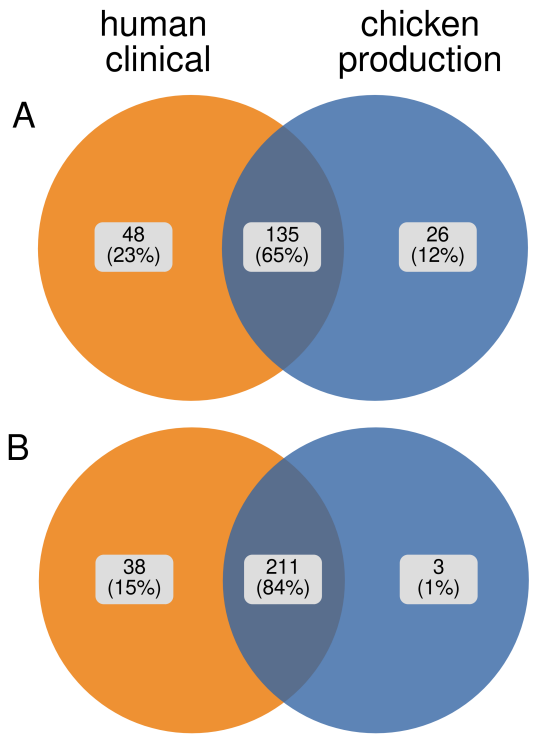



**Figure 3.5.** Venn diagrams depicting the identified antimicrobial resistance genes identified via Reads2Resistome (RGI) and The Comprehensive Antibiotic Resistance Database and Resfinder (A) and virulence factors identified via Reads2Resistome (ABRICATE) and the Virulence Factor Database.

It is important to mention that the majority of AMR genes identified by CARD are likely chromosomally encoded and may not be mobilizable. Many of the AMR genes harbored on the chromosome confer levels of resistance that are not medically relevant but may confer fitness advantages to these isolates in their native environments (Bagel *et al.*, 1999; Girgis, Hottes and Tavazoie, 2009). Contrastingly, most of the AMR genes identified by ResFinder are more likely to be mobilized within bacterial populations, via plasmids or other mobile genetic elements (MGEs), therefore are more likely to spread between/within *E. coli* strains originating from human and food animal populations. We hypothesized that a majority of the AMR genes identified by ResFinder were located either on plasmid or chromosomal contigs, and not located in close proximity to other MGEs such as genomic islands, insertion sequences, gene cassettes,

etc. This hypothesis was proposed due to all but 3 ResFinder identified AMR genes being also identified by CARD which are majorly present on chromosomal contigs. To test this hypothesis, we submitted 4 isolates to VRProfile2 (Wang *et al.*, 2022), an online bacterial mobile element detection pipeline capable of predicting AMR gene presence on plasmid or chromosomal contigs or those in close proximity to other MGEs. VRProfile2 results indicated AMR genes found by Resfinder were located on plasmid contigs, as was the case for aph(3')-III. No AMR genes were identified in close proximity to any other MGE types (**Table S3.5**).

Overall, we observed a high prevalence of genes that can confer resistance to antimicrobials considered highly and critically important to human medicine in both chicken production and human clinical *E. coli* isolates. The prevalence of these genes in both human and chicken *E. coli* isolates indicates that these populations may serve as a possible reservoir of antimicrobial resistance.

### 3.4.5 The presence of Type III secretion system genes classified the E. coli isolates into two separate clusters

To determine the genes driving the separation of these isolates into the two distinct clusters, we performed Random Forests classification on the filtered AMR gene and VF table (**Table S3.1**). This reclassification resulted in 100% correct classification of all isolates to the two clusters (A and B) (**Table S3.6**). T3SS genes play a key role in the virulence of many Gram-negative bacterial pathogens (Tseytin *et al.*, 2018) and were prevalent in all isolates from cluster B and completely absent from cluster A isolates. Twenty-five percent of all human clinical isolates (n = 339) carried genes for the T3SS operon (Tseytin *et al.*, 2018) whereas only 9.29% of all chicken production isolates (n = 352) harbored the T3SS operon. Along with genes from the T3SS operon, the intimin gene (*eae*) that is required for intimate adherence and virulence in both

humans and animals, was a high contributor to the mean decreasing accuracy and was present in 100% of cluster B isolates but completely absent from cluster A isolates (**Figure 3.7A**). Reducing the input AMR and virulence factors to only include the T3SS outer ring protein (*escD)* and the T3SS secretin *(escC),* two genes that encode oligomerizing proteins for the T3SS and that are essential for T3SS functionality (Slater *et al.*, 2018), resulted in 100% correct classification of all isolates to either cluster A or B. Our results suggest that cluster B isolates that harbor T3SS may have the potential to deliver effector proteins that promote bacterial colonization, replication, and transmission in human host cell cytoplasm, while cluster A isolates may lack this capability. Nevertheless, we found both clusters contained isolates from human clinical and chicken production settings.

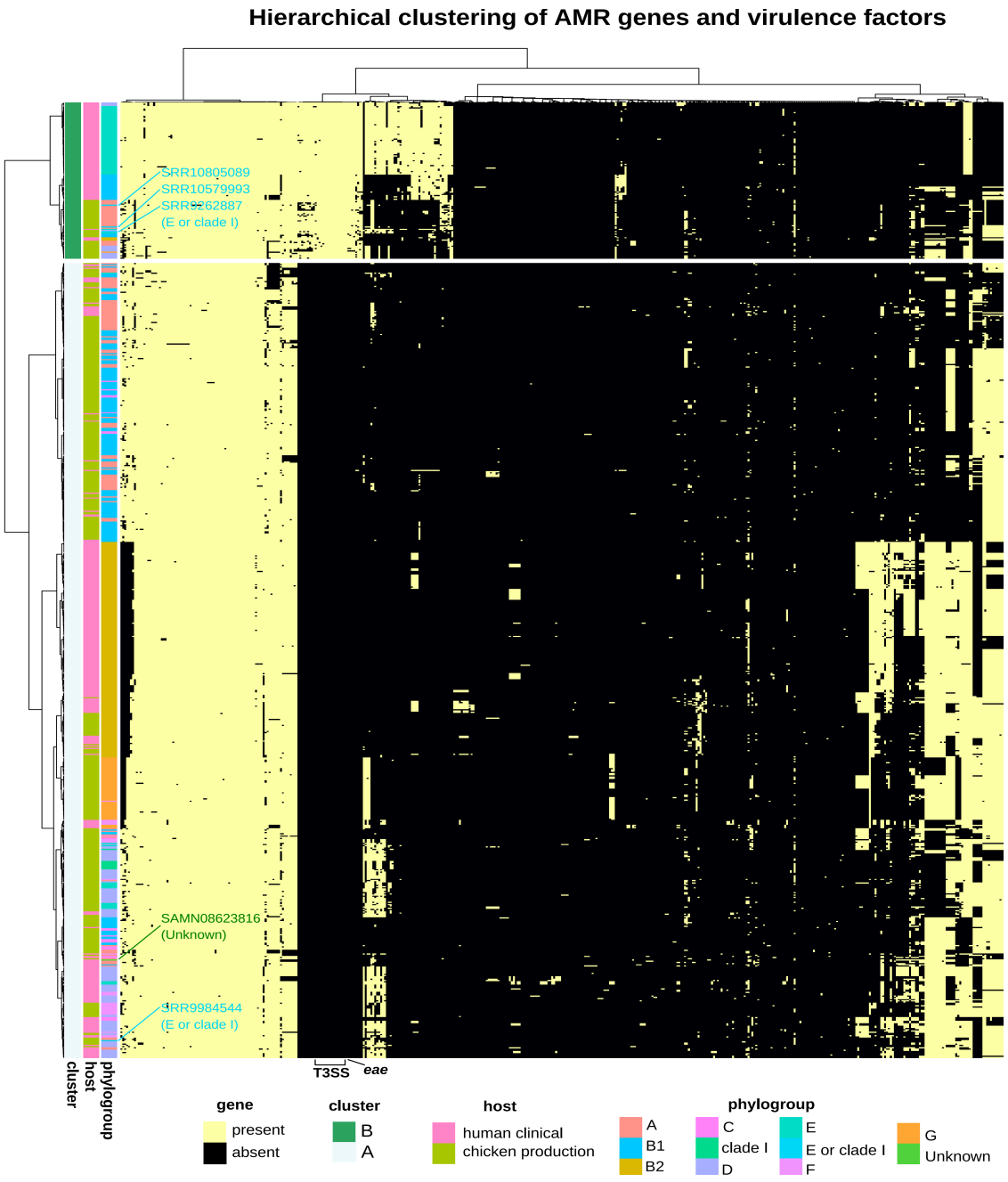



**Figure 3.6.** Heatmap of AMR genes and virulence factors across all isolates (Table S1). Heatmap was generated in R v4.0.4 with pheatmap v1.0.12 (clustering_method = "average" (UPGMA), clustering_distance_cols = "binary", clustering_distance_rows = "euclidean") using the filtered AMR gene and virulence factor table. Isolates from phylogroups Unknown and 'E or clade I' are highlighted.

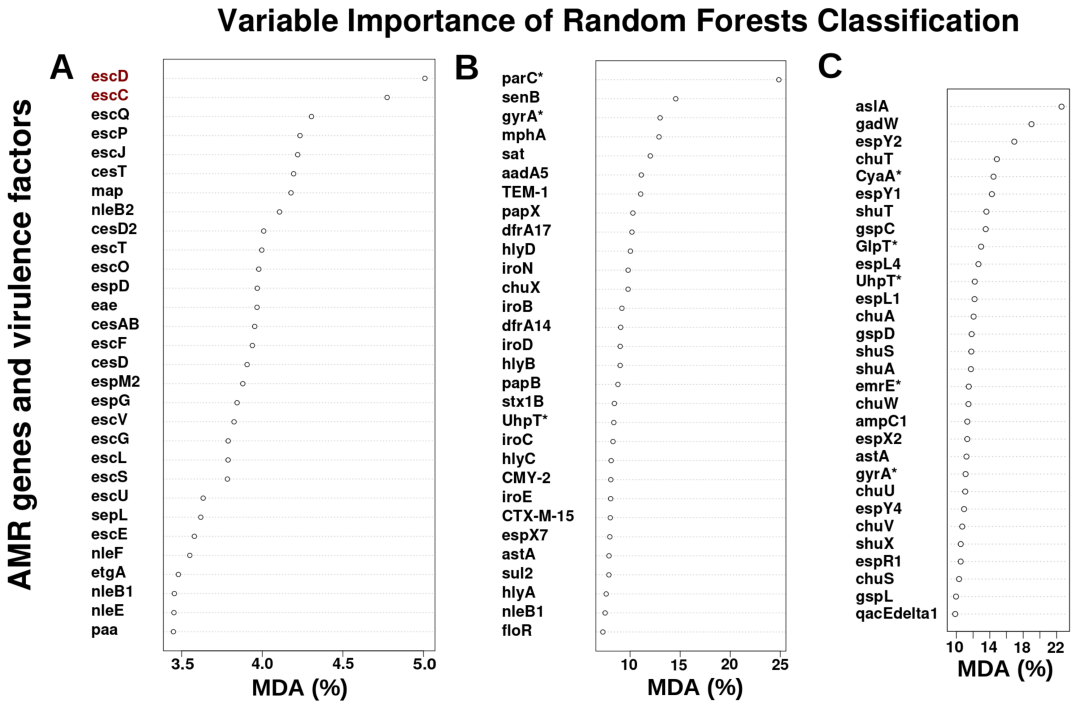



**Figure 3.7.** Mean decreasing accuracy (MDA) plot of variables in Random Forests classification models. MDA represents how much accuracy the model losses by exclusion of each variable (AMR gene or virulence factor). Random Forests classification was performed on AMR gene and virulence factor table (see materials and methods for details). (A) Resulting MDA plot after classification of isolates into hierarchical clustering identified clusters A and B. Genes highlighted in red, on their own, result in 100% classification of isolates into hierarchical clustering identified clusters A and B. (**B**) Resulting MDA plot after classification of isolates into chicken production isolates and human clinical isolates. (**C**) Resulting MDA plot after classification of isolates into ClermonTyping identified phylogroups. Random Forests classification in R was performed using the randomForest v4.6-14. Genes with (*****) indicate those with mutations: [parC(S80I), gyrA(S83L (n = 30), S84L (n = 2), D87Y (n= 275)), UtpT(E350Q), CyaA(S352T), GlpT(E448K)] (**Table S3.7**).

### 3.4.6 AMR genes and virulence factors classified E. coli isolates by their host origin

To assess which AMR genes and VF differentiate isolates based on their host origin we performed Random Forests classification (**Figure 3.7B**). This classification resulted in 1.1% (5 isolates) error in classifying chicken isolates and a 5.6% (19 isolates) error in classifying human

clinical isolates (**Table S3.8**). All misclassified isolates belonged to the hierarchical cluster A (**Table S3.9**) described earlier. The *E. coli* topoisomerase IV (*parC*) gene with the S80I mutation, was the highest contributor to the mean deceasing accuracy. The *parC* gene plays a critical role in DNA replication and confers reduced susceptibility to fluoroquinolones if the relevant parC and gyrase A *(gyrA*) mutations are present together (Bagel *et al.,* 1999; Del Mar Tavío *et al.,* 1999; Sáenz *et al.,* 2003; Johnning *et al.,* 2015). One hundred and eighty-five isolates (1 chicken production and 184 human clinical) had *parC* mutations and 215 isolates (25 chicken production and 190 human clinical) had *gyrA* mutations. Four human clinical isolates had *parC* mutations but no *gyrA* mutation, while 34 isolates (24 chicken production and 10 human clinical) had *gyrA* mutations but no *parC* mutation. One *E. coli* isolate, misclassified as a human isolate, harbored both the *parC* (S80I) and *gyrA* (S83L, S84L and D87Y) mutations (**Table S3.7**). The remaining 572 isolates did not have *parC* or *gyrA* mutation. In summary, *parC* and *gyrA* mutations that confer resistance to fluoroquinolones were present in the many of human clinical isolates and it was a major factor that separated the *E. coli* isolates into their respective host origins.

The iron acquisition genes (*iroN, iroD* and *iroE)* were high contributors to the mean decreasing accuracy of random forest classification. The catecholate siderophore uptake system (*iroBCDEN)* plays a critical role in virulence since iron is required for many cellular processes but it is limited in host sites of extraintestinal infections (Gao *et al.*, 2012). This gene cluster has been reported to be located in a pathogenicity island on the chromosome and has been found on plasmids (Sorsa *et al.*, 2003; Oladeinde *et al.*, 2022). However, this gene cluster was absent from all misclassified chicken isolates but was present in 56% and 7% of chicken production isolates and human clinical isolates, respectively (**Table S3.1**). The enterotoxin TieB protein (*senB)* was the second most important factor and was absent from all chicken production isolates regardless

of their Random Forests classification but present in 28% of human clinical isolates. Only one misclassified human clinical isolate harbored *senB*. These data suggest that the human and chicken isolates in our dataset carry distinct AMR determinants and VF which differentiate them from one another.

### 3.4.7 AMR genes and VF content can classify E. coli isolates by phylogroup

Next, we hypothesized that if the AMR genes and VF content is specific to each phylogroup, we should be able to classify the *E. coli* isolates correctly into their respective phylogroups based on their AMR and VF profile. To evaluate this, we used Random Forests to classify the isolates into their identified phylogroups. Classification resulted in the following error rates: phylogroup A (11%), B1(3%), B2(0%), C(48%), clade I (0%), D(0%), E(5%), 'E or clade I' (100%), F(4%), G(2%), and Unknown (100%) (**Table S3.10**). Unlike host origin classification, reducing the input table for phylogroup classification resulted in an increase in the misclassification error. Interestingly, the top 30 factors, needed to classify isolates into their respective phylogroups, consisted of 8 AMR genes and 22 VF (**Figure 3.7C**), suggesting that VF have a greater influence than AMR genes for phylogroup classification.

Six out of eleven VF functions found using the virulence factor database, (**Table S3.11**) were present in all phylogroups and the remaining functions were either entirely absent from individual phylogroups or present in some (**Table S3.2**). Interestingly, virulence functions relating to adherence, secretion systems and iron uptake were present in all phylogroups but functions relating to immune evasion were present in only phylogroups A, B1, B2, D and E. (**Table S3.2**). Thus, both chicken production and human clinical isolates have the potential to

serve as VF reservoirs. Albeit *E. coli* isolates of phylogroups C, 'E or clade I', clade I, G and F, lacked many of the VF functions compared to A, B1, B2, D and E.

We found that an isolate's AMR and VF profile was sufficient, in most cases, to determine its phylogroup. Additionally, misclassified isolates were supported by their AMR and VF similarity as seen using the UPGMA clustering approach (see Materials and Methods for details, **Figure 3.6**). For instance, the *E. coli* isolate that was identified by ClermonTyping as unknown was classified into phylogroup A based on its AMR gene and VF profile similarity with other phylogroup A, B1 and C isolates (**Figure 3.6**). The misclassification 'E or clade I' isolates (NCBI accession numbers: SRR10805089, SRR10579993, and SRR9262887) as phylogroups D, D, and E, respectively, was also supported by clustering in close proximity to other phylogroup D and E isolates (**Figure 3.6**). An additional 'E or clade I' isolate (NCBI accession no: SRR9984544) was misclassified as D but contained a similar AMR and VF profile as other phylogroup D isolates. The misclassification was due to the absence of iron uptake genes *shuS*, *shuA*, and *shuX*, as well as the non-Locus of Enterocyte Effacement (LEE)-encoded T3SS genes *espL4* and *espY1*. Although the AMR and VF harbored by an *E. coli* isolate provided sufficient signal for correct classification of phylogroups, it could not correctly classify all phylogroups identified by ClermonTyping. However, the misclassification of "Unknown" groups and E or clade I isolates was explained and supported by their clustering based on the similarity of their AMR gene and VF profiles (**Figure 3.6**). Similarly, the location of the "Unknown" isolate within the phylogenetic tree supports the Random Forests classification (**Figure 3.3**), i.e., clustered with phylogroup A isolates.

## 3.5 Discussion

In this study we utilized *E. coli* isolates obtained from the NARMS surveillance program to study differences in phylotype, AMR and VF between chicken production and human clinical isolates. To do this, we first developed a massively parallel pipeline (Reads2Resistome) to streamline our genomic analysis and to ensure that our results are repeatable and reproducible. We found that *E. coli* can serve as a reservoir of AMR genes and VF which supports previously reported findings (Cortés *et al.*, 2010; Katouli, 2010; Van Hoek *et al.*, 2011). We determined that a majority of AMR genes and VF identified were shared (65% and 84%, respectively) between human clinical isolates and chicken production isolates. Although there is a possibility of horizontal gene transfer in the environment, the E. coli isolates from this study were mostly host-specific, suggesting that strains from chicken production rarely colonize humans. We identified AMR genes that have been acquired, as well as those encoded on the chromosome that can acquire mutations that will result in reduced susceptibility to antibiotics. We confirmed the presence of 42 acquired AMR genes conferring resistance to 21 drug classes (**Table 3.2, Figure 3.4A, Figure 3.5A**). The drug classes identified include those considered critically and highly important to human medicine. For example, genes conferring resistance to fosfomycin, considered a high priority and critically important antimicrobial by the WHO (World Health Organization, 2019), were significantly more prevalent within the chicken production isolates. Utilizing the CARD database, a collection of all AMR genes, including those with mutations known to confer a level of resistance to known antimicrobials, we identified genes conferring resistance to 31 antibiotic drug classes in our human clinical isolates, 28 of which were observed in the chicken production isolates (**Figure 3.4B**). In addition, genes conferring resistance to these 31 drug classes were observed across a majority of identified phylogroups.

We determined that chicken production isolates carried a higher proportion of acquired AMR genes than human clinical isolates. Acquired AMR genes are those which have been obtained by a strain via mobile genetic elements such as plasmids, integrative conjugative elements, bacteriophages, insertion sequences and gene capture elements such as a site-specific recombination system (Van Hoek *et al.*, 2011). Acquired AMR genes are important in the context of animal-to-human transmission as they are more readily shared between bacteria and provide a route for dissemination of AMR genes. The high proportion of benzalkonium chloride is not surprising due to its high usage, as a disinfectant, in the food and medical industries (García and Cabo, 2018). The gene conferring resistance to benzalkonium chloride, *mdfA*, a multi-drug efflux system, also confers resistance to tetracycline and rhodamine. When comparing the presence of AMR genes from CARD, which includes genes with mutations resulting in a resistance genotype to acquired resistance genes from ResFinder, we observed a drastic decline in the number of drug classes a given isolate may present resistance to. The prevalence of tetracycline in human clinical and chicken production isolates is expected due to its widespread usage in both agriculture and human clinical medicine and is of concern because of its high importance to human medicine.

Of particular interest is our finding of widespread potential resistance to fluoroquinolone antibiotics within the chicken production isolates as their usage was banned in poultry in 2005 (Price *et al.*, 2007), especially given that our data were obtained from samples between 2018 and 2020. Fluoroquinolone is a critically important antimicrobial drug class to human medicine due to its use in treatment of *Campylobacter* spp., invasive *Salmonella* spp. and multidrug resistant *Shigella* spp. infections (World Health Organization, 2019). Further investigation of the causes for fluoroquinolone resistance persistence in poultry is needed to ensure its efficacy in treating human infections. Taken together, we identified many antimicrobial drug classes which were

highly present, but expected, in our study population. However, the presence of AMR genes conferring resistance to antibiotics banned for use in poultry, or food animal production in general, raises concern for their future effectiveness in treating human infections and diseases.

We found 11 virulence functions, conferred by 251 identified VFs, (**Table S3.11, Figure 3.5B**) to be present in both chicken and human clinical isolates. Based on hierarchical clustering of T3SS genes, the *E. coli* isolates were classified into two distinct clusters, A and B. The T3SS is associated with pathogens which can adhere to the epithelial surface of the host, many of which can cause diarrheal disease; the second leading cause of death in children globally (Deborah Chen and Frankel, 2005; Black *et al.*, 2010). Cluster A contained a higher representation of chicken production isolates while cluster B comprised mainly of human clinical isolates. This result suggests that *E. coli* isolates from chicken production and human clinical isolates differ in their carriage of T3SS genes. The T3SS is essential for virulence and colonization of the human gut and is the genetic basis for enteropathogenic *E. coli* classification (Deborah Chen and Frankel, 2005). Additionally, the T3SS in poultry contributes to the virulence in avian pathogenic *E. coli (Wang et al.*, 2016). This result suggests that the T3SS genes can be used as targets for a rapid on-farm and clinical diagnostic detection of virulent *E. coli* strains in chicken production and hospital settings. Identifying a lower proportion of chicken production isolates harboring the T3SS is a positive finding as the T3SS increases virulence in both humans and chickens.

Classification into host origin was successful with little error. Two of the most important factors for this separation were the *parC* gene (S80I) and *gyrA* [S83L (n = 30), S84L (n = 2) and D87Y (n = 275)] which are both linked to fluoroquinolone resistance. It has been documented that these resistance mutants are commonly found in environmental *E. coli* even in the absence

of fluoroquinolone selective pressures (Del Mar Tavío *et al.*, 1999). We speculate that these genes, and their respective mutations, which relate to DNA replication and transcription, may confer a fitness advantage. The secreted enterotoxin TieB, encoded by *senB*, was the second most important factor in separating human and chicken isolates. *senB* was only identified in human clinical isolates and it was expected since *senB* is typically associated with enteroinvasive and uropathogenic *E. coli* infection in humans (Mao *et al.*, 2012). Similarly, *TEM-1*, a broad-spectrum beta-lactamase conferring resistance to penicillins and first generation cephalosphorins, was also a high contributor to the separation of hosts and was only present in human linked isolates. While *TEM-1* has been identified in poultry isolates (Seo and Lee, 2018), cephalosporin usage is more common in beef production (FDA, 2018).

This study reemphasizes the utility of phylogroup classification as a convenient way to identify pathogenic strains. The most recent version of ClermonTyper (Clermont *et al.*, 2019) relies on five genes (*arpA*, *chuA*, yj*aA*, TspE4.C2, and *ybgD*) for phylogroup classification, only one of the quadruplex genes, *chuA* (outer membrane protein responsible for heme uptake) was present in our presence/absence table of identified AMR and VF. Even though *yjaA* (gene that differentiates phylogroups B2 and D isolates) was absent from the AMR gene and VF table, we were able to reclassify B2 and D isolates with 100% accuracy. Similarly, TspE4.C2 gene differentiates phylogroups A from B1 and its absence from our AMR gene and VF table resulted in 11% classification error for A isolates and 3% error for B1 isolates. The absence of 4 of the 5 ClermonTyping genes might have contributed to the 48% error in classifying phylogroup C. The ability to separate phylogroups F and G from B2 was recently accomplished with the addition of the genes *cfaB* (CFA/I fimbrial subunit B), specific to phylogroup G strains, and *ybgD* (fimbrial-like adhesin protein), specific to phylogroup F strains (Lu *et al.*, 2016; Clermont *et al.*, 2019).

The absence of these genes in our table only resulted in one phylogroup F (4 % error rate) and one phylogroup G misclassified isolates (2 % error rate). Our results show that phylotype classification based on AMR and VF could be achieved even when the majority of the genes used for Clermont *E. coli* phylotyping are missing.

The use of whole genome sequencing (WGS) is an effective tool to predict the AMR and virulence potential of a bacterium, however WGS still has its limitations. The depth of coverage is crucial for genome assembly and gene annotation, both of which can be significantly influenced by a coverage of less than 30X. Our study excluded many isolates from both chicken production and the human clinical settings due to low sequencing coverage. Due to the nature of the database chosen for isolate selection i.e., a surveillance system for food-borne and other enteric bacteria, we expected an over representation of isolates harboring AMR genes and VF, and an under representation of susceptible commensal and environmental *E. coli* isolates in our datasets.

Additionally, the choice of AMR and virulence databases is important as not all databases have the same entries and methods of detection. Here, we implemented Reads2Resistome which allowed us to incorporate the results from multiple databases. Our final AMR gene and VF were comprised of the highest percent hit based on similarity to the reference gene (cut off for gene coverage > 80%) from each database. Another limitation of this study is the lack of AMR or virulence phenotype data for each isolate. Therefore, the presence an AMR gene or VF in a genome is not a confirmation that the isolate is resistant to the antibiotic/drug class predicted or that the isolate can cause diseases.

## 3.6 Materials and methods

### 3.6.1 Study isolates, normalization, and assembly

One thousand, seven-hundred and eleven *Escherichia coli i*solates including 874 human clinical and 837 chicken production isolates, were selected from the United States Food and Drug Administration’s GenomeTrackr network via NCBI's Pathogen Detection Network. Agricultural poultry isolates were selected based on the following criteria: Species: "*E.coli and Shigella*", Location: "USA", Target Creation: "01/01/2018 - 03/30/2020", Host: "Poultry", "poultry", "*Gallus gallus*", "*Gallus gallus domesticus*". Human clinical isolates were selected based on the following criteria: Species: "*E.coli and Shigella*", Location: "USA", Target Creation: "01/01/2018 - 03/30/2020", Host: "*Homo sapiens*", "*Homosapiens*", "*homo sapiens*", "Human", "Human (infant)". Only isolates with a “Strain” denoted as “E. coli” were retained for downstream analysis, isolates identified as *Shigella* were not retained. If available, Sequence Read Archive (SRA) reads, in FASTQ format, were downloaded based on the BioSample ID for both agricultural and human isolates. In the event the SRA reads were not available, FASTA assemblies were downloaded from the associated BioProject submission. To ensure accurate and unbiased antimicrobial resistance identification, SRA reads were further filtered based on coverage greater than or equal to 20X. To guarantee inclusion of quality genome assemblies, FASTA assemblies were filtered based on having an N50 greater than 10,000bp. Isolates, 3 in total, were identified to be identical and were removed from downstream statistical analysis.

To compare genetic structure and contents, specifically AMR genes, virulence factors, and plasmid replicons across our study isolates we assembled the SRA FASTQ reads into genomes, annotated these genomes and annotated the downloaded FASTA assemblies. To reduce bias in

gene identification and assembly quality across isolates, SRA reads were normalized to 30X coverage using the insilico_read_normalization.pl script from Trinity RNA-Seq v2.11.0 (Trinity's GitHub). SRA reads were assembled using Reads2Resistome v0.0.2 (options: slidingwindow = 4:20, assembly = nonhybrid, threads 60). Resulting SRA assemblies were further filtered based on N50 (>10,000bp) and final coverage (>=20X) like the FASTA assemblies described earlier. The final study data set consisted of R2R assembled isolate genomes and downloaded FASTA assemblies totaling 791 *E. coli* isolates (452 chicken production and 339 human clinical, **Figure 3.2A**). Maps presenting the geographic distribution of isolates (Figure 2C) was generated using the 'Filled Map' chart type in Microsoft Excel from Microsoft Office 365 (Microsoft Corporation).

### 3.6.2 Phylogenetic and sequence type analysis

To evaluate the relatedness of the isolates with respect to their assembled genomes we constructed a phylogenetic tree utilizing Single Nucleotide Polymorphisms (SNPs). These SNPs were identified by comparing the *E. coli* K12-MG1655 (accession U00096) reference genome to each of the 791 *E. coli* isolates using Snippy v4.6.0 (Torsten Seemann, 2015). The resulting whole genome SNP alignment was then used to construct a maximum-likelihood phylogenetic tree using the GTR+GAMMA model of evolution and utilizing RAxML (RAxMLHPC-PTHREADS) v8.2.12 (Stamatakis, 2006; Lutteropp, Kozlov and Stamatakis, 2019). In addition, we performed 100 bootstraps to assess nodal support. Only 100 bootstraps were performed due to time and computational limitations; it took 75.49 days utilizing 64 compute cores and 504Gb memory to construct the phylogenetic tree and perform the 100 bootstraps. To determine which of the seven main phylogroups the isolates belong to, we used ClermonTyping v1.4.0 (Clermont *et al.*, 2019), utilizing the default settings, to determine the phylogroup designation (A, B1, B2,

C, D, E and F) for each isolate. Tree visualization, along with phylogroup and metadata overlay was done in R v4.0.4 using ggtree v2.4.1 (Yu, 2020), factoextra v1.0.7 (Alboukadel Kassambara, Fabian Mundt, 2020) and ggnewscale v0.4.5 (Campitelli, 2022). Sequence type of each isolate was determined using ECTyper v1.0.0 (Bessonov *et al.*, 2021) using default settings.

### 3.6.3 Antimicrobial resistance gene and virulence-factor characterization using Reads2Resistome

Reads2Resistome (R2R) (See CHAPTER 3) was used to characterize the antimicrobial resistance genes, virulence factors and prophage regions. Reads2Resistome utilizes ABRICATE v0.5 (Seemann, no date) to screen the assembled contigs against the following databases via BLAST and reports a percentage identity to the reference sequence. Databases used by ABRICATE, and provided by R2R, are: ARG-ANNOT antibiotic resistance gene database (Gupta *et al.*, 2014), the Comprehensive Antibiotic Resistance Database (CARD) (Alcock *et al.*, 2020), MEGARes Antimicrobial Database for High-Throughput Sequencing (Lakin *et al.*, 2017), NCBI AMRFinderPlus (Feldgarden *et al.*, 2019), PlasmidFinder (Carattoli *et al.*, 2014), and VirulenceFinder database (Liu *et al.*, 2019). Downloaded FASTA assemblies' resistome characterization was performed by the modified Reads2Resistome (R2R-0.0.1-Fasta-QC-Ann-Only.nf). Additionally, resistance gene identification was performed using the Resistance Gene Identifier v5.1.1 (Alcock *et al.*, 2020) (RGI) and acquired resistance genes were identified using ResFinder (Bortolaia *et al.*, 2020)r. AMR genes and virulence factors with the highest percent identity across all databases were selected for further analysis. The heatmap of AMR genes and virulence factors, using the table (described below, **Table S3.1**) was generated using pheatmap v1.0.12 in R v4.0.4 using the ‘average’ [unweighted pair group method with arithmetic mean (UPGMA)] clustering method (**Figure S3.1**) (Miller *et al.*, 2019). (See Supplemental Materials

and Methods for further description of R2R). Mobile genetic elements were determined for selected isolates using VRProfile2 (updated most recently on October 2nd, 2021 (Wang *et al.*, 2022).

### 3.6.4 Hierarchical clustering and statistical analysis of resistome and virulence factors

To evaluate the relationships between poultry and human isolates with respect to their AMR and virulence factors we employed hierarchical clustering and ordination analysis. AMR and virulence database hits from R2R and RGI were filtered utilizing >=95% and >=80% sequence identity to the reference database query, respectively. ResFinder results were filtered utilizing >=85% identity to the reference database query. A table was then generated based on the presence/absence of identified VF and AMR genes from all isolates (**Table S3.1**). A distance matrix was generated using the tanimoto metric via the Distance() (Miller *et al.*, 2019) function from the IntClust v0.1.0 (Marijke Van Moerbeke, 2019) package. hclust() from the stats v3.6.2 package (R Core Team, 2021) was then utilize to perform the hierarchical clustering under the 'average' (UPGMA) method using the determined optimal number of clusters. We identified the optimal number of clusters which separate the 791 isolates using the silhouette method implemented by the fviz_nbclust() function from the factoextra v1.0.7 package (Alboukadel Kassambara, Fabian Mundt, 2020). Drug classes associated with each identified AMR gene were determined using the output from RGI, in combination with the CARD database (Alcock *et al.*, 2020), and virulence factor associated functions were determined using the comparative tables from the Virulence Factor Database (Liu *et al.*, 2019). Genes conferring resistance to drug classes were enumerated for each isolate and a proportion was calculated using the total number of genes in the study population conferring resistance to a given drug class. Drug classes

significantly differing between hosts were determined by Wilcoxon rank-sum test using the wilcox.test() function from the stats v3.6.2 package (R Core Team, 2021). P-value adjustment for multiple comparisons was conducted using the p.adjust() function, using the Benjamini and Hochberg (“BH”) method, from the stats v3.6.2 package (R Core Team, 2021). Principal component analysis was conducted using the prcomp() function from the stats v3.6.2 package (R Core Team, 2021) and plots were generated using ggplot2 (Hadley, 2016). All analyses were done in R v4.0.4 utilizing RStudio v1.2.1106 (RStudio Team, 2016).

### 3.6.5 Random Forests classification

To determine the most influential AMR and virulence factors driving the separation of the isolates into 1) clusters identified via hierarchical clustering, 2) isolate host origin, and 3) identified phylogroups we employed Random Forests classification (Breiman, 2001). The described presence/absence table of AMR genes and virulence factors was used as input (**Table S3.1**). TuneRF() (Andy Liaw, Matthew Wiener, 2002) was used to determine the optimal mtry, with respect to out-of-bag error estimate, for RandomForest(). RandomForest() v4.6-14 (Andy Liaw, Matthew Wiener, 2002) was used for classification with nttree=200 based on the filtered AMR and virulence-factor presence/absence table. All plots were generated using ggplot2 v3.3.3 (Hadley, 2016). All analyses were done in R v4.0.4 utilizing RStudio v1.2.1106 (RStudio Team, 2016).

## 3.7 Data Availability

All raw isolate data, FASTQ and FASTA files, were obtained from and are available through NCBI using the ‘ncbi-accession’ numbers in **Table S3.1**. All other data are available upon request.

## 3.10 REFERENCES

Ahumada-Santos, Y.P., Báez-Flores, M.E., Díaz-Camacho, S.P., Uribe-Beltrán, M. de J., Eslava-Campos, C.A., Parra-Unda, J.R., Delgado-Vargas, F., 2020. Association of phylogenetic distribution and presence of integrons with multidrug resistance in Escherichia coli clinical isolates from children with diarrhoea. J. Infect. Public Health 13, 767–772. https://doi.org/10.1016/J.JIPH.2019.11.019

Alboukadel Kassambara, Fabian Mundt, 2020. factoextra: Extract and Visualize the Results of Multivariate Data Analyses.

Alcock, B.P., Raphenya, A.R., Lau, T.T.Y., Tsang, K.K., Bouchard, M., Edalatmand, A., Huynh, W., Nguyen, A.L.V., Cheng, A.A., Liu, S., Min, S.Y., Miroshnichenko, A., Tran, H.K., Werfalli, R.E., Nasir, J.A., Oloni, M., Speicher, D.J., Florescu, A., Singh, B., Faltyn, M., Hernandez-Koutoucheva, A., Sharma, A.N., Bordeleau, E., Pawlowski, A.C., Zubyk, H.L., Dooley, D., Griffiths, E., Maguire, F., Winsor, G.L., Beiko, R.G., Brinkman, F.S.L., Hsiao, W.W.L., Domselaar, G.V., McArthur, A.G., 2020. CARD 2020: Antibiotic resistome surveillance with the comprehensive antibiotic resistance database. Nucleic Acids Res. 48, D517–D525. https://doi.org/10.1093/nar/gkz935

Andy Liaw, Matthew Wiener, 2002. Classification and Regression by Randomforest.

Bagel, S., Hüllen, V., Wiedemann, B., Heisig, P., 1999. Impact of gyrA and parC mutations on quinolone resistance, doubling time, and supercoiling degree of Escherichia coli. Antimicrob. Agents Chemother. 43, 868–875. https://doi.org/10.1128/AAC.43.4.868/ASSET/F739B9CE-E8B5-442A-B569-4072FF1489AA/ASSETS/GRAPHIC/AC0490515001.JPEG

Bessonov, K., Laing, C., Robertson, J., Yong, I., Ziebell, K., Gannon, V.P.J., Nichani, A., Arya, G., Nash, J.H.E., Christianson, S., 2021. ECTyper: in silico Escherichia coli serotype and species prediction from raw and assembled whole-genome sequence data. Microb. Genomics 7. https://doi.org/10.1099/MGEN.0.000728

Biswas, K., Sharma, P., Joshi, S.R., 2019. Co-occurrence of antimicrobial resistance and virulence determinants in enterococci isolated from traditionally fermented fish products. J. Glob. Antimicrob. Resist. 17, 79–83. https://doi.org/10.1016/j.jgar.2018.11.012

Black, R.E., Cousens, S., Johnson, H.L., Lawn, J.E., Rudan, I., Bassani, D.G., Jha, P., Campbell, H., Walker, C.F., Cibulskis, R., Eisele, T., Liu, L., Mathers, C., 2010. Global, regional, and national causes of child mortality in 2008: a systematic analysis. The Lancet 375, 1969–1987. https://doi.org/10.1016/S0140-6736(10)60549-1

Bortolaia, V., Kaas, R.S., Ruppe, E., Roberts, M.C., Schwarz, S., Cattoir, V., Philippon, A., Allesoe, R.L., Rebelo, A.R., Florensa, A.F., Fagelhauer, L., Chakraborty, T., Neumann, B., Werner, G., Bender, J.K., Stingl, K., Nguyen, M., Coppens, J., Xavier, B.B., Malhotra-Kumar, S., Westh, H., Pinholt, M., Anjum, M.F., Duggett, N.A., Kempf, I., Nykäsenoja, S., Olkkola, S., Wieczorek, K., Amaro, A., Clemente, L., Mossong, J., Losch, S., Ragimbeau, C., Lund, O., Aarestrup, F.M., 2020. ResFinder 4.0 for predictions of phenotypes from genotypes. J. Antimicrob. Chemother. 75, 3491–3500. https://doi.org/10.1093/JAC/DKAA345

Boyle, D.P., Zembower, T.R., 2015. Epidemiology and Management of Emerging Drug-Resistant Gram-Negative Bacteria: Extended-Spectrum β-Lactamases and Beyond. Urol. Clin. 42, 493–505. https://doi.org/10.1016/J.UCL.2015.05.005

Breiman, L., 2001. Random Forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324

Campitelli, E., 2022. ggnewscale: Multiple Fill and Color Scales in ggplot2. https://doi.org/10.5281/ZENODO.2543762

Caniça, M., Manageiro, V., Abriouel, H., Moran-Gilad, J., Franz, C.M.A.P., 2019. Antibiotic resistance in foodborne bacteria. Trends Food Sci. Technol. 84, 41–44. https://doi.org/10.1016/j.tifs.2018.08.001

Carattoli, A., Zankari, E., Garciá-Fernández, A., Larsen, M.V., Lund, O., Villa, L., Aarestrup, F.M., Hasman, H., 2014. In Silico detection and typing of plasmids using plasmidfinder and plasmid multilocus sequence typing. Antimicrob. Agents Chemother. 58, 3895–3903. https://doi.org/10.1128/AAC.02412-14

CDC, 2019. Surveillance for Foodborne Disease Outbreaks United States , 2014 : Annual Report 1–16.

Clermont, O., Christenson, 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, 58–65. https://doi.org/10.1111/1758-2229.12019

Clermont, O., Dixit, O.V.A., Vangchhia, B., Condamine, B., Dion, S., Bridier-Nahmias, A., Denamur, E., Gordon, D., 2019. Characterization and rapid identification of phylogroup G in Escherichia coli, a lineage with high virulence and antibiotic resistance potential. Environ. Microbiol. 21, 3107–3117. https://doi.org/10.1111/1462-2920.14713

Comery, R., Thanabalasuriar, A., Garneau, P., Portt, A., Boerlin, P., Reid-Smith, R.J., Harel, J., Manges, A.R., Gruenheid, S., 2013. Identification of potentially diarrheagenic atypical enteropathogenic Escherichia coli strains present in Canadian food animals at slaughter and in retail meats. Appl. Environ. Microbiol. 79, 3892–3896. https://doi.org/10.1128/AEM.00182-13/SUPPL_FILE/ZAM999104460SO1.PDF

Cortés, P., Blanc, V., Mora, A., Dahbi, G., Blanco, J.E., Blanco, M., López, C., Andreu, A., Navarro, F., Alonso, M.P., Bou, G., Blanco, J., Llagostera, M., 2010. Isolation and Characterization of Potentially Pathogenic Antimicrobial-Resistant Escherichia coli Strains from Chicken and Pig Farms in Spain. Appl. Environ. Microbiol. 76, 2799–2805. https://doi.org/10.1128/AEM.02421-09

Croxen, M.A., Law, R.J., Scholz, R., Keeney, K.M., Wlodarska, M., Finlay, B.B., 2013. Recent advances in understanding enteric pathogenic Escherichia coli. Clin. Microbiol. Rev. 26, 822–880. https://doi.org/10.1128/CMR.00022-13/ASSET/1535F48F-03D1-4FCC-A2D1-996296B3EB69/ASSETS/GRAPHIC/ZCM9990924420008.JPEG

Deborah Chen, H., Frankel, G., 2005. Enteropathogenic Escherichia coli: unravelling pathogenesis. FEMS Microbiol. Rev. 29, 83–98. https://doi.org/10.1016/J.FEMSRE.2004.07.002

Del Mar Tavío, M., Vila, J., Ruiz, Joaquím, Ruiz, José, Martín-Sánchez, A.M., Jiménez De Anta, M.T., 1999. Mechanisms involved in the development of resistance to fluoroquinolones in Escherichia coli isolates. J. Antimicrob. Chemother. 44, 735–742. https://doi.org/10.1093/JAC/44.6.735

Dewey-Mattia, D., Manikonda, K., Hall, A.J., Wise, M.E., Crowe, S.J., 2018. Surveillance for Foodborne Disease Outbreaks - United States, 2009-2015. MMWR Surveill. Summ. 67, 1–11. https://doi.org/10.15585/MMWR.SS6710A1

Dho-Moulin, M., Morris Fairbrother, J., 1999. Avian pathogenic Escherichia coli (APEC). Vet. Res. BioMed Cent. 30.

Ewers, C., Antão, E.M., Diehl, I., Philipp, H.C., 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, 184–192. https://doi.org/10.1128/AEM.01324-08/SUPPL_FILE/TABLE_S3_AEM_01234_08.DOC

Ewers, C., Li, G., Wilking, H., Kießling, S., Alt, K., Antáo, E.M., Laturnus, C., Diehl, I., Glodde, S., Homeier, T., Böhnke, U., Steinrück, H., Philipp, H.C., Wieler, L.H., 2007. Avian pathogenic, uropathogenic, and newborn meningitis-causing Escherichia coli: How closely related are they? Int. J. Med. Microbiol. 297, 163–176. https://doi.org/10.1016/J.IJMM.2007.01.003

FDA, 2018. Summary Report on Antimicrobials Sold or Distributed for Use in Food-Producing Animals. Cent. Vet. Med. 1–14.

Feldgarden, M., Brover, V., Haft, D.H., Prasad, A.B., Slotta, D.J., Tolstoy, I., Tyson, G.H., Zhao, S., Hsu, C.H., McDermott, P.F., Tadesse, D.A., Morales, C., Simmons, M., Tillman, G., Wasilenko, J., Folster, J.P., Klimke, W., 2019. Validating the AMRFINder tool and resistance gene database by using antimicrobial resistance genotype-phenotype correlations in a collection of isolates. Antimicrob. Agents Chemother. 63. https://doi.org/10.1128/AAC.00483-19

Fenske, G.J., Scaria, J., 2021. Analysis of 56,348 Genomes Identifies the Relationship between Antibiotic and Metal Resistance and the Spread of Multidrug-Resistant Non-Typhoidal Salmonella. Microorg. 2021 Vol 9 Page 1468 9, 1468. https://doi.org/10.3390/MICROORGANISMS9071468

Ferri, M., Ranucci, E., Romagnoli, P., Giaccone, V., 2017. Antimicrobial resistance: A global emerging threat to public health systems. Crit. Rev. Food Sci. Nutr. 57, 2857–2876. https://doi.org/10.1080/10408398.2015.1077192

Gao, Q., Wang, X., Xu, H., Xu, Y., Ling, J., Zhang, D., Gao, S., Liu, X., 2012. Roles of iron acquisition systems in virulence of extraintestinal pathogenic Escherichia coli: salmochelin and aerobactin contribute more to virulence than heme in a chicken infection model. https://doi.org/10.1186/1471-2180-12-143

García, M.R., Cabo, M.L., 2018. Optimization of E. coli inactivation by benzalkonium chloride reveals the importance of quantifying the inoculum effect on chemical disinfection. Front. Microbiol. 9, 1259. https://doi.org/10.3389/FMICB.2018.01259/FULL

Girgis, H.S., Hottes, A.K., Tavazoie, S., 2009. Genetic Architecture of Intrinsic Antibiotic Susceptibility. PLOS ONE 4, e5629. https://doi.org/10.1371/JOURNAL.PONE.0005629

Guabiraba, R., Schouler, C., 2015. Avian colibacillosis: Still many black holes. FEMS Microbiol. Lett. 362, 118. https://doi.org/10.1093/femsle/fnv118

Gupta, S.K., Padmanabhan, B.R., Diene, S.M., Lopez-Rojas, R., Kempf, M., Landraud, L., Rolain, J.M., 2014. ARG-annot, a new bioinformatic tool to discover antibiotic resistance genes

in bacterial genomes. Antimicrob. Agents Chemother. 58, 212–220. https://doi.org/10.1128/AAC.01310-13

Gurevich, A., Saveliev, V., Vyahhi, N., Tesler, G., 2013. QUAST: Quality assessment tool for genome assemblies. Bioinformatics 29, 1072–1075. https://doi.org/10.1093/bioinformatics/btt086

Hadley, W., 2016. Ggplot2. Springer Science+Business Media, LLC, New York, NY.

Ingle, D.J., Clermont, O., Skurnik, D., Denamur, E., Walk, S.T., Gordon, D.M., 2011. Biofilm formation by and thermal niche and virulence characteristics of Escherichia spp. Appl. Environ. Microbiol. 77, 2695–2700. https://doi.org/10.1128/AEM.02401-10

Jakobsen, L., Garneau, P., Bruant, G., Harel, J., Olsen, S.S., Porsbo, L.J., Hammerum, A.M., Frimodt-Møller, N., 2012. Is Escherichia coli urinary tract infection a zoonosis? Proof of direct link with production animals and meat. Eur. J. Clin. Microbiol. Infect. Dis. 31, 1121–1129. https://doi.org/10.1007/S10096-011-1417-5/FIGURES/2

Jakobsen, L., Garneau, P., Kurbasic, A., Bruant, G., Stegger, M., Harel, J., Jensen, K.S., Brousseau, R., Hammerum, A.M., Frimodt-Møller, 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. https://doi.org/10.1099/jmm.0.033993-0

Johnning, A., Kristiansson, E., Fick, J., Weijdegård, B., Larsson, D.G.J., 2015. Resistance Mutations in gyrA and parC are Common in Escherichia Communities of both Fluoroquinolone-

Polluted and Uncontaminated Aquatic Environments. Front. Microbiol. 6, 1355. https://doi.org/10.3389/FMICB.2015.01355

Johnson, T.J., Kariyawasam, S., Wannemuehler, Y., Mangiamele, P., Johnson, S.J., Doetkott, C., Skyberg, J.A., Lynne, A.M., Johnson, J.R., Nolan, L.K., 2007. The genome sequence of avian pathogenic Escherichia coli strain O1:K1:H7 shares strong similarities with human extraintestinal pathogenic E. coli genomes. J. Bacteriol. 189, 3228–3236. https://doi.org/10.1128/JB.01726-06/ASSET/85203383-A758-4288-B983-30303F75676B/ASSETS/GRAPHIC/ZJB0080766960003.JPEG

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, 745–758. https://doi.org/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., 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, 7043–7050. https://doi.org/10.1128/AEM.01395-08/ASSET/35350ED5-B35D-41B0-A1E2-D880DD80BCFF/ASSETS/GRAPHIC/ZAM0220893520003.JPEG

Kaper, J.B., Nataro, J.P., Mobley, H.L.T., 2004. Pathogenic Escherichia coli. Nat. Rev. Microbiol. 2004 22 2, 123–140. https://doi.org/10.1038/nrmicro818

Katouli, M., 2010. Population structure of gut Escherichia coli and its role in development of extra-intestinal infections. Iran. J. Microbiol. 2, 59.

Khairy, R.M., Mohamed, E.S., Ghany, H.M.A., Abdelrahim, S.S., 2019. Phylogenic classification and virulence genes profiles of uropathogenic E. coli and diarrhegenic E. coli strains isolated from community acquired infections. PLOS ONE 14, e0222441. https://doi.org/10.1371/JOURNAL.PONE.0222441

Lakin, S.M., Dean, C., Noyes, N.R., Dettenwanger, A., Ross, A.S., Doster, E., Rovira, P., Abdo, Z., Jones, K.L., Ruiz, J., Belk, K.E., Morley, P.S., Boucher, C., 2017. MEGARes: An antimicrobial resistance database for high throughput sequencing. Nucleic Acids Res. 45, D574–D580. https://doi.org/10.1093/nar/gkw1009

Li, G., Cai, W., Hussein, A., Wannemuehler, Y.M., Logue, C.M., Nolan, L.K., 2012. Proteome response of an extraintestinal pathogenic Escherichia coli strain with zoonotic potential to human and chicken sera. J. Proteomics 75, 4853–4862. https://doi.org/10.1016/J.JPROT.2012.05.044

Liu, B., Zheng, D., Jin, Q., Chen, L., Yang, J., 2019. VFDB 2019: A comparative pathogenomic platform with an interactive web interface. Nucleic Acids Res. 47, D687–D692. https://doi.org/10.1093/nar/gky1080

Lu, S., Jin, D., Wu, S., Yang, J., Lan, R., Bai, X., Liu, S., Meng, Q., Yuan, X., Zhou, J., Pu, J., Chen, Q., Dai, H., Hu, Y., Xiong, Y., Ye, C., Xu, J., 2016. Insights into the evolution of pathogenicity of Escherichia coli from genomic analysis of intestinal E. coli of Marmota himalayana in Qinghai–Tibet plateau of China. Emerg. Microbes Infect. 5, e122. https://doi.org/10.1038/EMI.2016.122

Lutteropp, S., Kozlov, A.M., Stamatakis, A., 2019. A Fast and Memory-Efficient Implementation of the Transfer Bootstrap. bioRxiv. https://doi.org/10.1101/734848

Lynch, M., Painter, J., Woodruff, R., Braden, C., 2006. Surveillance for Foodborne-Disease Outbreaks — United States, 1998–2002. MMWR Surveill. Summ. 55, 1–34.

Mainil, J., 2013. Escherichia coli virulence factors. Vet. Immunol. Immunopathol. 152, 2–12. https://doi.org/10.1016/J.VETIMM.2012.09.032

Mao, B.H., Chang, Y.F., Scaria, J., Chang, Chih Ching, Chou, L.W., Tien, N., Wu, J.J., Tseng, C.C., Wang, M.C., Chang, Chao Chin, Hsu, Y.M., Teng, C.H., 2012. Identification of Escherichia coli Genes Associated with Urinary Tract Infections. J. Clin. Microbiol. 50, 449. https://doi.org/10.1128/JCM.00640-11

Marijke Van Moerbeke, 2019. IntClust: integration of multiple data sets with clustering techniques. Version 0.1.0.

Mellata, M., 2013. Human and avian extraintestinal pathogenic escherichia coli: Infections, zoonotic risks, and antibiotic resistance trends. Foodborne Pathog. Dis. https://doi.org/10.1089/fpd.2013.1533

Miller, D.L., Rexstad, E., Thomas, L., Marshall, L., Laake, J.L., 2019. Distance Sampling in *R*. J. Stat. Softw. 89. https://doi.org/10.18637/jss.v089.i01

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, 1177–1187. https://doi.org/10.1128/AEM.03524-14

Moulin-Schouleur, M., Schouler, C., Tailliez, P., Kao, M.R., Brée, A., Germon, P., Oswald, E., Mainil, J., Blanco, M., Blanco, J., 2006. Common virulence factors and genetic relationships

between O18:K1:H7 Escherichia coli isolates of human and avian origin. J. Clin. Microbiol. 44, 3484–3492. https://doi.org/10.1128/JCM.00548-06/ASSET/A4A5E7AD-BABB-4FEA-B194-89BEA5862078/ASSETS/GRAPHIC/ZJM0100667760003.JPEG

Nyachuba, D.G., 2010. Foodborne illness: is it on the rise? Nutr. Rev. 68, 257–269. https://doi.org/10.1111/J.1753-4887.2010.00286.X

Oladeinde, A., Abdo, Z., Zwirzitz, B., Woyda, R., Lakin, S.M., Press, M.O., Cox, N.A., Thomas, J.C., Looft, T., Rothrock, M.J., Zock, G., Plumblee Lawrence, J., Cudnik, D., Ritz, C., Aggrey, S.E., Liachko, I., Grove, J.R., Wiersma, C., 2022. Litter Commensal Bacteria Can Limit the Horizontal Gene Transfer of Antimicrobial Resistance to Salmonella in Chickens. Appl. Environ. Microbiol. https://doi.org/10.1128/aem.02517-21

Pan, Y., Zeng, J., Li, L., Yang, J., Tang, Z., Xiong, W., Li, Y., Chen, S., Zeng, Z., 2020. Coexistence of Antibiotic Resistance Genes and Virulence Factors Deciphered by Large-Scale Complete Genome Analysis. mSystems 5. https://doi.org/10.1128/MSYSTEMS.00821-19/SUPPL_FILE/MSYSTEMS.00821-19-S0001.DOC

Panth, Y., 2019. Colibacillosis in poultry: A review. J. Agric. Nat. Resour. 2, 301–311. https://doi.org/10.3126/JANR.V2I1.26094

Price, L.B., Lackey, L.G., Vailes, R., Silbergeld, E., 2007. The Persistence of Fluoroquinolone-Resistant Campylobacter in Poultry Production. Environ. Health Perspect. 115, 1035. https://doi.org/10.1289/EHP.10050

R Core Team, 2021. R: A language and environment for statistical computing.

Ramos, S., Silva, V., de Lurdes Enes Dapkevicius, M., Caniça, M., Tejedor-Junco, M.T., Igrejas, G., Poeta, P., 2020. Escherichia coli as Commensal and Pathogenic Bacteria among Food-Producing Animals: Health Implications of Extended Spectrum β-Lactamase (ESBL) Production. Anim. Open Access J. MDPI 10, 1–15. https://doi.org/10.3390/ANI10122239

RStudio Team, 2016. RStudio: Integrated Development Environment for R.

Russo, T.A., Johnson, J.R., 2003. Medical and economic impact of extraintestinal infections due to Escherichia coli: focus on an increasingly important endemic problem. Microbes Infect. 5, 449–456. https://doi.org/10.1016/S1286-4579(03)00049-2

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. Chemother. 51, 1001–1005. https://doi.org/10.1093/JAC/DKG168

Seemann, T., n.d. ABRICATE. Github.

Seo, K.W., Lee, Y.J., 2018. Prevalence and Characterization of β-Lactamases Genes and Class 1 Integrons in Multidrug-Resistant Escherichia coli Isolates from Chicken Meat in Korea. https://home.liebertpub.com/mdr 24, 1599–1606. https://doi.org/10.1089/MDR.2018.0019

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 abilities to kill chicken embryos, grow in human urine, and colonize the murine kidney. Infect. Immun. 74, 6287–6292. https://doi.org/10.1128/IAI.00363-06/ASSET/1107A68C-A398-4DD0-BB6E-52A326A03724/ASSETS/GRAPHIC/ZII0110662750002.JPEG

Slater, S.L., Sågfors, A.M., Pollard, D.J., Ruano-Gallego, D., Frankel, G., 2018. The type III secretion system of pathogenic escherichia coli. Curr. Top. Microbiol. Immunol. 416, 51–72. https://doi.org/10.1007/82_2018_116/TABLES/1

Smati, M., Clermont, O., Le Gal, F., Schichmanoff, O., Jauréguy, F., Eddi, A., Denamur, E., Picard, B., 2013. Real-time PCR for quantitative analysis of human commensal Escherichia coli populations reveals a high frequency of subdominant phylogroups. Appl. Environ. Microbiol. 79, 5005–5012. https://doi.org/10.1128/AEM.01423-13/SUPPL_FILE/ZAM999104614SO1.PDF

Sorsa, L.J., Dufke, S., Heesemann, J., Schubert, S., 2003. Characterization of an iroBCDEN gene cluster on a transmissible plasmid of uropathogenic Escherichia coli: Evidence for horizontal transfer of a chromosomal virulence factor. Infect. Immun. 71, 3285–3293. https://doi.org/10.1128/IAI.71.6.3285-3293.2003

Stamatakis, A., 2006. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688–2690. https://doi.org/10.1093/bioinformatics/btl446

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, 3412–3419. https://doi.org/10.1128/IAI.00347-10/ASSET/D143C0CA-4259-4FFE-A5D1-120CDAA88CFA/ASSETS/GRAPHIC/ZII9990987270004.JPEG

Torsten Seemann, 2015. Snippy, Rapid haploid variant calling and core genome alignment. GitHub Repos.

Tseytin, I., Dagan, A., Oren, S., Sal-Man, N., 2018. The role of EscD in supporting EscC polymerization in the type III secretion system of enteropathogenic Escherichia coli. Biochim. Biophys. Acta BBA - Biomembr. 1860, 384–395. https://doi.org/10.1016/J.BBAMEM.2017.10.001

Van Hoek, A.H.A.M., Mevius, D., Guerra, B., Mullany, P., Roberts, A.P., Aarts, H.J.M., 2011. Acquired antibiotic resistance genes: An overview. Front. Microbiol. 2, 203. https://doi.org/10.3389/FMICB.2011.00203/BIBTEX

Wang, M., Goh, Y.-X., Tai, C., Wang, H., Deng, Z., Ou, H.-Y., 2022. VRprofile2: detection of antibiotic resistance-associated mobilome in bacterial pathogens. Nucleic Acids Res. 50. https://doi.org/10.1093/nar/gkac321

Wang, S., Liu, X., Xu, X., Yang, D., Wang, D., Han, X., Shi, Y., Tian, M., Ding, C., Peng, D., Yu, S., 2016. Escherichia coli type III secretion system 2 ATPase EivC Is involved in the motility and virulence of avian pathogenic Escherichia coli. Front. Microbiol. 7, 1387. https://doi.org/10.3389/FMICB.2016.01387/BIBTEX

World Health Organization, G.W.H., 2019. WHO | WHO list of Critically Important Antimicrobials (CIA).

Yang, S.C., Lin, C.H., Aljuffali, I.A., Fang, J.Y., 2017. Current pathogenic Escherichia coli foodborne outbreak cases and therapy development. Arch. Microbiol. 2017 1996 199, 811–825. https://doi.org/10.1007/S00203-017-1393-Y

Yu, G., 2020. Using ggtree to Visualize Data on Tree-Like Structures. Curr. Protoc. Bioinforma. 69. https://doi.org/10.1002/cpbi.96

# CHAPTER 4: GENOMIC CHARACTERIZATION OF *CAMPYLOBACTER* SPECIES ISOLATED FROM REUSED PEANUT HULL BROILER LITTER[3]

## 4.1 Summary

*Campylobacter* infections are a leading cause of bacterial diarrhea in humans globally. Infections are due to consumption of contaminated food products and are highly associated with chicken meat, with chickens being an important reservoir for *Campylobacter*. Here, we characterized the genetic diversity of *Campylobacter* species detected in broiler chicken litter over three consecutive flocks and determined their antimicrobial resistance and virulence factor profiles. Antimicrobial susceptibility testing and whole genome sequencing were performed on *Campylobacter jejuni* (n = 39) and *Campylobacter coli* (n = 5) isolates. All *C. jejuni* isolates were susceptible to all antibiotics tested while *C. coli* (n =4) were resistant to only tetracycline and harbored the tetracycline-resistant ribosomal protection protein (TetO). Virulence factors differed within and across grow houses but were explained by the isolates' flock cohort, species and multilocus sequence type. Virulence factors involved in the ability to invade and colonize host tissues and evade host defenses were absent from flock cohort 3 *C. jejuni* isolates as compared to flock 1 and 2 isolates. Our results show that virulence factors and antimicrobial resistance genes differed by the isolates' multilocus sequence type and by the flock cohort they were present in. We hypothesize that the house environment and litter management practices

3 This work has been previously published on a preprint server: Woyda, R., Oladeinde, A., Endale, D., Strickland, T., Lawrence, J.P., Abdo, Z., 2023b. Broiler house environment and litter management practices impose selective pressures on antimicrobial resistance genes and virulence factors of Campylobacter (preprint). Genetics. https://doi.org/10.1101/2023.02.02.526821

performed imposed selective pressures on antimicrobial resistance genes and virulence factors. In particular, the absence of key virulence factors within the final flock cohort 3 isolates suggests litter reuse selected for *Campylobacter* strains that are less likely to colonize the chicken host. Further studies are needed to understand the effect of reused litter on antimicrobial resistance genes, either at the microcosm scale or in controlled, small scale, broiler chicken houses.

## 4.2 Importance

*Campylobacter* is a leading cause of foodborne illness in the United States due to the consumption of contaminated food products or from mishandling of food products, often associated with chicken meat. *Campylobacter* is common in the microbiota of avian and mammalian gut; however, the acquisition of antimicrobial resistance genes and virulence factors may result in strains that pose a significant threat to public health. Although there are studies that have investigated the genetic diversity of *Campylobacter* strains isolated from post-harvest chicken samples, there is limited data on the genome characteristics of isolates recovered from pre-harvest broiler production. In this study, we show that *Campylobacter jejuni* and *Campylobacter coli* that differ in their carriage of antimicrobial resistance and virulence factors may differ in their ability to evade host defense mechanisms and colonize the gut of chickens and humans. Furthermore, we found that differences in virulence factor profiles were explained by the species of *Campylobacter* and its multilocus sequence type.

## 4.3 Introduction

Campylobacteriosis is a leading cause of diarrheal illness worldwide and poultry are the major reservoir of *Campylobacter* species (Young, Davis and DiRita, 2007; De Vries *et al.*, 2018). The

Centers for Disease Control and Prevention (CDC, 2019) estimate that 1.5 million United States residents are affected by campylobacteriosis each year (CDC, 2019). While it rarely results in long-term health problems, studies estimate that 5-20% of campylobacteriosis cases develop irritable bowel syndrome, 1-5% develop arthritis and, in very rare instances, campylobacteriosis may cause Guillain-Barré syndrome (Mishu 1993, Hansson 2016). Economic burden from *Campylobacter* infections was estimated in 2012 to be $1.56 billion and *Campylobact*er, specifically from poultry, was ranked as the leading pathogen-food combination to cause health risks to humans and to negatively impact the economy (Scharff, 2012; Hoffmann, Maculloch and Batz, 2015). Transmission of *Campylobacter* species occurs through consumption or handling of contaminated food products, direct contact with farm or domesticated animals and abattoir workers not practicing good handwashing and food safety practices (Hansson *et al.*, 2018; Igwaran and Okoh, 2019; Mourkas *et al.*, 2020).

*Campylobacter* pathogenicity, disease severity and treatment options are influenced by the repertoire of virulence factors (VFs) and antimicrobial resistance genes (ARGs) they carry. The global rise of antimicrobial resistance (AMR) has impaired effective treatment of *Campylobacter* infections especially when *Campylobacter* strains harbor ARGs that confer resistance to critically important antibiotics (Florez-Cuadrado *et al.*, 2016; Chen *et al.*, 2018; Liu *et al.*, 2019; Zachariah *et al.*, 2021; Liao *et al.*, 2022). *Campylobacter* ARGs may be acquired through mutations (for example the C257T change in *gyrA* (DNA gyrase)) resulting in resistance to fluoroquinolones), encoded on plasmids, or located within multidrug resistance genomic islands (MDRGIs) such as the *erm*(B), which confers high levels of macrolide resistance (Shen *et al.*, 2018).

Importantly, ARGs located on plasmids or MDRGIs are generally transferable across *Campylobacter* species which may lead to the emergence of multidrug resistant strains. Similarly, *Campylobacter* may carry VFs that increase their pathogenicity and the ability to survive within a given host, which can exacerbate disease severity (Lopes *et al.*, 2021). Like ARGs, VFs may be accumulated in Campylobacters leading to strains that are highly virulent and pathogenic (Ghatak *et al.*, 2017; de Fátima Rauber Würfel *et al.*, 2020; Lopes *et al.*, 2021). This may lead to their persistence through pre-harvest and post-harvest, thereby posing a risk to the public (Ghatak *et al.*, 2017; Al Hakeem *et al.*, 2022). Furthermore, strains carrying ARGs conferring resistance to critically important antibiotics for humans, as well as possessing VFs that increase their ability to colonize host tissues, will be harder to treat in the event contaminated food products reach consumers (Montgomery *et al.*, 2018; Liu *et al.*, 2019; Béjaoui *et al.*, 2022).

Bacterial pathogens such as *Campylobacter* have been shown to persist in poultry litter that is reused to grow multiple flocks of broiler chickens (Rauber Würfel *et al.*, 2019). The copraphagic nature of these birds makes the litter one of the first broiler sourced material ingested upon placement in a broiler house. Therefore, it is plausible that broiler chicks get exposed to *Campylobacter* during pecking, bathing, or resting activities. Many studies have characterized *Campylobacter* presence in commonly used bedding materials such as pine shavings, sawdust and ricehulls (Kelley *et al.*, 1995; Pope and Cherry, 2000; Willis, Murray and Talbott, 2000; Stern *et al.*, 2001; Line, 2002, p. 2; Line and Bailey, 2006; Kassem *et al.*, 2010; Rauber Würfel *et al.*, 2019). However, none of these studies performed an in-depth genomic characterization of the *Campylobacter* isolates found.

We previously showed that *C. coli and C. jejuni* were unequally distributed across the litter of four co-located broiler houses on a single farm (Oladeinde *et al.*, 2022). We also showed that the probability of detecting *Campylobacter* in litter was higher for the first broiler flock cohort raised on litter compared to cohorts 2 and 3 (Oladeinde *et al.*, 2022). In the present study, we performed antimicrobial susceptibility testing and whole genome sequencing on forty-four *Campylobacter* isolates recovered from the study (Oladeinde *et al.*, 2022). An in-depth genomic characterization and phylogenetic analysis revealed that isolates clustered based on their VFs. We hypothesize that the litter environment exacts a selective pressure on VFs harbored by *Campylobacter* species however, further studies are needed to verify that the litter environment is responsible for influencing VF flux.

## 4.4 Results

### 4.4.1 *Campylobacter jejuni* and *coli* are present over three flock cohorts

The objective of this study was to understand the changes in ARGs and VFs in *Campylobacter* isolates obtained from litter used to raise multiple flocks of birds. Isolates were obtained from peanut hull litter samples collected from three consecutive flock cohorts within 4 co-located broiler houses. To determine the prevalence of *Campylobacter* in the collected litter samples (Oladeinde *et al.*, 2023), we performed direct and selective enrichment plating of litter eluate onto Cefex agar (Oladeinde *et al.*, 2023). *Campylobacter* was detected in 9.38 % (27/288) of litter samples. Next, we selected  a total of 44 *Campylobacter* isolates for whole genome sequencing (Oladeinde *et al.*, 2023). At least one Campylobacter isolate was selected from each Campylobacter positive litter sample. Additionally, if there were multiple positive isolates obtained from the same litter sample using the different isolation methods (i.e., direct plating,

filter method, or enrichment), the filter and enrichment isolates were chosen over the direct plating isolates. The sequenced *Campylobacter* population consisted of 5 *Campylobacter coli* (CC) and 39 *Campylobacter jejuni* (CJ) isolates as determined by taxonomic classification via Kraken2 (Wood, Lu and Langmead, 2019). *C. coli* was isolated from flocks 1 (n=1) and 2 (n=4) samples but absent from flock 3 (**Table 4.1**), while *C. jejuni* was isolated from each of the 3 flock cohorts. Occurrence of *C. jejuni* across flock cohorts was predominantly in flocks 1 (n=30) and 3 (n=8), while only 1 *C. jejuni* was isolated in flock 2. The 4 co-located broiler houses had unequal representation of both *C. coli* and *C. jejuni* (**Figure 4.1,** (Oladeinde *et al.*, 2023)). Houses 1, 2, 3 and 4 harbored 3, 3, 17 and 16 *C. jejuni* isolates respectively, while houses 1, 3 and 4 harbored 1, 3, and 1 *C. coli* isolates, respectively. Only broiler house 4 harbored *C. jejuni* over 3 consecutive flock cohorts and no house harbored *C. coli* over successive flocks. All *C. jejuni* isolated during flocks 1 and 2 were ST464 while all flock 3 isolates were ST48 (**Table S4.1**). One *C. coli* isolate was identified as ST9450, while the remaining 4 *C. coli* isolates did not match any multilocus sequence type (MLST) profiles (**Table S4.1**) due to poor sequencing coverage. Taken together, *C. jejuni* and *C. coli* were isolated from peanut hull-based litter reused to raise 3 consecutive flocks of birds within 4 co-located broiler houses.

**Table 4.1.** Occurrence of successfully isolated and sequenced *Campylobacter* species from peanut hull litter from 3 consecutive grow-out cycles across 4 co-located broiler houses.

| **Broiler House** | **Flock cohort 1** | | **Flock cohort 2** | | **Flock cohort 3** | |
|---|---|---|---|---|---|---|
| | ***C. coli*** | ***C. jejuni*** | ***C. coli*** | ***C. jejuni*** | ***C. coli*** | ***C. jejuni*** |
| **1** | 0 | 0 | 0 | 0 | 0 | 3 |
| **2** | 1 | 0 | 0 | 0 | 0 | 3 |
| **3** | 0 | 17 | 3 | 0 | 0 | 0 |
| **4** | 0 | 13 | 1 | 1 | 0 | 2 |

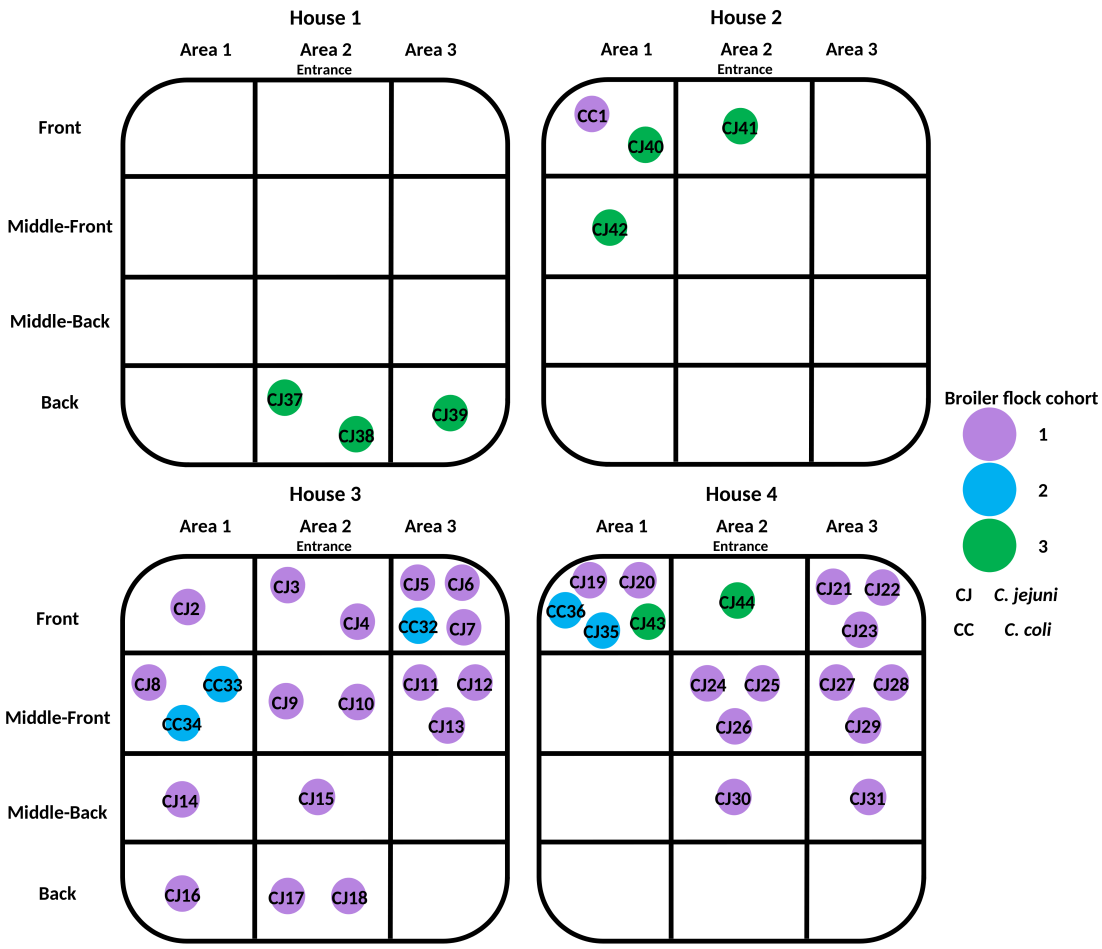



**Figure 4.1.** Visual representation of *Campylobacter* species isolated from peanut hull litter within each section of each broiler house. Samples were taken from 4 co-located broiler houses from 3 consecutive flock cohorts. Circles represent individual isolates labeled by their prospective species: *Campylobacter jejuni* (CJ) and *Campylobacter coli* (CC). Circle color indicates which flock cohort an isolate was obtained from: flock cohort 1 (purple), flock cohort 2 (blue) and flock cohort 3 (green).

### 4.4.2 Virulence factor and antimicrobial resistance profiles differed by species and by multilocus sequence type

We sought to understand the relationship between an isolate's repertoire of ARGs and VFs, its spatial distribution within a given broiler house, and how it is impacted by broiler house environmental factors. Environmental parameters measured included house temperature, litter moisture, and litter pH. Hierarchical clustering based on the presence and absence of all

identified VFs and ARGs (**Table S4.1**) revealed ARG and VF profiles were grouped by flock, *Campylobacter* species and by isolates' multilocus sequence type (**Figure 4.2**). Correspondence analysis revealed overlapping 95% confidence ellipses for isolates by species (**Figure 4.3A**). While hierarchical clustering grouped *C. jejuni* isolates by flock, it is important to note *C. jejuni* isolates from flock 1 encompass a single MLST, ST464, while flock 3 isolates are all ST48. This was recapitulated through correspondence analysis which identified non-overlapping 95% confidence ellipses for isolates by MLST (**Figure 4.3B**). Taken together, these results suggest that the species and sequence type of *Campylobacter* are the main factors explaining the differences observed in ARG and VF profiles.

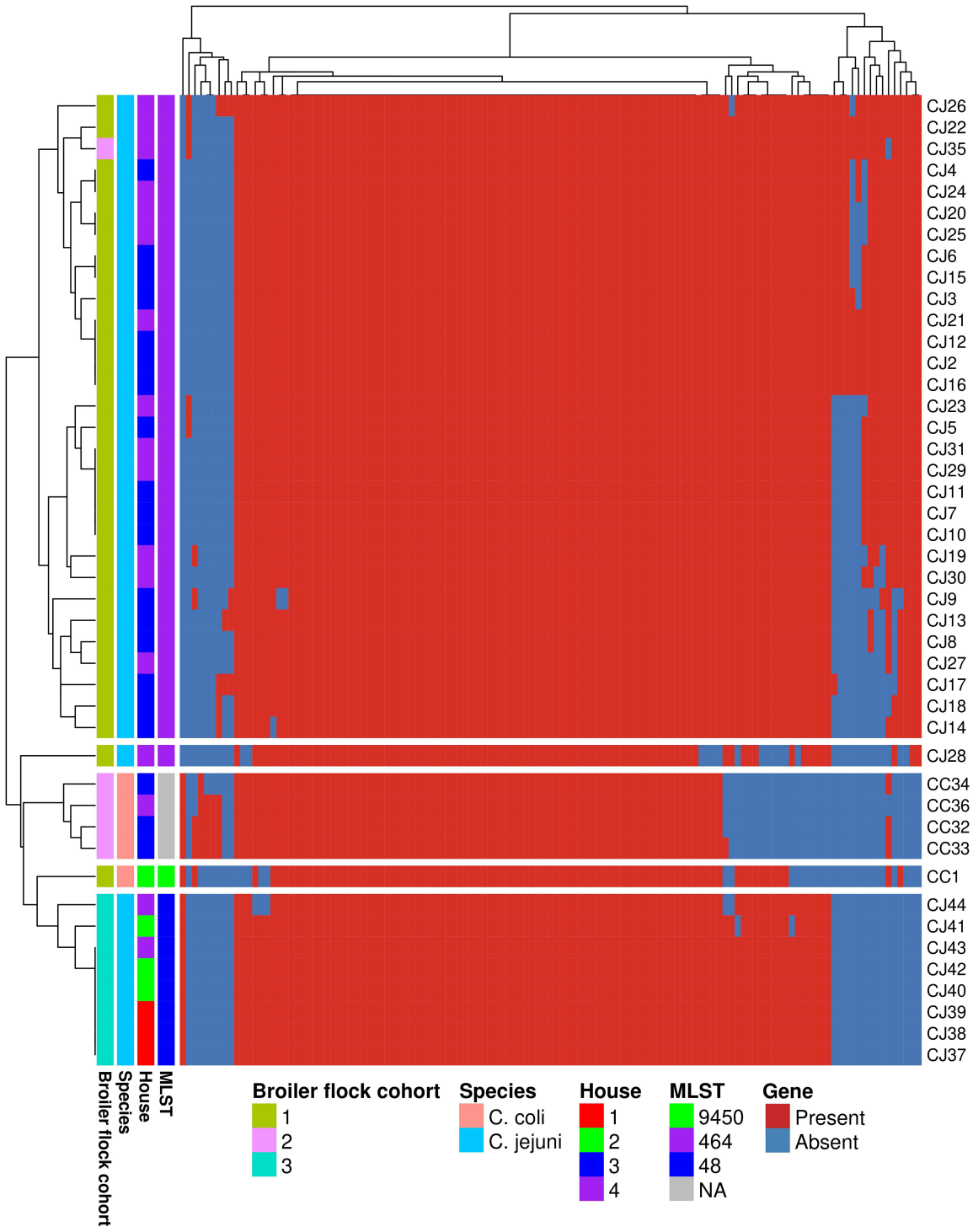

CJ26
CJ22
CJ35
CJ4
CJ24
CJ20
CJ25
CJ6
CJ15
CJ3
CJ21
CJ12
CJ2
CJ16
CJ23
CJ5
CJ31
CJ29
CJ11
CJ7
CJ10
CJ19
CJ30
CJ9
CJ13
CJ8
CJ27
CJ17
CJ18
CJ14
CJ28
CC34
CC36
CC32
CC33
CC1
CJ44
CJ41
CJ43
CJ42
CJ40
CJ39
CJ38
CJ37
Broiler flock cohort
Species
House
MLST
Broiler flock cohort
1
2
3
Species
C. coli
C. jejuni
House
1
2
3
4
MLST
9450
464
48
NA
Gene
Present
Absent

**Figure 4.2.** Heatmap of antimicrobial resistance genes and virulence factors across all isolates. Hierarchical clustering revealed ARG and VF profiles were grouped by flock, *Campylobacter* species and by isolates' multilocus sequence type (MLST). Heatmap was generated in R v4.0.4 with pheatmap v1.0.12 (clustering_method = "average" (UPGMA), clustering_distance_cols = "binary") using the filtered antimicrobial resistance gene and virulence factor table (**Table S4.1**). Columns on the right-hand side display the metadata associated with each isolate (labels on left-hand side). 'House' represents the broiler house for which an isolate was obtained from. 'Broiler flock cohort' corresponds to the flock cohort for which an isolates was obtained from.

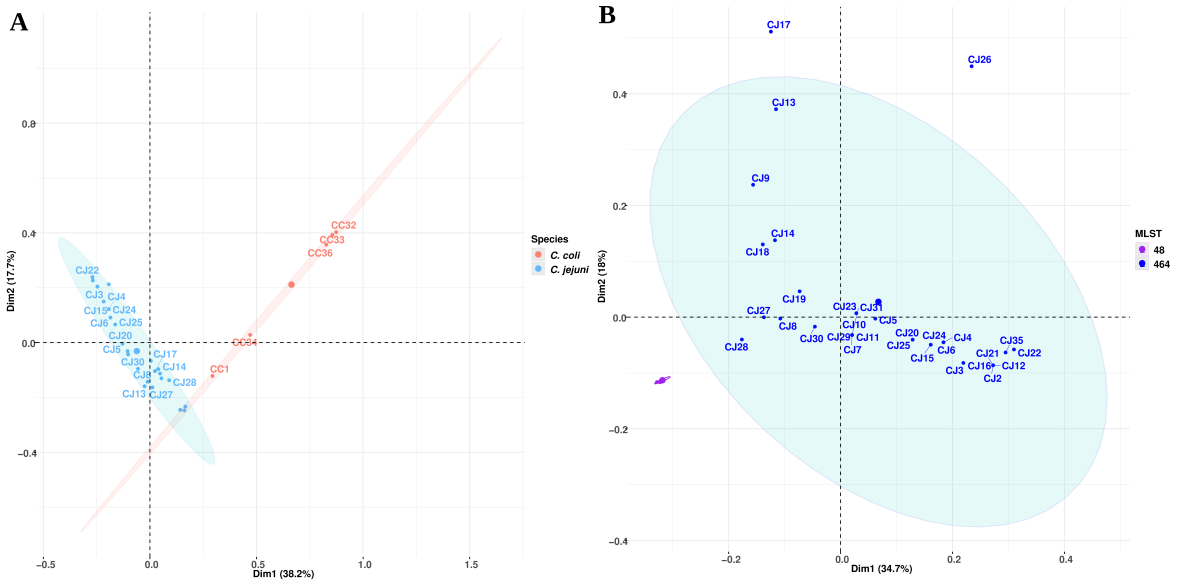



**Figure 4.3.** Correspondence analysis using presence/absence ARG and virulence factor profiles. Correspondence analysis revealed overlapping 95% confidence ellipses for isolates by species and non-overlapping 95% confidence ellipses for isolates by MLST. **(A)** Correspondence analysis on all *Campylobacter* isolates (n=44). Circle color corresponds to the verified species: *C. coli* (red) and *C. jejuni* (blue). Species labels are denoted with CC (*C. coli*) and CJ (*C. jejuni*). **(B)** Correspondence analysis on *Campylobacter jejuni* isolates (n = 39). Circle color corresponds to the identified multilocus sequence type (MLST). Ellipses represent a 95% confidence ellipsoid. Correspondence analysis on the presence/absence table of ARGs and virulence factors was conducted in R using factoextra v1.0.7, FactoMineR v2.4 and corrplot v0.2-0 packages.

### 4.4.3 Virulence factor profiles differed between species and multilocus sequence type

Correspondence and clustering analysis revealed that VF and ARG differences were explained by species and sequence type. Therefore, we sought to further investigate the functional differences of VFs across these parameters. The 112 VFs identified were grouped into 10 VF functional categories (**Table S4.1, Table S4.2**). Sixty-three (56%) of the VFs were present in all isolates (**Table S4.1**). Across species, *C. jejuni* isolates harbored significantly higher average

proportions of VFs than *C. coli* (P< 0.01) with functions relating to toxins, adherence, invasion, colonization and immune evasion, and motility and export apparatus (**Figure 4.4, Table 4.2**). *C. coli* isolate CC1 was the sole *C. coli* isolate that harbored *Cj1415*/*cysC* (cytidine diphosphoramidate kinase) - a toxin-related VF. *Cj1415*/*cysC* is involved in polysaccharide modification and contributes to serum resistance and the invasion of epithelial cells (Taylor and Raushel, 2018). Contrastingly, *Cj1415*/*cysC* was present in all *C. jejuni* isolates except for isolate 28 (CJ28). The *ctdA*, *ctdB* and *ctdC* VFs encoding for the cytolethal-distending toxin and responsible for cellular distension and death in the epithelial cell layer, were present in all *C. jejuni* isolates. The CheVAWY (*cheA*, *cheY, cheV* and *cheW*) system that is involved in adherence, motility, and chemotaxis (Reuter *et al.*, 2021), was also present in all *C. jejuni* isolates. Moreover, studies have shown that gene deletions, or insertional inactivation of *cheY* can result in the attenuation of growth within the chicken gastrointestinal tract (Hendrixson and DiRita, 2004) and the inability to colonize in murine or ferret disease models (Yao, Burr and Guerry, 1997; Bereswill *et al.*, 2011).

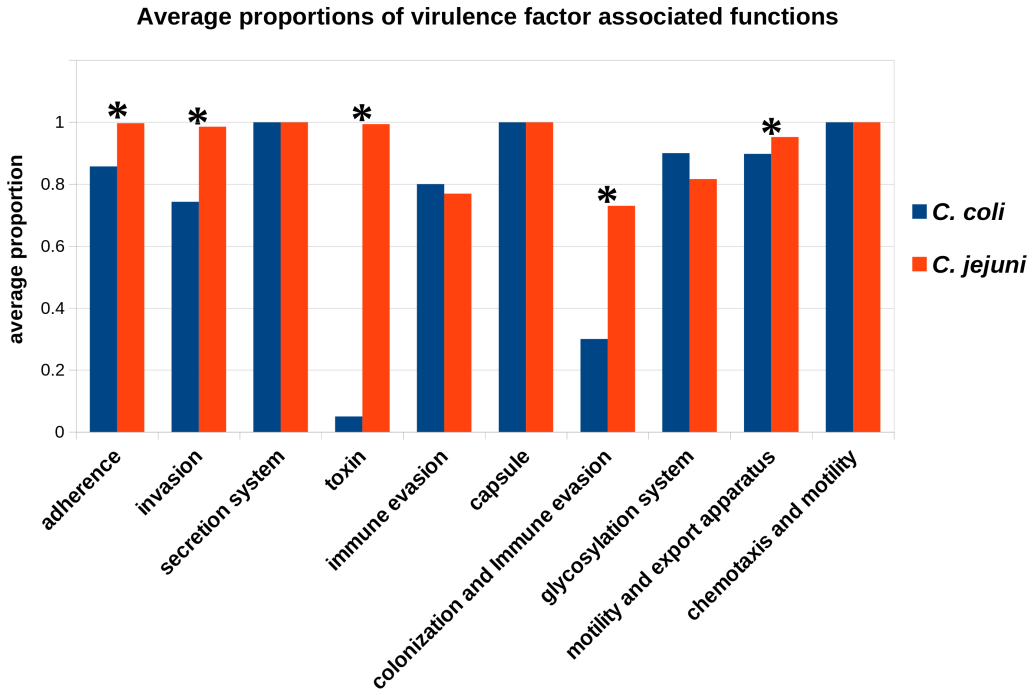



**Figure 4.4.** Average proportion of virulence factor-associated functions between *C. coli* and *C. jejuni* isolates. *C. jejuni* isolates harbored significantly higher average proportions of VFs with functions relating to toxins, adherence, invasion, colonization and immune evasion, and motility and export apparatus. Comparison was performed using the Wilcoxon rank-sum test. (*****) indicates an adjusted p value > 0.05. Virulence genes associated with each function were enumerated for each isolate and a proportion was calculated using the total number of genes in the study population with the given function. Adjusted p value adjustment was performed by the Benjamini-Hochberg false discovery rate correction.

**Table 4.2.** Average proportion of virulence factor-associated functions between *C. coli* and *C. jejuni* isolates. Comparison was performed using the Wilcoxon rank-sum test. Virulence genes associated with each function were enumerated for each isolate and a proportion was calculated using the total number of genes in the study population with the given function. Adjusted p value adjustment was performed using the Benjamini-Hochberg false discovery rate correction method. 'NaN' values are due to the inability to compute p values due to average proportion values being identical for all isolates.

| **Wilcoxon rank-sum test between *C. jejuni* and *C. coli*** | |
|---|---|
| **Virulence Factor Function** | **adjusted p value** |
| Toxin | 1.76E-08 |
| Adherence | 4.16E-08 |
| Invasion | 3.01E-06 |
| Motility and export apparatus | 3.44E-06 |
| Colonization and Immune evasion | 0.0062 |
| Glycosylation system | 0.31 |
| Immune evasion | 1 |
| Capsule | NaN |
| Chemotaxis and motility | NaN |
| Secretion system | NaN |

The capsular polysaccharrides (CPS) of *C. jejuni* are involved in virulence and are essential for survival in certain host environments. The CPS transporter gene (*kpsE*) was present in all *C. jejuni* isolates and in *C. coli* isolate 1 (CC1). *C. jejuni* lipopolysaccharide (LPS) is a known VF that mediates adhesion to epithelial cells while *kpsM* and *kpsT* are involved in LPS export (Karlyshev *et al.*, 2002). Both *kpsM* and *kpsT* were present in all *C. jejuni* and *C. coli* isolates. However, *kpsC* that is responsible for capsule modification, was only present in *C. jejuni* isolates and in a single *C. coli* isolate, CC1.

When comparing the presence of VFs in *C. jejuni* isolates across houses, only VF functions relating to immune evasion and glycosylation system were significantly different ($P<0.01$) (**Table 4.3, Table 4.4**). *C. jejuni* isolates from house 3 harbored more VF functions relating to

immune evasion and glycosylation than *C. jejuni* isolates from houses 1 (**Table 4.3**) and 2 (Table 4). No significant differences in VF functions were found between house 1 and house 2 *C. jejuni* isolates, or between house 4 and any other house (**Table S4.3, S4.4, S4.5, S4.6**). When compared across flocks, significant differences in VF functions relating to immune evasion, glycosylation system, and 'colonization and immune evasion' were found between *C. jejuni* isolates from flock 1 and flock 3 (**Table 4.5**). No significant differences in VF functions were found between flock 1 and 2 *C. jejuni* isolates (**Table S4.7**) albeit, only 1 *C. jejuni* isolate was sequenced from flock 2. Similarly, no significant differences in VF were identified between flock 2 and flock 3 *C. jejuni* isolates (**Table S4.8**). There were 7 VFs (*Cj1426c, fcl, pseE, kfiD, Cj1432c, Cj1440c* and *glf*) that were present in flock 1 (ST464) isolates but absent in *C. jejuni* isolates from flock 3 (ST48) (**Table S4.9**). *Cj1426c, kfiD, Cj1432c, Cj1440c* and *glf* are all involved in capsule biosysnthesis and transport, *fcl* (putative fucose synthase) is involved in LPS biosynthesis, and *pseE* is involved in O-linked flagellar glycosylation. Taken together, *C. jejuni* isolates from flock 3, all identified as ST48, harbored fewer VF than *C. jejuni* isolates from flocks 1 and 2, identified as ST464.

**Table 4.3.** Comparison of average proportions of virulence factor-associated functions for *C. jejuni* isolates which significantly differed in proportion between broiler houses 1 and 3 as determined by the Wilcoxon rank-sum test. Virulence genes associated with each function were enumerated for each isolate and a proportion was calculated using the total number of genes in the study population with the given function. Adjusted p value adjustment was performed by the Benjamini-Hochberg false discovery rate correction method. 'NaN' values are due to the inability to compute p values due to average proportion values being identical for all isolates.

| **Wilcoxon rank-sum test: house 1 vs house 3 (*C. jejuni*)** | |
|---|---|
| **Function** | **adjusted p value** |
| Glycosylation system | 0.0015 |
| Immune evasion | 0.0034 |
| Colonization and Immune evasion | 0.070 |
| Invasion | 1 |
| Motility and export apparatus | 1 |
| Adherence | NaN |
| Capsule | NaN |
| Chemotaxis and motility | NaN |
| Secretion system | NaN |
| Toxin | NaN |

**Table 4.4.** Comparison of average proportions of virulence factor-associated functions for *C. jejuni* isolates which significantly differed in proportion between broiler houses 2 and 3 as determined by the Wilcoxon rank-sum test. Virulence genes associated with each function were enumerated for each isolate and a proportion was calculated using the total number of genes in the study population with the given function. Adjusted p value adjustment was performed by the Benjamini-Hochberg false discovery rate correction method. 'NaN' values are due to the inability to compute p values due to average proportion values being identical for all isolates.

| **Wilcoxon rank-sum test: house 2 vs house 3** | |
|---|---|
| **Function** | **adjusted p value** |
| Glycosylation system | 0.0015 |
| Immune evasion | 0.0034 |
| Colonization and Immune evasion | 0.070 |
| Adherence | 0.25 |
| Invasion | 1 |
| Motility and export apparatus | 1 |
| Capsule | NaN |
| Chemotaxis and motility | NaN |
| Secretion system | NaN |
| Toxin | NaN |

**Table 4.5.** Comparison of average proportions of virulence factor-associated functions for *C. jejuni* isolates which significantly differed in proportion between flock cohorts 1 and 3 as determined by the Wilcoxon rank-sum test. Virulence genes associated with each function were enumerated for each isolate and a proportion was calculated using the total number of genes in the study population with the given function. Adjusted p value adjustment was performed by the Benjamini-Hochberg false discovery rate correction method. 'NaN' values are due to the inability to compute p values due to average proportion values being identical for all isolates.

| **Wilcoxon by flock 1 vs flock 3 C. jejuni (house 1, 2, 3 and 4)** | |
|---|---|
| **Function** | **adjusted p value** |
| Immune evasion | 1.46E-06 |
| Glycosylation system | 3.91E-06 |
| Colonization and Immune evasion | 0.00056 |
| Adherence | 0.61 |
| Invasion | 1 |
| Motility and export apparatus | 1 |
| Toxin | 1 |
| Secretion system | NaN |
| Capsule | NaN |
| Chemotaxis and motility | NaN |

### 4.4.4 The presence of Type IV and type VI secretion systems differentiated *Campylobacter* species

Both type IV and type VI secretion systems (T4SS and T6SS) enable delivery of bacterial effector proteins into neighboring bacterial and eukaryotic cells (Bleumink-Pluym 2013) and are commonly present in *C. jejuni* and *C. coli* (Bleumink-Pluym 2013, Daya Marasini 2020). Although genes encoding these secretion systems may be chromosomally encoded, they are commonly identified on plasmids (Cascales, 2008; Lertpiriyapong *et al.*, 2012; Bleumink-Pluym *et al.*, 2013; Ghatak *et al.*, 2017). Using the *C. jejuni* strain WP2202 plasmid pCJDM202 (NZ_CP014743) that has both the T4SS and T6SS operons as a reference, we determined that

T4SS was present only in *C. coli* isolates and T6SS was present only in *C. jejuni* isolates. *C. coli* isolates CC32, CC34 and CC36 harbored the largest segments of the reference query (>88%) while CC33 only contained 33% of the reference query (**Table S4.10**). Sixteen *C. jejuni* isolates (41%) harbored >85% of the T6SS operon with >99% pairwise identity (**Table S4.11**). Eight *C. jejuni* isolates contained between 14% and 76% of the T6SS operon, with >98% pairwise identity, while the remaining 15 isolates harbored <10% of the operon. No isolates from the final flock cohort, flock 3, harbored any T6SS genes. In summary, T4SS and T6SS were *Campylobacter* species-specific and were only found in isolates recovered from the litter of flocks 1 and 2. As these secretion systems aid in colonization in chickens, as well as humans, this data suggests isolates that lack these systems may have a decreased ability to infect chickens in subsequent flocks.

### 4.4.5 ARG carriage differs by species and flock

Antimicrobial susceptibility testing (AST) was performed on all *Campylobacter* isolates (**Table S4.12**). All *C. jejuni* isolates were susceptible to all drugs tested, while *C. coli* isolates (4/5) were resistant to tetracycline. Eleven ARGs were identified across all isolates after gene annotation (**Table S4.13**). All *C. coli* isolates with phenotypic resistance to tetracycline harbored *tetO* (tetracycline resistance gene). *tetO* is known to confer tetracycline resistance in *C. coli* (Sougakoff *et al.*, 1987). To determine if *tetO* was harbored on a plasmid, we used *C. jejuni* strain WP2202 plasmid pCJDM202 (NZ_CP014743) as a reference for a BLAST search. We identified regions on the WP2202 plasmid that had high percent identity (99.5%) to *C. coli* contigs containing *tetO* from this study. This result suggests that *tetO* is either located on a

plasmid or has been integrated into the chromosome, even though no plasmid replicons were identified.

The class D beta-lactamase structural gene, *bla*$_{\text{OXA-61}}$, that confers resistance to penams, cephalosporins and carbapenems (Alfredson and Korolik, 2005) was found in all 5 *C. coli* isolates and in 9 *C. jejuni* isolates. The *C. coli* isolates were isolated from house 2 of flock 1 and houses 3 and 4 of flock 2, whereas *C. jejuni* isolates carrying *bla*$_{\text{OXA-61}}$ were from houses 1, 2 and 4 of flock 3 **(Table S4.1**). *bla*$_{\text{OXA-61}}$ has been previously identified in both *C. jejuni* and *C. coli* isolates obtained from poultry production as well as humans (Griggs *et al.*, 2009; De Vries *et al.*, 2018; Gharbi *et al.*, 2021). Truncation of the upstream sequence of *bla*$_{\text{OXA-61}}$ -35 region has been reported by Alfredson et al (2005) to result in wild-type *C. jejuni* beta-lactam-susceptibility (Alfredson and Korolik, 2005). The upstream conserved sequence of *bla*$_{\text{OXA-61}}$ of *C. coli* isolates CC32, CC33 and CC36 was 100% identical to the upstream sequence of *bla*$_{\text{OXA-61}}$ harbored on a recombinant plasmid (NCBI: pGU0401) (Alfredson and Korolik, 2005). The other two *C. coli* isolates (CC1 and CC34) had a single T to G mutation at base 66 and CC1 had an additional T to C mutation at base 23. Neither of these mutations were located within the ribosomal binding site or the -10 or -35 promoter regions. We do not know if the presence of *bla*$_{\text{OXA-61}}$ conferred the expected resistance phenotype in these isolates since the NARMS *Campylobacter* AST panels used did not include antibiotics classified as penams, cephalosporins or carbapenems. Furthermore, resistance-nodulation-cell division-type multidrug efflux pumps (CmeABC and CmeDEF operons) that confer resistance to antimicrobials and toxic compounds (Akiba *et al.*, 2006) were found in *Campylobacter* isolates. CmeABC was present in all *C. jejuni* isolates while CmeDEF was found in both *C. coli* and *C. jejuni* isolates. CmeR is a known transcriptional regulator of CmeABC and when absent or mutated leads to the overexpression of the CmeABC

efflux pump and increases levels of resistance to several antimicrobials (Lin *et al.*, 2005). While the *cmeR* gene was present in all *C. jejuni* isolates, it was absent from *C. coli* isolates CC32, CC33, CC34 and CC36. In addition to ARGs, we found genes that contribute to arsenic resistance within *C. jejuni* isolates. The arsenical-resistance membrane transporter *acr3,* the putative membrane permease *arsP* and the arsenical pump membrane protein *arsB* were found in all *C. jejuni* isolates from flocks 1 and 2 but absent from flock 3 isolates.

In summary, AST revealed that phenotypic resistance to antibiotics was only present for tetracycline in *C. coli* isolates. Genomic characterization of ARGs discovered the presence of a beta-lactamase gene in flock 2 and 3 isolates that was absent from flock 1 isolates. In addition, two multidrug efflux pumps, one of which was only present in *C. coli* isolates were identified. Therefore, the distribution of ARGs within the *Campylobacter* isolates differed by flock cohort and by *Campylobacter* species.

### 4.4.6 Core genome analysis reveal limited genetic diversity among *C. jejuni* isolates

The following *C. jejuni* isolates were found to be identical based on the alignment of their core genome: CJ5, CJ7 and CJ10; CJ4, CJ11, CJ12 and CJ16; CJ14 and CJ23; CJ19 and 28; CJ6, CJ20, CJ21, CJ22 and CJ24; and CJ39, CJ40, CJ41, CJ42 and CJ43. Each set of identical isolates were from the same flock and of the same MLST but not from the same house (**Figure 4.1**). The high number of identical core genomes suggest that there is limited genetic diversity within the core genomes (genes present in >=95% of isolates) of isolates of the same species or MLST. The core genome for *C. jejuni* isolates consisted of 1116 genes and the core genome for *C. coli* was 1561 genes (**Table S4.14)**.

Using the Roary-generated core genome alignment, we estimated a maximum likelihood phylogenetic tree using RAxML under a GTR substitution matrix. The resulting phylogeny clustered isolates of the same species or MLST into separate clades. For instance, separate clades were identified for *C. coli* isolates, ST48 *C. jejuni* isolates from flock 3 and *C. jejuni* ST464 isolates from flocks 1 and 2 (**Figure 4.5**).

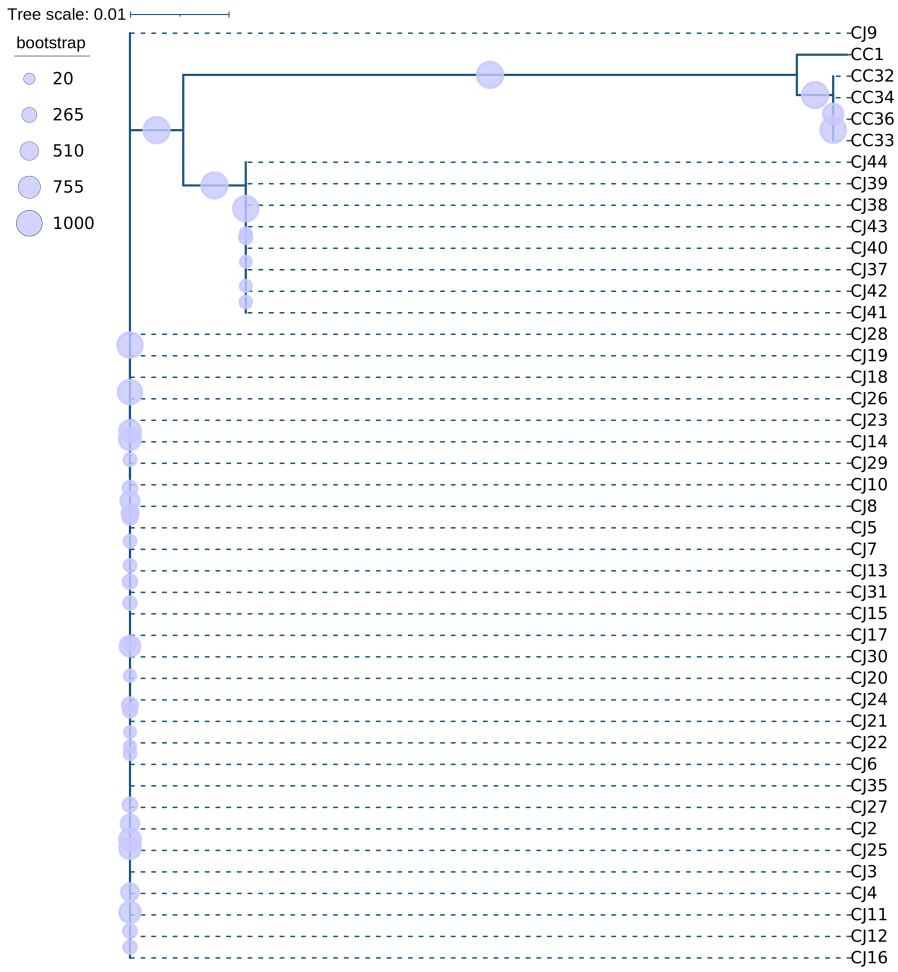
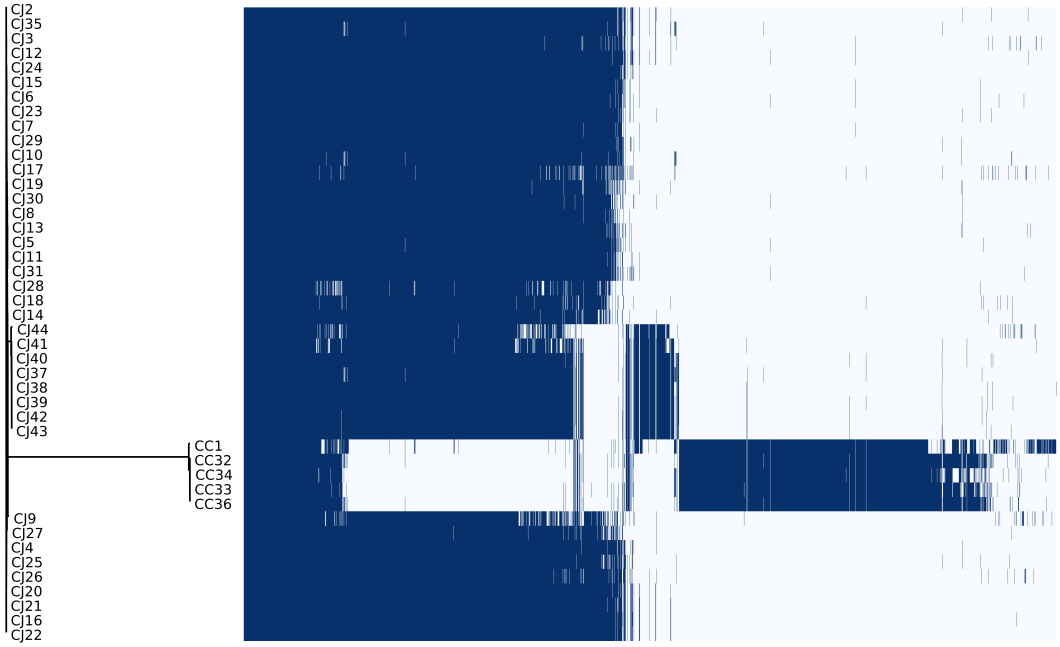



**Figure 4.5.** RAxML maximum likelihood phylogenetic tree estimated from the Roary core genome alignment under a GTR model of evolution. 1000 bootstraps were performed to ensure nodal support. Tree was visualized using the Interactive Tree of Life (iTOL).

*C. coli* isolates had a distinct accessory genome (genes present in <95% of the isolates) profile as seen by the Roary-produced accessory genome tree which was plotted alongside the gene

presence/absence information **(Figure 4.6, top).** This consisted of 1,223 genes including 557 associated with hypothetical proteins and 666 annotated genes. Additionally, isolates recovered from flock 3 carried a set of accessory genes lowly present in flock 1 and 2 isolates. This collection of 188 accessory genes consists of 147 associated hypothetical proteins and 41 genes with annotations. Although many isolates shared an identical core genome, all isolates were dissimilar from one another based on their accessory genome. Correspondence analysis based on the presence/absence of accessory genes, VFs and ARGs indicated a separation between flock 3 (ST48) isolates and flock 1 and 2 isolates (ST464) (**Figure 4.6, bottom**). Thus, both the accessory genome and core genomes grouped *Campylobacter* isolates based on their species and multilocus sequence type.


CJ2
CJ35
CJ3
CJ12
CJ24
CJ15
CJ6
CJ23
CJ7
CJ29
CJ10
CJ17
CJ19
CJ30
CJ8
CJ13
CJ5
CJ11
CJ31
CJ28
CJ18
CJ14
CJ44
CJ41
CJ40
CJ37
CJ38
CJ39
CJ42
CJ43
CC1
CC32
CC34
CC33
CC36
CJ9
CJ27
CJ4
CJ25
CJ26
CJ20
CJ21
CJ16
CJ22

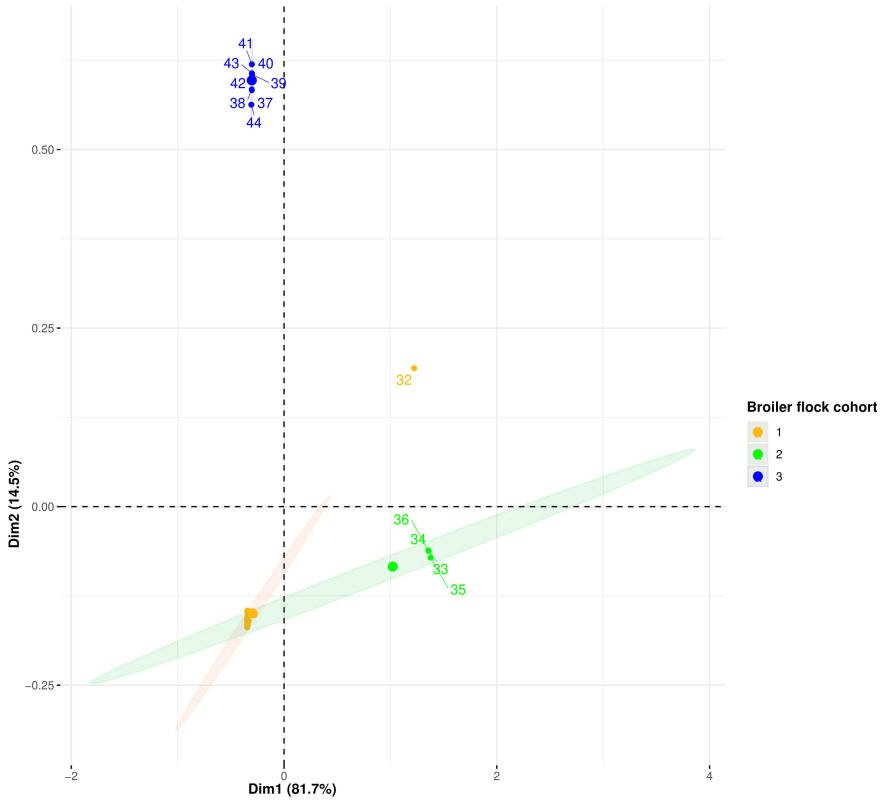

41
43
40
42
39
38
37
44
32
36
34
33
35
0.50
0.25
0.00
−0.25
−2
0
2
4
Dim1 (81.7%)
Dim2 (14.5%)
Broiler flock cohort
1
2
3

**Figure 4.6.** Core and accessory genome analysis. Species labels are denoted with CC (*C. coli*) and CJ (*C. jejuni*). *C. coli* isolates, *C. jejuni* ST48 isolates and *C. jejuni* ST464 isolates harbor distinct accessory genomes. **(Top)** Gene presence/absence matrix of core and accessory genes. Rows are labeled with each corresponding *Campylobacter* isolate and columns represent genes present (blue) or absent (white) within each isolates core and accessory genome. The matrix was produced using roary_plots.py and the Roary-generated gene_presence_absence.csv (**Table S4.14**) and accessory_binary_genes.fa.newick files. **(Bottom)** Roary, ARG, virulence factor presence/absence correspondence analysis. Correspondence analysis was performed on the combination of the Roary-generated gene_presence_absence.csv and the presence/absence table of ARGs and VFs. Correspondence analysis was conducted in R using factoextra v1.0.7, FactoMineR v2.4 and corrplot v0.2-0 packages.

## 4.5 Discussion

The purpose of this study was to characterize the ARGs and VFs of *Campylobacter* isolates recovered from litter during 3 consecutive flock cohorts of broiler chickens from 4 co-located broiler houses. Our objective was to identify the ARGs and VFs harbored by these isolates as well as understand how management and environmental factors can lead to genomic changes over the course of multiple flocks.

We found VFs and VF functions that significantly differed across species. VFs relating to adhesion, immune evasion and toxin production differed between *C. jejuni* and *C. coli* (**Table S4.2**). In general, *C. coli* isolates harbored fewer VF relating to toxin production, adherence, invasion, motility, colonization, and immune evasion than *C. jejuni* (**Figure 4.4, Table 4.2**). This higher number of VFs in *C. jejuni* may explain why they are more widespread than *C. coli* in broiler production (Powell *et al.*, 2012; Whitehouse *et al.*, 2018; Tang *et al.*, 2020). We also observed differences between *C. jejuni* and *C. coli* in their carriage of T4SS and T6SS. The T4SS aids in the invasion of epithelial cells and has been shown to support intraspecies and interspecies conjugative DNA transfer in *Campylobacter fetus* (Kienesberger *et al.*, 2011; Gokulan *et al.*, 2013; Van Der Graaf-Van Bloois *et al.*, 2016). Similarly, the T6SS is an important

VF for *C. jejuni* and it is involved in cell adhesion, cytotoxicity, and invasion *(Lertpiriyapong et al.*, 2012). *C. jejuni* isolates carrying a T6SS have been identified in poultry and human clinical settings (Bleumink-Pluym *et al.*, 2013; Ghatak *et al.*, 2017; Kanwal *et al.*, 2019; Marasini *et al.*, 2020). Kanwal et al. (2019) determined that *C. jejuni* possessing *hcp*, a T6SS gene and important effector protein, had higher hemolytic activity and higher competitive growth advantage against *Helicobacter pullorum*, a bacterium which inhabits a similar physiological niche in chickens (Kanwal *et al.*, 2019). No T6SS genes were identified in isolates from the final flock cohort, flock 3, and suggests these genes may impose a fitness cost when in a litter environment. As both the T4SS and T6SS are important virulence factors for the colonization of both chickens and humans, our data suggest that Campylobacter isolates in litter that lack secretion systems will be less likely to infect chickens and therefore less likely to enter into the production facility and consumer-borne food products.

We also found that *C. coli* and *C. jejuni* differed in their susceptibility to antibiotics. *C. coli* isolates (4/5) were resistant to tetracycline while all *C. jejuni* isolates were susceptible to all drugs tested. We identified ARGs and metal resistance genes encoding tetracycline resistance (*tetO)*, arsenic resistance (*arsP* and *acr3)*, multidrug efflux pumps (CmeABC and CmeDEF operons) and class D beta-lactamase structural gene ($bla_{\text{OXA-61}}$). Tetracycline is both approved for use in food-producing animals and classified as medically important antimicrobial (Center for Veterinary Medicine, 2022). Tetracycline accounts for the largest volume of sales for antimicrobials in food-producing animals and second highest used antibiotic in poultry (Center for Veterinary Medicine, 2022). Thus, it is not surprising we observed tetracycline resistant genes within our isolates. NARMS reporting for Campylobacters does not include beta-lactam/beta-lactamase inhibitor combination agents (Center for Veterinary Medicine, 2022) however,

resistance to beta-lactams has been identified in both humans and chickens (Lachance *et al.*, 1991; Thwaites and Frost, 1999). For example, $bla_{\text{OXA-61}}$ has been identified in both human and chicken isolates (Griggs *et al.*, 2009; Casagrande Proietti *et al.*, 2020).

Species differences were also observed based on the presence of the transcriptional repressor gene for the MDR pump CmeABC, *cmeR*. Mutations, or absence thereof, of the transcriptional repressor, CmeR can lead to enhanced production of the MDR pumps. CmeABC overexpression can lead to reduced susceptibility to tetracycline, ampicillin, cefotaxime, erythromycin, and fusidic acid in *Campylobacter jejuni* (Lin *et al.*, 2005) and therefore could explain resistance seen in isolates lacking CmeR. The CmeABC MDR pump and the corresponding regulator, *cmeR*, were present in all *C. jejuni* isolates. *C. coli* isolates that harbored *tetO* harbored CmeABC but not the *cmeR* regulator. The absence of the *cmeR* regulator could have contributed to the level of tetracycline resistance observed in *C. coli* isolates. The presence of this MDR pump, along with its regulator *cmeR*, in the *C. jejuni* isolates suggests that it does not confer resistance, above the epidemiological cutoff value, to the antibiotics tested in our AST panel as all *C. jejuni* were pan susceptible. Overall, *Campylobacter* isolates from this study pose no significant ARG threat and this observation may be attributed to the management program enacted by the producer. Here, the farmer adopted a “No Antibiotics Ever” program after a complete cleanout of the houses were done (Oladeinde *et al.*, 2022).

Our previous results (Oladeinde *et al.*, 2022) indicated *Campylobacter* was most prevalent during the grow-out of the first flock cohort compared to flock 3 (Oladeinde *et al.*, 2022). We observed flock differences with respect to VFs and ARGs. We determined that isolates recovered from the same flock cohort had similar VFs and VF associated functions. Isolates from flock 1 were raised on fresh peanut hull while flocks 2 and 3 were raised on the reused litter from flocks

1 and 2, respectively. We observed a significant difference in VFs in flock 3 (ST48) isolates compared to flock 1 isolates (ST464). We hypothesize that as the litter was reused over multiple flock cohorts the litter microbiome underwent significant flux. The VFs lost over the multiple flock cycles may have imposed a fitness cost resulting in ST464 isolates not being detected past the second flock cohort and ST48 isolates being detected in the final flock cohort, flock 3. The VFs which were absent from flock 3 (ST48) isolates are associated with survival within the chicken gut: capsule biosynthesis and motile phenotype and may not be essential for survival within the peanut hull litter. For example, the *glf* gene, encoding UDP-galactopyranose mutase, is involved in capsule polysaccharide biosynthesis (Poulin *et al.*, 2010) and is a known determinant for invasion, serum resistance, adherence, colonization and modulation of host immune responses (Rojas *et al.*, 2019). Alternatively, the introduction of ST48 *C. jejuni* isolates into the final flock cohort could have occurred through other means: 1) human or rodent transmission of these strains into the broiler houses, 2) isolation of present isolates was unsuccessful in the previous flock cohorts, or 3) incoming chicks harbored new strains which were then isolated.

We also found that *Campylobacter* prevalence differed between houses. For instance, there was a higher probability of detecting *Campylobacter* in houses 3 and 4 compared to houses 1 and 2 (Oladeinde *et al.*, 2022). Observed house differences are unlikely due to differences in day-old chicks because all chicks originated from the same hatchery and were randomly assigned to houses. VFs significantly differed across grow-out houses (**Table 4.3, Table 4.4**) and this suggests the house environment played an integral role in selecting for these strains. It is possible that these strains were introduced from the hatchery, as a new ST was detected following the introduction of chicks in the final, flock 3, cohort. It is also plausible these strains were residual

contamination from the previous flocks and the cleanup performed was not sufficient to remove them. Consequently, upon placement of new chicks, these strains were able to efficiently colonize the naive gastrointestinal tract and spread through houses 3 and 4. Recently, Yi Fan et al. (2022) showed that cleaning broiler houses with water increased activity of the gut microbiota and reduced *Campylobacter* transmission relative to a full disinfection (Fan *et al.*, 2022). Therefore, it is possible that the cleaning procedure used had differential effects on the resident bacterial population in each house. For instance, strains from houses 3 and 4 carried VFs that also play a role in organized biofilm formation (*cheA, cheY, cheV and cheW*), which may allow them to adhere to surfaces and persist through cleaning. Additionally, isolates from house 3, which were all from the first flock cohort, harbored a higher proportion of VFs with functions related to immune evasion, glycosylation system and colonization and immune evasion. Flagellin glycosylation has been shown to affect the adherence and invasion of human epithelial cells (Guerry *et al.*, 2006). We hypothesize that the altered litter environment, a result from its reuse in raising multiple cohorts of broilers, is imposing a selective pressure on these VFs and that isolates harboring these VFs have a decreased ability to persist over multiple flock cohorts.

We have provided new data on the genome characteristics of *C. jejuni* and *C. coli* isolates recovered from the litter of broiler chickens. We demonstrated that the presence of VFs and ARGs varied by species and by flock. While significantly more VFs were present in isolates from house 3, these isolates were not detected in the final flock (flock 3). Additionally, isolates that were found in the litter of flock 3 were missing several VFs that increase an isolate's ability to colonize and survive within the chicken host including VFs for capsule biosynthesis, motility and the T6SS. The presence of isolates harboring these genes within the litter of the first and second cohorts suggests these isolates may have been deposited through fecal excretion and

reingested through the coprophagic nature of the birds. However, isolates which are retained within the litter and are not reingested are at a disadvantage within the litter environment due to the retention of these colonization-related virulence factors and therefore are removed due to negative selection. We hypothesize this is why isolates from the final flock cohort are devoid of these virulence factors and were able to survive and be successfully isolated from the litter. These data suggest that the house environment and management practices including the initial house cleaning procedure and the reuse of the peanut hull litter over multiple flocks may be capable of imposing selective pressures on VFs. Therefore, more studies are warranted to understand the characteristics of reused litter, both microbial and physiochemical, which may confer selective pressures against *Campylobacter* harboring VFs relating to host colonization and survival within a host. Nonetheless, there are several limitations of the study which could have biased our interpretation of the results including the small number of flocks, unknown broiler house conditions, as well as the limit of detection of our sampling methodology for *Campylobacter* isolation. Lastly, results from this study are based on one farm and may not be representative of all farms which reuse peanut hull-based litter for broiler chicken production.

## 4.6 Materials and methods

Details of methods used for sampling on farms, litter management and bacterial isolation have been described before (Oladeinde *et al.,* 2022). We briefly re-describe some of these methods and present others below.

### 4.6.1 Study design

Four broiler houses on a farm in Central Georgia, each containing 22,000 to 24,000 broilers per flock, were selected for this study. Three cohorts of broiler flocks were raised in succession in

each of the 4 broiler houses between February and August 2018. Before the start of the study a complete litter cleanout was performed in each of the four houses. Before the first broiler flock was introduced fresh peanut hull litter was prepared in each house. Each successive flock, after the first, was raised on the previous flock cohorts' litter without any cleanout between flock cohorts. During the downtime between flocks the litter was mechanically conditioned by removing the caked portions. Additionally, during the downtime the litter was treated for ammonium control (typically 1 week before sampling) via topical application of a commercial litter acidifier. For the first 14 days of each flock cohort half house brooding was practiced; chicks were only allowed to occupy the front section of the broiler house until after 14 days. Copper sulfate was added to drinking water. All management procedures used are within the scope for routine industry practices.

### 4.6.2 Litter sample collection

From February 2018 to August 2018, a total of 288 poultry litter (PL) samples were collected from 4 co-located broiler houses throughout the study. This represents 96 PL sample collections from poultry houses per cohort of broiler flock raised on the same litter. For each broiler cohort, PL samples were collected both early (< 14 days) and late (days 32 - 38) during the grow-out phase at three different sampling times. During each sampling time from each of the four poultry houses, PL samples were collected from four sections: front, mid-front, mid-back and back. From each section, a pool of three litter grabs were collected, bagged litter was transported in a cooler with icepacks until arrival at the laboratory. Litter moisture content was determined for each litter sample by initially weighing 1 g, drying at ~106°C overnight, and re-weighing to measure dry weight. Moisture content was determined by the difference. Litter pH was obtained

by mixing litter (10 g) with 20 ml water, immersing pH probe into mixture, and recording the reading. Poultry house temperature was also collected during each sampling time.

### 4.6.3 Bacterial isolation and identification

For *Campylobacter* species detection, appropriate dilutions of the litter mixture were direct plated to Cefex agar (Remel, Lenexa, KS). Plates were incubated in a microaerobic, hydrogen enriched atmosphere (7.5 % H2, 2.5 % O2, 10 % CO2, and 80 % N2) at 42°C for 48 h. Additionally, aliquots of the litter mixture (4 x 50 ul drops) were placed onto a 0.65 µm cellulose acetate filter placed on Cefex agar. Filters were allowed to dry 30 min before being removed and plates were incubated as above. Enrichment was also performed by adding 1 ml of litter mixture to 9 ml bolton's broth and incubated in a microaerobic atmosphere at 42°C for 48 h before being transferred to Cefex agar and incubated as above. Presumptive positive colonies were selected based on typical cellular morphology and motility using phase contrast microscopy. Isolates were confirmed using the *Campylobacter* BAX® real-time PCR Assay (Hygiena; Wilmington, DE) according to manufacturer's directions. Twenty-seven unique litter samples were positive for *Campylobacter* and a total of 53 *Campylobacter* isolates were obtained following the different isolation methods described above. For whole genome sequencing, at least one *Campylobacter* positive isolate was selected from the 27 litter samples. Additionally, if there were multiple positive isolates obtained from the same litter sample using the different isolation methods (i.e., direct plating, filter method, or enrichment), the filter and enrichment isolate were chosen over the direct plating isolate. A total of 44 *Campylobacter* positive samples were chosen for whole genome sequencing.

### 4.6.4 Antibiotic susceptibility testing

We performed antimicrobial susceptibility testing (AST) on 5 *Campylobacter coli* and 39 *Campylobacter jejuni* isolates recovered from the litter of broiler chicks following the National Antimicrobial Resistance Monitoring System (NARMS) protocol for Gram-negative bacteria. The following antimicrobials were tested: Azithromycin, Ciprofloxacin, Clindamycin, Erythromycin, Florfenicol, Gentamicin, Nalidixic Acid, Telithromycin and Tetracycline. Antimicrobial susceptibility of *Campylobacter* isolates was determined using the Sensititre semi-automated system (Thermo Fisher Scientific, Kansas City, KS) according to manufacturer's instructions. Briefly, bacterial suspensions equivalent to a 0.5 McFarland suspension were prepared, aliquoted into a CAMPY panel and incubated at 42°C for 24 h under microaerobic conditions. Minimum inhibitory concentrations were determined and categorized as resistant according to breakpoints based on epidemiological cut-off values as used by the National Antimicrobial Resistance Monitoring System (NARMS; https://www.fda.gov/media/108180/download).

### 4.6.5 Whole genome sequencing and processing

Illumina short read sequencing was performed on DNA extracted from *Campylobacter* isolates recovered from litter. Libraries were prepared using Nextera XT DNA library preparation kits (Illumina, Inc., San Diego, CA) following the manufacturers protocol. Libraries were sequenced on the MiSeq platform with 250-bp paired end reads. Genome assembly, antimicrobial resistance gene identification, virulence factor identification, plasmid replicon identification, phage region identification and genome annotation were done using Reads2Resistome pipeline v0.0.2 (Woyda, Oladeinde and Abdo, 2023). Online ResFinder (Bortolaia *et al.*, 2020) for annotation of acquired

resistance genes and additional resistance gene identification was performed using the Resistance Gene Identifier v5.1.1 (RIG) and antibacterial biocide and metal resistance genes were identified with BacMet (Pal *et al.*, 2014). MLST was determined using the mlst software (Jolley and Maiden, 2010; Seemann, no date), which utilizes the PubMLST website (https://pubmlst.org/). It was required for all reference database hits needed to meet a minimum identity match of 85%. Reference Genbank plasmids AF301164.1, CP044166.1, CP020775.1 and CP014743 were used to determine the exact locations of type 4 and type 6 secretion systems. Verification of genes using Megablast was performed in Geneious Prime version 2022.2.2.

### 4.6.7 Statistical analysis

ARGs and virulence database hits from R2R, ResFinder, BacMet and RGI were filtered utilizing >=85% sequence identity to the reference database query and a table was generated based on the presence/absence of identified genes from all isolates. Correspondence analysis on the presence/absence table of ARGs and virulence factors was done in R using factoextra v1.0.7 (Alboukadel Kassambara, Fabian Mundt, 2020), FactoMineR v2.4 (Lê, Josse and Husson, 2008) and corrplot v0.2-0 (Wei, Taiyun and Simko, Viliam, 2021) packages. Heatmaps were generated using the pheatmap (Raivo Kolde, 2019) package in R. A distance matrix was generated using the jaccard metric via the vegdist() function from the vegan v2.6-4 package (Jari Oksanen *et al.*, 2022). Optimal number of clusters was identified using the silhouette method implemented by the fviz_nbclust function from the factoextra v1.0.7 package. hclust() from the stats v3.6.2 (R Team, 2015) package was then utilized to perform hierarchical clustering under the ‘average’ (UPGMA) method using the determined optimal number of clusters. All analyses were done in R v4.0.4 (R Team, 2015) utilizing RStudio v1.2.1106 (RStudio Team, 2016).

### 4.6.8 Taxonomic classification and phylogenetic analysis

Taxonomic classification of *Campylobcater* isolates using the quality controlled Ilumina short-read sequences was performed with Kraken2 (Wood, Lu and Langmead, 2019, p. 2). Pangenome analysis of annotated assemblies was performed with Roary (Page *et al.*, 2015). A phylogenetic tree of the core genome alignment from Roary (core_gene_alighment.aln) was constructed using RAxML with the maximum likelihood method under a GTR model with 1000 bootstraps (Stamatakis, 2006, 2014). The following command was used for phylogentic tree estimation using RAxML: “raxmlHPC -m GTR -p 12345 -s core_gene_alignment.aln -f a -x 12345 -N 1000 -T 48”. Tree visualization was conducted using the Interactive Tree of Life (Letunic and Bork, 2021).

## 4.10 REFERENCES

Akiba, M. *et al.* (2006) ‘Interaction of CmeABC and CmeDEF in conferring antimicrobial resistance and maintaining cell viability in Campylobacter jejuni’, *Journal of Antimicrobial Chemotherapy*, 57(1), pp. 52–60. Available at: https://doi.org/10.1093/JAC/DKI419.

Al Hakeem, W.G. *et al.* (2022) ‘Campylobacter jejuni in Poultry: Pathogenesis and Control Strategies’, *Microorganisms*, 10(11), p. 2134. Available at: https://doi.org/10.3390/microorganisms10112134.

Alboukadel Kassambara, Fabian Mundt (2020) ‘factoextra: Extract and Visualize the Results of Multivariate Data Analyses’. Available at: https://CRAN.R-project.org/package=factoextra.

Alfredson, D.A. and Korolik, V. (2005) ‘Isolation and expression of a novel molecular class D β-lactamase, OXA-61, from Campylobacter jejuni’, *Antimicrobial Agents and Chemotherapy*, 49(6), pp. 2515–2518. Available at: https://doi.org/10.1128/AAC.49.6.2515-2518.2005.

Béjaoui, A. *et al.* (2022) ‘Virulence Profiling, Multidrug Resistance and Molecular Mechanisms of Campylobacter Strains from Chicken Carcasses in Tunisia’, *Antibiotics*, 11(7), p. 830. Available at: https://doi.org/10.3390/antibiotics11070830.

Bereswill, S. *et al.* (2011) ‘Novel Murine Infection Models Provide Deep Insights into the “Ménage à Trois” of Campylobacter jejuni, Microbiota and Host Innate Immunity’, *PLoS ONE*. Edited by D.W. Metzger, 6(6), p. e20953. Available at: https://doi.org/10.1371/journal.pone.0020953.

Bleumink-Pluym, N.M.C. *et al.* (2013) ‘Identification of a Functional Type VI Secretion System in Campylobacter jejuni Conferring Capsule Polysaccharide Sensitive Cytotoxicity’, *PLOS Pathogens*, 9(5), p. e1003393. Available at: https://doi.org/10.1371/JOURNAL.PPAT.1003393.

Bortolaia, V. *et al.* (2020) ‘ResFinder 4.0 for predictions of phenotypes from genotypes’, *The Journal of antimicrobial chemotherapy*, 75(12), pp. 3491–3500. Available at: https://doi.org/10.1093/JAC/DKAA345.

Casagrande Proietti, P. *et al.* (2020) ‘Beta-lactam resistance in Campylobacter coli and Campylobacter jejuni chicken isolates and the association between blaOXA-61 gene expression and the action of β-lactamase inhibitors’, *Veterinary Microbiology*, 241, p. 108553. Available at: https://doi.org/10.1016/j.vetmic.2019.108553.

Cascales, E. (2008) ‘The type VI secretion toolkit’, *EMBO reports*, 9(8), pp. 735–741. Available at: https://doi.org/10.1038/EMBOR.2008.131.

CDC (2019) ‘Surveillance for Foodborne Disease Outbreaks United States , 2014 : Annual Report’, pp. 1–16.

Center for Veterinary Medicine (2022) *Antimicrobial Use and Resistance in Animal Agriculture*. Summary report. United States Food and Drug Administration, Center for Veterinary Medicine. Available at: https://www.fda.gov/media/159544/download.

Chen, J.C. *et al.* (2018) ‘Report of erm(B) campylobacter jejuni in the United States’, *Antimicrobial Agents and Chemotherapy*, 62(6). Available at: https://doi.org/10.1128/AAC.02615-17/ASSET/6D441FD3-B142-4933-A074-8E96C22084A9/ASSETS/GRAPHIC/ZAC0061872190001.JPEG.

De Vries, S.P.W. *et al.* (2018) ‘Phylogenetic analyses and antimicrobial resistance profiles of Campylobacter spp. from diarrhoeal patients and chickens in Botswana’, *PLOS ONE*, 13(3), p. e0194481. Available at: https://doi.org/10.1371/JOURNAL.PONE.0194481.

Fan, Y. *et al.* (2022) ‘The Use of Disinfectant in Barn Cleaning Alters Microbial Composition and Increases Carriage of Campylobacter jejuni in Broiler Chickens’, *Applied and Environmental Microbiology*, 88(10), p. e0029522. Available at: https://doi.org/10.1128/AEM.00295-22.

de Fátima Rauber Würfel, S. *et al.* (2020) ‘Campylobacter jejuni isolated from poultry meat in Brazil: in silico analysis and genomic features of two strains with different phenotypes of antimicrobial susceptibility’, *Molecular Biology Reports*, 47(1), pp. 671–681. Available at: https://doi.org/10.1007/S11033-019-05174-Y/FIGURES/5.

Florez-Cuadrado, D. *et al.* (2016) ‘Description of an erm(B)-carrying Campylobacter coli isolate in Europe’, *Journal of Antimicrobial Chemotherapy*, 71(3), pp. 841–843. Available at: https://doi.org/10.1093/JAC/DKV383.

Gharbi, M. *et al.* (2021) ‘Distribution of virulence and antibiotic resistance genes in Campylobacter jejuni and Campylobacter coli isolated from broiler chickens in Tunisia’, *Journal of Microbiology, Immunology and Infection* [Preprint]. Available at: https://doi.org/10.1016/J.JMII.2021.07.001.

Ghatak, S. *et al.* (2017) ‘Whole genome sequencing and analysis of Campylobacter coli YH502 from retail chicken reveals a plasmid-borne type VI secretion system’, *Genomics Data*, 11, pp. 128–131. Available at: https://doi.org/10.1016/J.GDATA.2017.02.005.

Gokulan, K. *et al.* (2013) ‘Impact of Plasmids, Including Those EncodingVirB4/D4 Type IV Secretion Systems, on Salmonella enterica serovar Heidelberg Virulence in Macrophages and Epithelial Cells’, *PLoS ONE*. Edited by N.J. Mantis, 8(10), p. e77866. Available at: https://doi.org/10.1371/journal.pone.0077866.

Griggs, D.J. *et al.* (2009) ‘β-lactamase-mediated β-lactam resistance in Campylobacter species: Prevalence of Cj0299 (blaOXA-61) and evidence for a novel β-lactamase in C. jejuni’, *Antimicrobial Agents and Chemotherapy*, 53(8), pp. 3357–3364. Available at: https://doi.org/10.1128/AAC.01655-08/ASSET/583D8EB8-73FE-4922-B484-3C77CD579710/ASSETS/GRAPHIC/ZAC0080983230003.JPEG.

Guerry, P. *et al.* (2006) ‘Changes in flagellin glycosylation affect Campylobacter autoagglutination and virulence’, *Molecular Microbiology*, 60(2), pp. 299–311. Available at: https://doi.org/10.1111/j.1365-2958.2006.05100.x.

Hansson, I. *et al.* (2018) ‘Knowledge gaps in control of Campylobacter for prevention of campylobacteriosis’, *Transboundary and Emerging Diseases*, 65, pp. 30–48. Available at: https://doi.org/10.1111/TBED.12870.

Hendrixson, D.R. and DiRita, V.J. (2004) ‘Identification of Campylobacter jejuni genes involved in commensal colonization of the chick gastrointestinal tract’, *Molecular Microbiology*, 52(2), pp. 471–484. Available at: https://doi.org/10.1111/J.1365-2958.2004.03988.X.

Hoffmann, S., Maculloch, B. and Batz, M. (2015) *Economic burden of major foodborne illnesses acquired in the United States*, *Economic Cost of Foodborne Illnesses in the United States*, pp. 1–74. Available at: www.ers.usda.gov.

Igwaran, A. and Okoh, A.I. (2019) ‘Human campylobacteriosis: A public health concern of global importance’, *Heliyon*, 5(11), p. e02814. Available at: https://doi.org/10.1016/j.heliyon.2019.e02814.

Jari Oksanen *et al.* (2022) ‘vegan: Community Ecology Package’. Available at: https://CRAN.R-project.org/package=vegan.

Kanwal, S. *et al.* (2019) ‘Variation in antibiotic susceptibility and presence of type VI secretion system (T6SS) in Campylobacter jejuni isolates from various sources’, *Comparative Immunology, Microbiology and Infectious Diseases*, 66, p. 101345. Available at: https://doi.org/10.1016/J.CIMID.2019.101345.

Karlyshev, A.V. *et al.* (2002) ‘Genetic and biochemical evidence of a Campylobacter jejuni capsular polysaccharide that accounts for Penner serotype specificity: Genetics and biochemistry of C. jejuni LPS biosynthesis’, *Molecular Microbiology*, 35(3), pp. 529–541. Available at: https://doi.org/10.1046/j.1365-2958.2000.01717.x.

Kassem, I. *et al.* (2010) ‘Use of bioluminescence imaging to monitor Campylobacter survival in chicken litter’, *Journal of Applied Microbiology*, 109(6), pp. 1988–1997. Available at: https://doi.org/10.1111/J.1365-2672.2010.04828.X.

Kelley, T.R. *et al.* (1995) ‘Bacterial Pathogens and Indicators in Poultry Litter during Re-Utilization’, *Journal of Applied Poultry Research*, 4(4), pp. 366–373. Available at: https://doi.org/10.1093/japr/4.4.366.

Kienesberger, S. *et al.* (2011) ‘Interbacterial Macromolecular Transfer by the *Campylobacter fetus* subsp. *venerealis* Type IV Secretion System’, *Journal of Bacteriology*, 193(3), pp. 744–758. Available at: https://doi.org/10.1128/JB.00798-10.

Lachance, N. *et al.* (1991) ‘Role of the beta-lactamase of Campylobacter jejuni in resistance to beta-lactam agents’, *Antimicrobial Agents and Chemotherapy*, 35(5), pp. 813–818. Available at: https://doi.org/10.1128/AAC.35.5.813.

Lê, S., Josse, J. and Husson, F. (2008) ‘FactoMineR : An *R* Package for Multivariate Analysis’, *Journal of Statistical Software*, 25(1). Available at: https://doi.org/10.18637/jss.v025.i01.

Lertpiriyapong, K. *et al.* (2012) ‘Campylobacter jejuni Type VI Secretion System: Roles in Adaptation to Deoxycholic Acid, Host Cell Adherence, Invasion, and In Vivo Colonization’, *PLoS ONE*. Edited by S. Bereswill, 7(8), p. e42842. Available at: https://doi.org/10.1371/journal.pone.0042842.

Letunic, I. and Bork, P. (2021) ‘Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation’, *Nucleic Acids Research*, 49(W1), pp. W293–W296. Available at: https://doi.org/10.1093/NAR/GKAB301.

Liao, Y.S. *et al.* (2022) ‘Antimicrobial Resistance in Campylobacter coli and Campylobacter jejuni from Human Campylobacteriosis in Taiwan, 2016 to 2019’, *Antimicrobial Agents and Chemotherapy*, 66(1). Available at: https://doi.org/10.1128/AAC.01736-21/SUPPL_FILE/AAC.01736-21-S0001.PDF.

Lin, J. *et al.* (2005) ‘CmeR functions as a transcriptional repressor for the multidrug efflux pump CmeABC in Campylobacter jejuni’, *Antimicrobial Agents and Chemotherapy*, 49(3), pp. 1067–1075. Available at: https://doi.org/10.1128/AAC.49.3.1067-1075.2005.

Line, J. (2002) ‘Campylobacter and Salmonella populations associated with chickens raised on acidified litter’, *Poultry Science*, 81(10), pp. 1473–1477. Available at: https://doi.org/10.1093/ps/81.10.1473.

Line, J.E. and Bailey, J.S. (2006) ‘Effect of On-Farm Litter Acidification Treatments on Campylobacter and Salmonella Populations in Commercial Broiler Houses in Northeast Georgia’, *Poultry Science*, 85(9), pp. 1529–1534. Available at: https://doi.org/10.1093/ps/85.9.1529.

Liu, D. *et al.* (2019) ‘Emerging Erm(B)-mediated macrolide resistance associated with novel multidrug resistance genomic islands in Campylobacter’, *Antimicrobial Agents and Chemotherapy*, 63(7). Available at: https://doi.org/10.1128/AAC.00153-19/SUPPL_FILE/AAC.00153-19-S0004.XLSX.

Lopes, G.V. *et al.* (2021) ‘Virulence factors of foodborne pathogen Campylobacter jejuni’, *Microbial Pathogenesis*, 161, p. 105265. Available at: https://doi.org/10.1016/J.MICPATH.2021.105265.

Marasini, D. *et al.* (2020) ‘Molecular characterization of megaplasmids encoding the type VI secretion system in Campylobacter jejuni isolated from chicken livers and gizzards’, *Scientific Reports 2020 10:1*, 10(1), pp. 1–10. Available at: https://doi.org/10.1038/s41598-020-69155-z.

Montgomery, M.P. *et al.* (2018) ‘Multidrug-Resistant *Campylobacter jejuni* Outbreak Linked to Puppy Exposure — United States, 2016–2018’, *MMWR. Morbidity and Mortality Weekly Report*, 67(37), pp. 1032–1035. Available at: https://doi.org/10.15585/mmwr.mm6737a3.

Mourkas, E. *et al.* (2020) ‘Agricultural intensification and the evolution of host specialism in the enteric pathogen *Campylobacter jejuni*’, *Proceedings of the National Academy of Sciences*, 117(20), pp. 11018–11028. Available at: https://doi.org/10.1073/pnas.1917168117.

Oladeinde, A. *et al.* (2022) ‘Management and environmental factors influence the prevalence and abundance of food-borne pathogens and commensal bacteria in peanut hull-based broiler litter’, *Poultry Science*, p. 102313. Available at: https://doi.org/10.1016/j.psj.2022.102313.

Page, A.J. *et al.* (2015) ‘Roary: rapid large-scale prokaryote pan genome analysis’, *Bioinformatics*, 31(22), pp. 3691–3693. Available at: https://doi.org/10.1093/bioinformatics/btv421.

Pal, C. *et al.* (2014) ‘BacMet: antibacterial biocide and metal resistance genes database’, *Nucleic Acids Research*, 42(D1), pp. D737–D743. Available at: https://doi.org/10.1093/nar/gkt1252.

Pope, M.J. and Cherry, T.E. (2000) ‘An evaluation of the presence of pathogens on broilers raised on poultry litter treatment®-treated litter’, *Poultry Science*, 79(9), pp. 1351–1355. Available at: https://doi.org/10.1093/ps/79.9.1351.

Poulin, M.B. *et al.* (2010) ‘Characterization of a bifunctional pyranose-furanose mutase from Campylobacter jejuni 11168’, *Journal of Biological Chemistry*, 285(1), pp. 493–501. Available at: https://doi.org/10.1074/jbc.M109.072157.

Powell, L.F. *et al.* (2012) ‘The prevalence of *Campylobacter* spp. in broiler flocks and on broiler carcases, and the risks associated with highly contaminated carcases’, *Epidemiology and Infection*, 140(12), pp. 2233–2246. Available at: https://doi.org/10.1017/S0950268812000040.

R Team (2015) *R Team: R: A language and environment for statistical programming*, *URL http://www. R-project. org*. Available at: https://scholar.google.com/scholar?cluster=17691937679756112449&hl=en&oi=scholarr http://scholar.google.com/scholar?q=related:hf0sGkdUijQJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5.

Raivo Kolde (2019) ‘pheatmap: Pretty Heatmaps’. Available at: https://CRAN.R-project.org/package=pheatmap.

Rauber Würfel, S.D.F. *et al.* (2019) ‘Population Dynamics of Thermotolerant Campylobacter in Broilers Reared on Reused Litter’, *Foodborne Pathogens and Disease*, 16(11), pp. 738–743. Available at: https://doi.org/10.1089/fpd.2019.2645.

Reuter, M. *et al.* (2021) ‘Inactivation of the core *cheVAWY* chemotaxis genes disrupts chemotactic motility and organised biofilm formation in *Campylobacter jejuni*’, *FEMS Microbiology Letters*, 367(24), p. fnaa198. Available at: https://doi.org/10.1093/femsle/fnaa198.

Rojas, J.D. *et al.* (2019) ‘Distribution of Capsular Types of Campylobacter jejuni Isolates from Symptomatic and Asymptomatic Children in Peru’, *The American Journal of Tropical Medicine and Hygiene*, 101(3), p. 541. Available at: https://doi.org/10.4269/AJTMH.18-0994.

RStudio Team (2016) ‘RStudio: Integrated Development Environment for R’. Boston, MA: RStudio, Inc.

Scharff, R.L. (2012) ‘Economic Burden from Health Losses Due to Foodborne Illness in the United States’, *Journal of Food Protection*, 75(1), pp. 123–131. Available at: https://doi.org/10.4315/0362-028X.JFP-11-058.

Shen, Z. *et al.* (2018) ‘Antimicrobial Resistance in Campylobacter spp’, *Microbiology Spectrum*, 6(2). Available at: https://doi.org/10.1128/MICROBIOLSPEC.ARBA-0013-2017/FORMAT/EPUB.

Sougakoff, W. *et al.* (1987) ‘Nucleotide sequence and distribution of gene tetO encoding tetracycline resistance in Campylobacter coli’, *FEMS Microbiology Letters*, 44(1), pp. 153–159. Available at: https://doi.org/10.1111/j.1574-6968.1987.tb02260.x.

Stamatakis, A. (2006) ‘RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models’, *Bioinformatics*, 22(21), pp. 2688–2690. Available at: https://doi.org/10.1093/bioinformatics/btl446.

Stamatakis, A. (2014) ‘RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies’, *Bioinformatics*, 30(9), pp. 1312–1313. Available at: https://doi.org/10.1093/BIOINFORMATICS/BTU033.

Stern, N.J. *et al.* (2001) ‘Distribution of Campylobacter spp. in Selected U.S. Poultry Production and Processing Operations’, *Journal of Food Protection*, 64(11), pp. 1705–1710. Available at: https://doi.org/10.4315/0362-028X-64.11.1705.

Tang, Y. *et al.* (2020) ‘Characterization and Prevalence of Campylobacter spp. From Broiler Chicken Rearing Period to the Slaughtering Process in Eastern China’, *Frontiers in Veterinary Science*, 7, p. 227. Available at: https://doi.org/10.3389/fvets.2020.00227.

Taylor, Z.W. and Raushel, F.M. (2018) ‘Cytidine Diphosphoramidate Kinase: An Enzyme Required for the Biosynthesis of the O -Methyl Phosphoramidate Modification in the Capsular Polysaccharides of Campylobacter jejuni’, *Biochemistry*, 57(15), pp. 2238–2244. Available at: https://doi.org/10.1021/ACS.BIOCHEM.8B00279/ASSET/IMAGES/ACS.BIOCHEM.8B00279.SOCIAL.JPEG_V03.

Thwaites, R.T. and Frost, J.A. (1999) ‘Drug resistance in Campylobacter jejuni, C coli, and C lari isolated from humans in north west England and Wales, 1997’, *Journal of Clinical Pathology*, 52(11), pp. 812–814. Available at: https://doi.org/10.1136/jcp.52.11.812.

Van Der Graaf-Van Bloois, L. *et al.* (2016) ‘Campylobacter fetus Subspecies Contain Conserved Type IV Secretion Systems on Multiple Genomic Islands and Plasmids’, *PLOS ONE*, 11(4), p. e0152832. Available at: https://doi.org/10.1371/JOURNAL.PONE.0152832.

Wei, Taiyun and Simko, Viliam (2021) ‘R package “corrplot”: Visualization of a Correlation Matrix’. Available at: https://github.com/taiyun/corrplot.

Whitehouse, C.A. *et al.* (2018) ‘Use of whole-genome sequencing for Campylobacter surveillance from NARMS retail poultry in the United States in 2015’, *Food Microbiology*, 73, pp. 122–128. Available at: https://doi.org/10.1016/j.fm.2018.01.018.

Willis, W.L., Murray, C. and Talbott, C. (2000) ‘Effect of delayed placement on the incidence of Campylobacter jejuni in broiler chickens’, *Poultry Science*, 79(10), pp. 1392–1395. Available at: https://doi.org/10.1093/PS/79.10.1392.

Wood, D.E., Lu, J. and Langmead, B. (2019) ‘Improved metagenomic analysis with Kraken 2’, *Genome Biology*, 20(1), p. 257. Available at: https://doi.org/10.1186/s13059-019-1891-0.

Woyda, R., Oladeinde, A. and Abdo, Z. (2023) ‘Chicken Production and Human Clinical *Escherichia coli* Isolates Differ in Their Carriage of Antimicrobial Resistance and Virulence Factors’, *Applied and Environmental Microbiology*. Edited by M. Vives, pp. e01167-22. Available at: https://doi.org/10.1128/aem.01167-22.

Yao, R., Burr, D.H. and Guerry, P. (1997) ‘CheY-mediated modulation of *Campylobacter jejuni* virulence’, *Molecular Microbiology*, 23(5), pp. 1021–1031. Available at: https://doi.org/10.1046/j.1365-2958.1997.2861650.x.

Young, K.T., Davis, L.M. and DiRita, V.J. (2007) ‘Campylobacter jejuni: molecular biology and pathogenesis’, *Nature Reviews Microbiology 2007 5:9*, 5(9), pp. 665–679. Available at: https://doi.org/10.1038/nrmicro1718.

Zachariah, O.H. *et al.* (2021) ‘Multiple drug resistance of Campylobacter jejuni and Shigella isolated from diarrhoeic children at Kapsabet County referral hospital, Kenya’, *BMC Infectious Diseases*, 21(1), pp. 1–8. Available at: https://doi.org/10.1186/S12879-021-05788-3/TABLES/4.

# CHAPTER 5: MANAGEMENT FACTORS INFLUENCE *SALMONELLA* PERSISTENCE IN REUSED POULTRY LITTER OVER THREE SUCCESSIVE FLOCKS[4]

## 5.1 Summary

*Salmonella* infections are a leading cause of bacterial food-borne illness worldwide. Infections are highly associated with the consumption of contaminated food, and in particular, chicken meat. Understanding how management practices and environmental factors influence *Salmonella* populations in broiler chicken production may aid in reducing the risk of food-borne illness in humans. Utilizing whole genome sequencing with antimicrobial and heavy metal resistance, virulence factor and plasmid identification, we have characterized the genetic diversity of *Salmonella enterica* isolates (n = 55) obtained from broiler chicken litter. *S. enterica* isolates were recovered from the litter of broiler chickens over three consecutive flocks in four broiler houses on a single integrated farm in Georgia, USA. The chickens were raised under a newly adopted “No Antibiotics Ever” program and copper sulfate was administered via drinking water. *In-silico* serovar prediction identified three *S. enterica* serovars: Enteritidis (n = 12), Kentucky (n = 40) and Senftenberg (n = 3). Antimicrobial susceptibility testing revealed that only one *S.* Kentucky isolate was resistant to streptomycin, while the remaining isolates were susceptible to all antibiotics tested. Metal resistance operons, including copper and silver, were identified chromosomally and on plasmids in serovar Senftenberg and Kentucky isolates, respectively.

4 This work has been previously published on a preprint server: Woyda, R., Oladeinde, A., Endale, D., Strickland, T., Lawrence, J.P., Abdo, Z., 2023b. Management factors influence Salmonella persistence in reused poultry litter over three successive flocks. (preprint). https://doi.org/10.1101/2023.03.03.531069

Serovar Kentucky isolates harboring metal resistance operons were the only *Salmonella* isolates recovered from the litter of third flock cohort. These results suggest the addition of copper sulfate to drinking water may have selected for *S.* Kentucky isolates harboring plasmid-borne copper resistance genes and may explain their persistence in litter from flock to flock.

## 5.2 Importance

*Salmonella* foodborne illnesses are the leading cause of hospitalizations and deaths, resulting in a high economic burden on the healthcare system. Globally, chicken meat is one of the highest consumed meats and is a predominant source of foodborne illness. The severity of *Salmonella* infections depends on the presence of antimicrobial resistance genes and virulence factors. While there are many studies which have investigated *Salmonella* strains isolated from post-harvest chicken samples, there is a gap in our understanding of the prevalence and persistence of *Salmonella* in pre-harvest and in particular their makeup of antibiotic resistance genes, virulence factors and metal resistance genes. The objective of this study was to determine how on-farm management practices and environmental factors influence *Salmonella* persistence, as well as the antimicrobial resistance genes and virulence factors they harbor. In this study we demonstrate that broiler chickens raised without antibiotics are less likely to harbor antibiotic resistance, however the practice of adding acidified copper sulfate to drinking water may select for strains carrying metal resistant genes.

## 5.3 Introduction

*Salmonella* is responsible for an estimated 1.2 million illnesses annually in the United States, 1 million of which are attributed to the consumption of contaminated food (Scallan *et al.*, 2011). Chicken meat is the most consumed meat in the United States (USDA, Economic Research

Service, 2022) and is a predominant source of food-borne illness (ADAS, 2016). *Salmonella* food-borne outbreaks are strongly associated with consumption of chicken meat (ADAS, 2016; Antunes *et al.*, 2016). Moreover, the annual economic loss due to *Salmonella* infections is an estimated $11 billion in the United States (Wernicki, Nowaczek and Urban-Chmiel, 2017).

*Salmonella enterica* encompasses the most pathogenic species and consists of thousands of different serovars, some of which are host-specific while some maintain the ability to infect a broad range of hosts (Mezal *et al.*, 2014). The United States Department of Agriculture Food Safety and Inspection Service surveillance fiscal year 2022 reports indicate the top *Salmonella* serovars isolates from domestic chicken samples are, in decreasing order, Infantis, Kentucky, Enteritidis and Typhimurium (USDA, 2022). In 2018, the Foodborne Diseases Active Surveillance Network (FoodNet) identified Enteritidis, Newport, and Typhimurium serovars were the most frequently identified as causing illness in humans (Tack et al., 2019). While some serovars are capable of infecting both chickens and humans (e.g., Enteritidis and Typhimurium), serovars such as Kentucky and Sofia are highly prevalent in poultry but have a low association with human outbreaks (Ferrari *et al.*, 2019). Thus, the threat of *Salmonella* outbreaks in humans is serovar dependent, however the severity of disease and the efficacy of treatment depends on the strain's virulence factor (VF) and antimicrobial resistance gene (ARG) repertoire. Importantly, VFs and ARGs may be shared across *Salmonella* serovars and, in some instances, across different bacterial genera if these genetic elements are located on specific plasmids or other mobile genetic elements (Fricke *et al.*, 2009; Oladeinde *et al.*, 2019).

*Salmonella* pathogenicity is dependent on the VF repertoire and pathogenesis proceeds through attachment to host tissues, invasion of host tissues, macrophage survival, replication and subsequent dissemination (Gao, Wang and Ogunremi, 2020). Virulence factors in *Salmonella* can

be accumulated and, in combination, allow for successful host colonization and bypassing of host defenses. These virulence factors are typically encoded within *Salmonella* pathogenicity islands or on plasmids (Jajere, 2019; Gao, Wang and Ogunremi, 2020).

Antimicrobial therapy is the first choice treatment for *Salmonella* infections. However, the global rise and threat of antimicrobial resistance, due to the misuse of antibiotics in food-producing animals and human medicine, has put pressure on the broiler production industry to reduce usage of antimicrobials (United States Food and Drug Administration, Center for Veterinary Medicine, 2022). Difficulties in treating *Salmonella* infections occur when strains are resistant to multiple antimicrobials from distinct antimicrobial classes; such strains have been identified in poultry, beef and pork products (Gieraltowski *et al.*, 2016; Bearson *et al.*, 2019; Feng *et al.*, 2020). Importantly, strains harboring ARGs which confer resistance to antimicrobials critically important to human medicine (World Health Organization, 2019), as well as those possessing VF increasing pathogenicity to humans, will pose a threat and will be harder to treat in the event contaminated food products reach consumers (Fricke *et al.*, 2009; Jajere, 2019; Zakaria *et al.*, 2022). Additionally, the use of heavy metals in chicken production, such as copper for growth promotion and/or sanitation of water lines, has led to the discovery that genes conferring metal resistance are often carried on the same mobile elements as ARGs and VFs (Dziewit *et al.*, 2015; Pal *et al.*, 2015; Bukowski *et al.*, 2019; Souza *et al.*, 2022). Resistance to metals may increase the virulence of a bacterial pathogen, e.g., overcoming iron toxicity and invading and colonizing different tissue types (Nairz *et al.*, 2015), as well as give a competitive advantage to strains in production environments where metals are used for antimicrobial agents (Bearson *et al.*, 2019). Thus, it is critical to understand how on-farm management practices and environmental factors influence strains harboring both ARGs, VFs and metal resistance genes to

develop prediction, or management, methods to limit their occurrence in post-harvest chicken production.

Broiler chickens are raised on litter for an entire grow-out (0 – 56 days). Broiler litter is a complex matrix consisting of a decomposing plant-based bedding material, chicken excreta, feathers, and feed. Broiler chickens are copraphagic by nature and therefore will start to consume the litter upon placement in a broiler house. Thus, it is reasonable to assume that chicks will be exposed to pathogens during both the initial stages of growth and throughout their grow-out through normal activities such as pecking and bathing. Many studies have reported the occurrence of *Salmonella* in broiler litter (Jones *et al.*, 1991; Kelley *et al.*, 1995, p. 19; Pope and Cherry, 2000; Line, 2002; Brooks *et al.*, 2010; Shepherd *et al.*, 2010; Velasquez *et al.*, 2018), but only a few studies have examined the genetic factors that allow *Salmonella* to persist in pre-harvest broiler production (Roberts *et al.*, 2013; Vaz *et al.*, 2017; Shang, Wei and Kang, 2018; Dunn *et al.*, 2022).

The objective of this study was to determine how management practices and environmental factors affect the persistence of *Salmonella*, as well as their VFs, ARGs and metal resistance genes, in pre-harvest broiler production. In this study, we performed an in-depth genomic characterization of *Salmonella* isolates recovered from the litter of four co-located broiler chicken houses over three consecutive flocks (Oladeinde *et al.*, 2023). We previously reported that these isolates were unequally distributed across the litter of the broiler houses and showed that the probability of detecting *Salmonella* was higher for the first broiler cohort raised on litter compared to cohorts 2 and 3 (Oladeinde *et al.*, 2023). Here, we performed antimicrobial susceptibility testing and whole genome sequencing on 55 isolates recovered from the litter (Oladeinde *et al.*, 2023). While antibiotic susceptibility testing resulted in one isolate displaying

phenotypic resistance, our genomic characterization identified a potential beta-lactamase gene as well as copper and silver resistance operons. Characterization of the antimicrobial resistance genes, biocide and metal resistance genes and virulence factors revealed a large core set of genes with differences explained by isolates' serovar. Only isolates harboring specific copper and silver resistance operons were detected over the three consecutive flock cycles. Together these results suggest the on-farm management practice of adding copper sulfate to the drinking water, commonly practiced in the industry, may provide means for *Salmonella* persistence over multiple flock cohorts.

## 5.4 Results

### 5.4.1 *Salmonella* serovar occurrence differed across grow-out periods, flocks, and broiler houses

*Salmonella* prevalence in litter collected from four broiler house floors was determined by direct and selective enrichment plating of litter eluate on both Brilliant Green Sulfur (BGS) and Xylose Lysine Tergitol-4 (XLT-4) agar (Oladeinde *et al.*, 2023). *Salmonella* was detected in 10.76% (31/288) of litter samples over 3 consecutive flock cohorts across 4 co-located broiler houses (Oladeinde *et al.*, 2023). Fifty-five *Salmonella* isolates were selected for whole genome sequencing. These isolates were obtained either through direct plating or enrichment methods and encompassed all unique serogroups (Oladeinde *et al.*, 2023). *In silico* serovar prediction identified three different serovars: *S.* Enteritidis (n = 12), *S.* Kentucky (n = 40), and *S.* Senftenberg (n = 3) (**Table S5.1**). A majority of isolates (85.5%, n = 47) were recovered during the late grow-out (32-38 days old chickens). Each serovar identified was present in both early (4-14 days old chickens) and late grow-out phases (**Table S5.1**). All *S.* Enteritidis isolates were

identified as ST11, all *S.* Kentucky isolates as ST152 (excluding SK34 and SK36 due to low sequencing coverage (< 20X)), and *S.* Senftenberg isolates were identified as ST14 (excluding SS26 due to low sequencing coverage (< 20X)). Multi-locus sequence typing (MLST) was performed using the PubMLST website and the MLST software (Jolley and Maiden, 2010).

*Salmonella* occurrence differed over multiple flock cohorts (**Table 5.1,** (Oladeinde *et al.*, 2023)). The 37 *Salmonella* isolates from flock 1 consisted of the following serovars: Enteritidis (n=9), Kentucky (n=25), Senftenberg (n=3) and a single unidentified isolate. The second flock cohort included 10 *Salmonella* isolates classified as Enteritidis (n=3) and Kentucky (n=7). The final flock cohort, flock 3, harbored only *S.* Kentucky isolates (n=8). The distribution of isolates and serovars also differed across the broiler houses (**Figure 5.1**). Houses 2 (33%, n = 18) and 3 (51%, n = 28) harbored majority of the sequenced isolates and the remaining 9 (16%) originated from house 4. *Salmonella* was not detected from any litter samples from house 1. *Salmonella* isolates sequenced from house 4 included a single *S.* Enteritidis isolate as well as a set of 8 *S.* Kentucky isolates which were all clustered together within the back section of the broiler house (**Figure 5.1**).

**Table 5.1.** Occurrence of *Salmonella* species in peanut hull litter over 3 grow-out cycles across 4 co-located broiler houses.

| **Broiler House** | **Flock cohort 1** | | | **Flock cohort 2** | | | **Flock cohort 3** | | |
|---|---|---|---|---|---|---|---|---|---|
| | ***S.* Enteritidis** | ***S.* Kentucky** | ***S.* Senftenberg** | ***S.* Enteritidis** | ***S.* Kentucky** | ***S.* Senftenberg** | ***S.* Enteritidis** | ***S.* Kentucky** | ***S.* Senftenberg** |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 8 | 8 | 0 | 0 | 2 | 0 | 0 | 0 | 0 |
| 3 | 1 | 17 | 3 | 2 | 5 | 0 | 0 | 8 | 0 |
| 4 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |

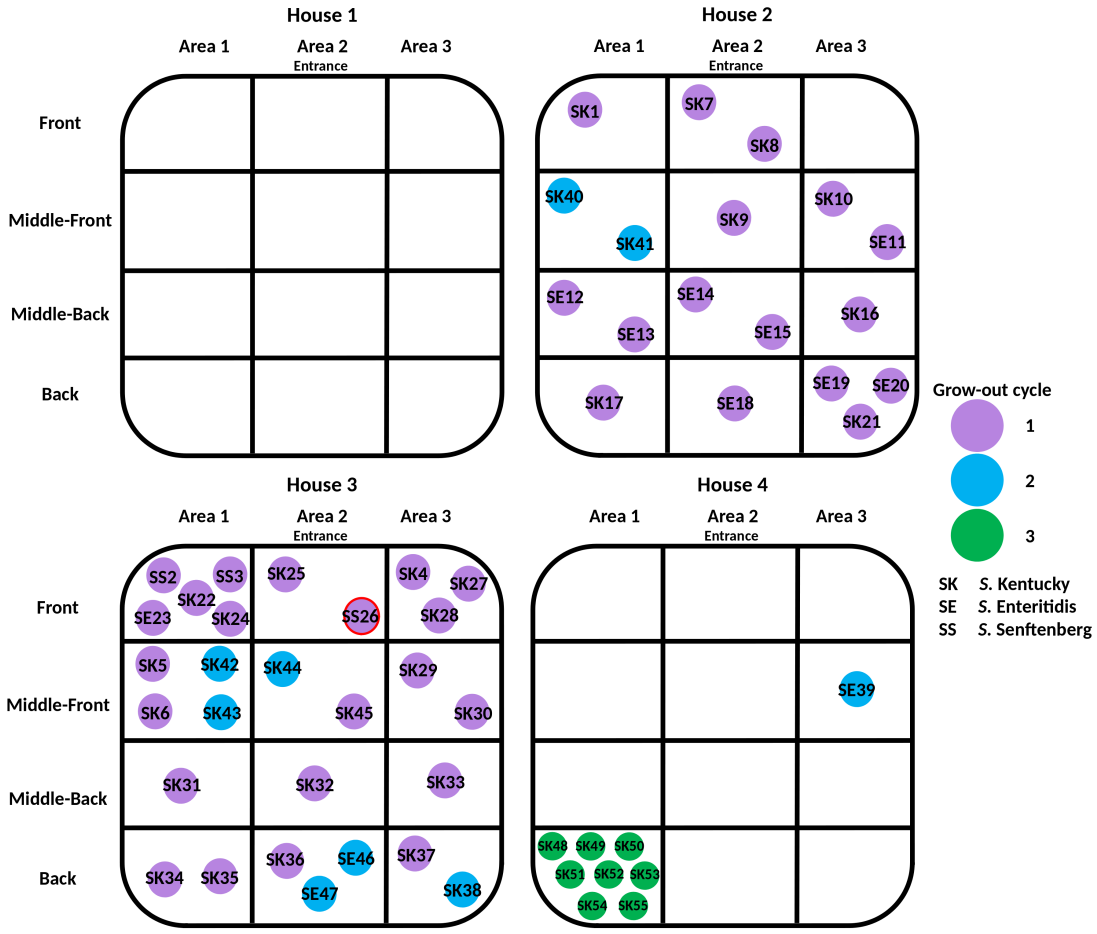



**Figure 5.1.** Visual representation of *Salmonella* species isolated from peanut hull litter within each section of each grow house. Samples were taken from 4 co-located grow houses over 3 grow-out cycles. Circles represent individual isolates labeled by their prospective serovar: *Salmonella* Kentucky (SK), *Salmonella* Enteritidis (SE) and *Salmonella* Senftenberg (SS). Circle color indicates which grow-out cycle an isolate was obtained from: grow-out cycle 1 (purple), grow-out cycle 2 (blue) and grow-out cycle 3 (green). Isolates outline in red indicate isolates not included in further analysis due to poor sequencing coverage (SE23).

### 5.4.2 Serovar Kentucky and Senftenberg harbored additional copper and silver resistance operons

Antibacterial biocides and metal resistance genes were identified using the experimentally confirmed resistance function BacMet database (Pal *et al.*, 2014). A total of 125 genes with >80% identity to the reference database entry were identified and collectively conferred

resistance to 74 unique compounds (**Table S5.2**). 108 of these genes were found to be present in >96.4% of isolates. Many copper and other metal homeostasis related genes were present in >96.4% of isolates (e.g., *copA* (detoxification of cytoplasmic Cu[I]) (Rensing and Grass, 2003), *cueP* (periplasm copper binding) (Pezza *et al.*, 2016), and *zraP/zraR* (involved in intrinsic antibiotic and zinc resistance) (Rome *et al.*, 2018)) (**Table S5.2**). In particular, the *silABCPRS* and *pcoABCESR* resistance operons (silver and copper, respectively) were identified only in serovar Kentucky and Senftenberg isolates. (**Figure 5.2**). These operons are often clustered together on a plasmid and compose the copper homeostasis and silver resistance island (CHARSI) (Staehlin *et al.*, 2016). While these operons appear to be chromosomally encoded within the Senftenberg isolates (contigs >390kbp), within Kentucky isolates these copper and silver operons appear to reside on an IncI1 plasmid.

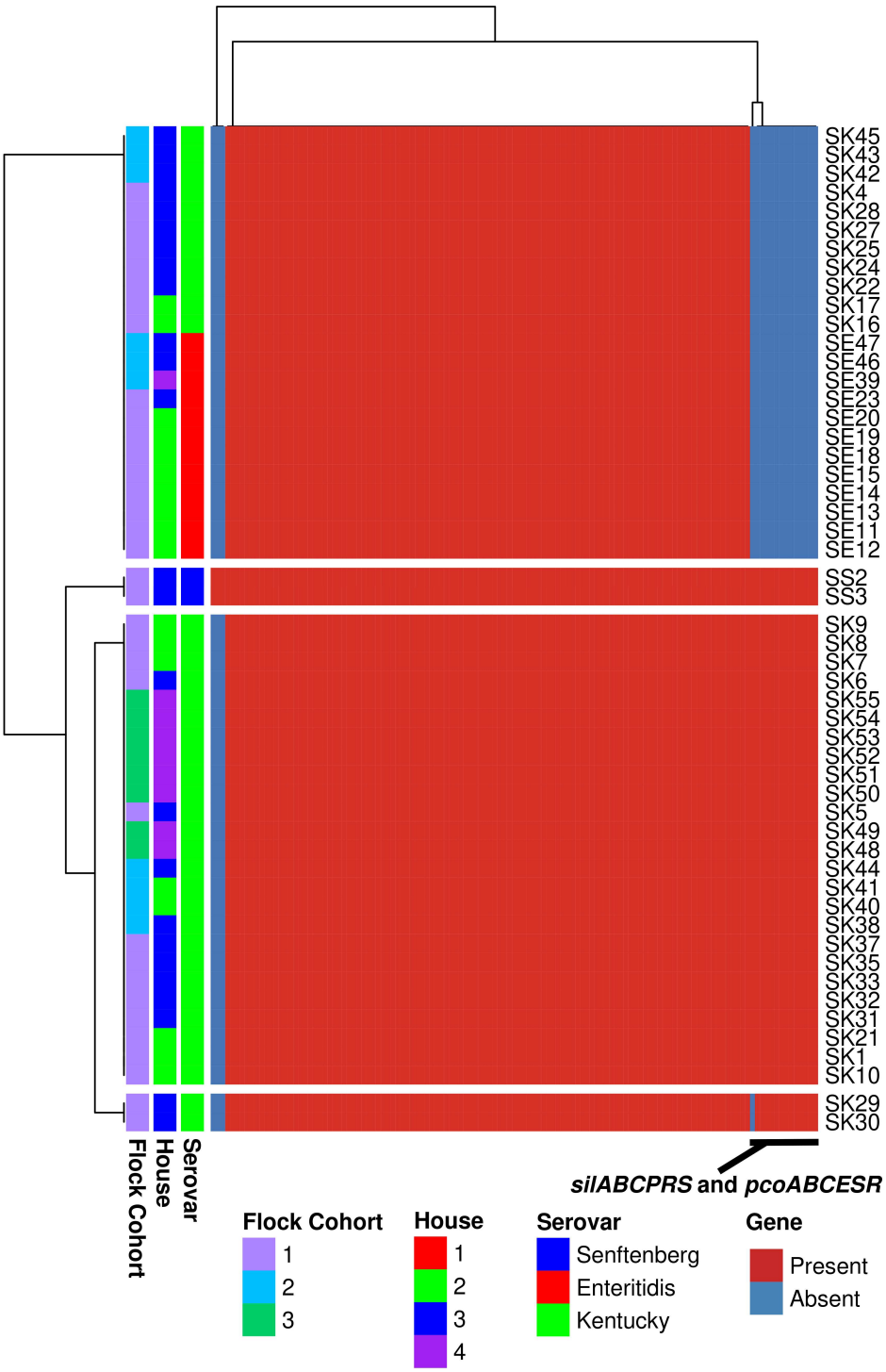



**Figure 5.2.** Heatmap of biocide and metal resistance genes. Genes conferring resistance to biocide and metal resistance were obtained via BacMet. Heatmap was generated in R v4.0.4 with pheatmap v1.0.12 (clustering_method = "average" (UPGMA)) using the filtered, >= 80% identity, presence/absence table of biocide and metal resistance genes (**Table S5.2**).

### 5.4.3 *Salmonella* isolates share a majority of identified virulence factors

A total of 119 VFs were identified utilizing the Virulence Factor Database (VFDB) via ABRICATE and 86 (72%) VFs were present in all isolates (**Table S5.1**). Of the remaining 33 VFs, 19 were present in the majority of isolates (observed within > 70% of those isolates but less than 100%) while the remaining 14 where only observed in 4% - 21% of all isolates. Hierarchical clustering revealed that differences in VF across all isolates could be explained by an isolate's serovar (**Figure 5.3**). The 119 VFs are involved in 12 different functions such as secretion systems (63%), fimbrial adherence determinants (20%), iron uptake (4%), adherence (2.5%) and magnesium uptake (1.7%) (**Table 5.2, Table S5.3**). Taken together, the *Salmonella* isolates harbored a similar core set of virulence factors and differences observed were explained by the isolates' serovar.

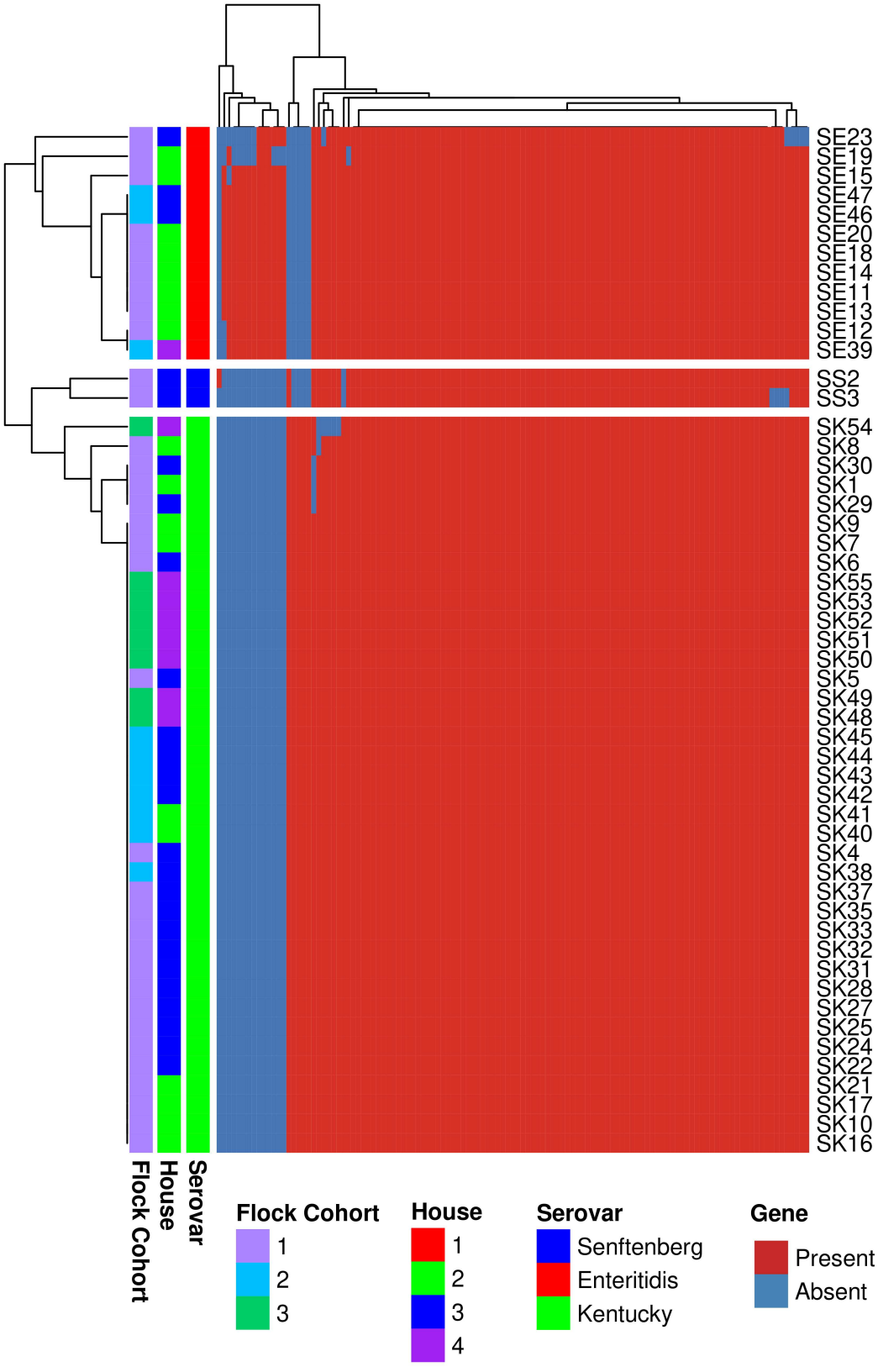

SE23
SE19
SE15
SE47
SE46
SE20
SE18
SE14
SE11
SE13
SE12
SE39
SS2
SS3
SK54
SK8
SK30
SK1
SK29
SK9
SK7
SK6
SK55
SK53
SK52
SK51
SK50
SK5
SK49
SK48
SK45
SK44
SK43
SK42
SK41
SK40
SK4
SK38
SK37
SK35
SK33
SK32
SK31
SK28
SK27
SK25
SK24
SK22
SK21
SK17
SK10
SK16
Flock Cohort
House
Serovar
Flock Cohort
1
2
3
House
1
2
3
4
Serovar
Senftenberg
Enteritidis
Kentucky
Gene
Present
Absent

**Figure 5.3.** Heatmap of virulence factor genes. Virulence factors were identified from the Virulence Factor Database via ABRICATE. Heatmap was generated in R v4.0.4 with pheatmap v1.0.12 (clustering_method = "average" (UPGMA)) using the virulence factor data from **Table S5.1**.

**Table 5.2.** Enumeration of virulence factor functional categories of virulence factors identified from the Virulence Factor Database (VFDB) (via ABRICATE (Reads2Resistome)).

| Function | Number of genes |
|---|---|
| Secretion system | 75 |
| Fimbrial adherence determinants | 24 |
| Iron uptake | 5 |
| Nonfimbrial adherence determinants | 4 |
| Adherence | 3 |
| Magnesium uptake | 2 |
| Invasion | 1 |
| Macrophage inducible genes | 1 |
| Regulation | 1 |
| Serum resistance | 1 |
| Stress adaptation | 1 |
| Toxin | 1 |

### 5.4.4 Serovar Kentucky and Senftenberg isolates harbored unique iron acquisition virulence factors

Serovar Kentucky isolates, regardless of the flock cohort they were obtained from, harbored similar virulence factor profiles (**Figure 5.3**). Unique to Kentucky isolates was the iron acquisition gene *entE* (involved in enterochelin synthesis) and shared with serovar Senftenberg isolates was the iron transport gene *fepC* (involved in enterochelin transport) (Bearson *et al.*, 2008). Iron regulation is important for intracellular pathogens such as *Salmonella* as it is required for growth. For example, tight regulation of intracellular iron is crucial due to the detrimental effect of excess intracellular iron (Andrews, Robinson and Rodríguez-Quiñones, 2003). These iron regulation-related genes, in addition to those uniquely harbored by Kentucky and Senftenberg isolates, suggest an enhanced ability to regulate iron acquisition.

### 5.4.5 Serovar Kentucky and Enteritidis isolates harbor distinct fimbrial adherence genes

Pathogenic bacteria employ a host of cell surface adhesins to adhere and eventually colonize novel host tissues such as the gasterointestinal tract (Kuijpers *et al.*, 2019; Božić *et al.*, 2020). Serovar Kentucky and Enteritidis isolates, each, harbored a distinct adhesion-related operon. Serovar Kentucky harbored the entire *fae* operon (*fae* of the K88 F4 fimbrial gene cluster as observed from our VFDB analysis) (**Table S5.1**).

Serovar Enteritidis isolates uniquely harbored 13 virulence factors which were present in a majority of Enteritidis isolates (**Figure 5.3**). These genes included the plasmid encoded fimbrial operon, pefABCD, responsible for adhesion to the intestinal epithelium (Seribelli *et al.*, 2020)

and 3 genes of the spvABCDR operon (*spvB, spvC,* and *spvR*) which is associated with survival and replication in macrophages (Rychlik, Gregorova and Hradecka, 2006). Lastly, the *rck* (resistance to complement killing) gene was identified in these isolates and is known to increase serum resistance and adhesion to epithelial cell lines (Heffernan *et al.*, 1994). These virulence genes aid in colonization and survival within host tissues and are similar to the functions of the unique VFs identified in Kentucky isolates. However, Kentucky isolates were isolated from the third flock cohort, while no Enteritidis isolates were isolates. These data suggest that the litter environment may impose a selective pressure on these virulence genes and the isolates that harbor them.

### 5.4.6 *Salmonella* isolates harbored similar AMR profile

Forty-one antimicrobial resistance genes (ARGs) were identified using the Resistance Gene Identifier (RGI), which utilizes the Comprehensive Antibiotics Resistance database (CARD) (Alcock *et al.*, 2020). Of these ARGs, 39 were present in 100% of the isolates (**Figure 5.4**). The 41 ARGs were predicted to confer resistance to 26 different antibiotic drug classes (**Table S5.4**). *kdpE* gene, a transcriptional activator of a two-component potassium transport system and $Bla_{\text{TEM-60}}$ (identified as an hypothetical protein using Prokka and NCBI BLAST search), an extended-spectrum beta-lactamase gene, were identified in 38 and 40 isolates, respectively. *kdpE* has been reported to regulate virulence genes (Hughes *et al.*, 2009; Zhao *et al.*, 2010; Freeman, Dorus and Waterfield, 2013) and was found in all serovars from this study. $Bla_{\text{TEM-60}}$ was present in all serovar Kentucky and Senftenberg isolates and absent from all Enteritidis isolates (**Table 5.3**).

Additionally, identification of acquired ARGs (genes likely to be harbored on a mobile genetic elements such as plasmids) via ResFinder (Bortolaia *et al.*, 2020) resulted in only two

hits; the *sitABCD* operon and the *aac(6')-Iaa* gene. The *sitABCD* operon was present in all serovar Enteritidis and Kentucky isolates and absent from Senftenberg (**Table 5.3**). However, the *sitABCD* operon was at 72% percent identity which was below the 80% coverage identity cutoff for inclusion in our analysis. SitABCD is a putative iron transporter which plays a role in iron acquisition during infection and is important for growth in tissues following invasion of intestinal epithelium (Janakiraman and Slauch, 2000). The *aac(6')-Iaa* gene, a chromosome-encoded aminoglycoside acetyltransferase (identified as *AAC(6')-Iy* gene using the CARD database*)* was also present in all isolates (**Table S5.1**). *aac(6')-type* genes have commonly been found in *S.* Kentucky and Typhimurium strains, however a study of 2,762 isolates harboring *aac(6')-type* or *aac(6')-Iy* gene found that only 11 isolates exhibited phenotypic resistance to aminoglycosides (Neuert *et al.*, 2018).

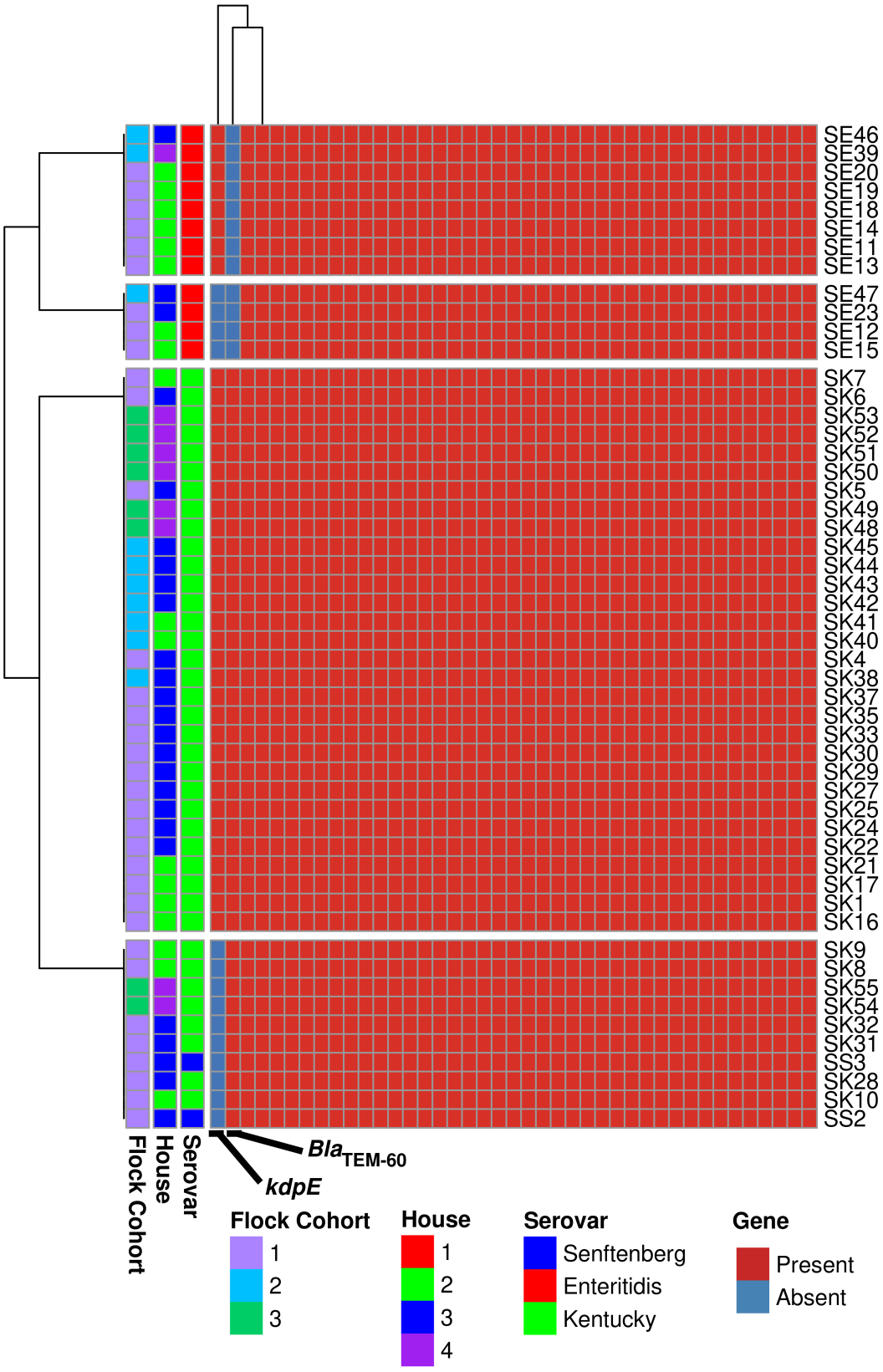

SE46
SE39
SE20
SE19
SE18
SE14
SE11
SE13
SE47
SE23
SE12
SE15
SK7
SK6
SK53
SK52
SK51
SK50
SK5
SK49
SK48
SK45
SK44
SK43
SK42
SK41
SK40
SK4
SK38
SK37
SK35
SK33
SK30
SK29
SK27
SK25
SK24
SK22
SK21
SK17
SK1
SK16
SK9
SK8
SK55
SK54
SK32
SK31
SS3
SK28
SK10
SS2
Flock Cohort
House
Serovar
Bla TEM-60
kdpE
Flock Cohort
1
2
3
House
1
2
3
4
Serovar
Senftenberg
Enteritidis
Kentucky
Gene
Present
Absent

**Figure 5.4.** Heatmap of antimicrobial resistance genes identified. Antimicrobial resistance genes were identified from RGI and ResFinder. Heatmap was generated in R v4.0.4 with pheatmap v1.0.12 (clustering_method = "average" (UPGMA)) using the RGI and ResFinder data from **Table S5.1**.

In summary, the majority of the identified ARGs were chromosomally encoded and did not confer phenotypic resistance to the predicted antibiotics. This conclusion was supported by antibiotic susceptibility testing (AST) results. AST showed that 54/55 isolates were susceptible to the fourteen antibiotics tested (**Table S5.5**), while one isolate (SK32) exhibited resistance to streptomycin. We did not find any known ARG or chromosomal mutation that could explain the streptomycin resistance observed in SK32. Taken together, the *Salmonella* isolates in this study do not harbor ARGs capable of conferring clinically relevant levels of antibiotic resistance except *S.* Kentucky strain SK32.

### 5.4.7 Co-occurrence of virulence factors and plasmid contigs

Five plasmid replicons were identified from the PlasmidFinder database (Carattoli *et al.*, 2014) within serovar Kentucky isolates; “IncI1_1_Alpha”, “IncX1_3”, “Col156”, “Col(BS512) “and “Col(MG828)” (**Table 5.3, Table S5.6**). IncI1 and IncX1 were identified in all Kentucky isolates (38/38). A serovar Kentucky isolate (SK4) also harbored the Col plasmid replicons Col156 and Col(BS512) on short contigs of lengths 3,904 bp and 3,137 bp, respectively (**Table S5.6**). A total of 129 genes were identified on the IncI1 and IncX1 contigs, using the contigs from isolates SK1 and SK16, respectively (**Table S5.7**). The IncI1 contigs harbored silver and copper resistance operons (i.e., *silABCPRS* and *pcoABCESR*, respectively) as well as conjugation proteins (PilI, PilJ, PilK, PilL, PilM, PilN, PilO, PilP, TraA, TraB and TraC). A serovar Kentucky isolate was selected for long read sequencing to confirm the presence of the IncI1 plasmid within serovar

Kentucky isolates. These results confirmed an 81,816 bp plasmid harboring both the *silABCPRS* and *pcoABCESR* operons (**Figure S5.1**).

IncFII(S), IncFIC(FII) and IncFIB(S) replicons were identified in serovar Enteritidis isolates (8/11, 8/11, 10/11 isolates, respectively) (**Table S5.6**). Rapid Annotation using Subsystem Technology (RAST) annotation of the isolate SE13 contig containing all 3 IncF replicons identified 97 genes (**Table S5.7**). This contig harbored the *spvABCDR* operon, *pefABCD* operon and *mig-5* genes in addition to the IncF plasmid conjugative proteins FinO, TraV, TrbD, TraA, TraB, TraK, TraE, TraL and TraA (**Figure S5.2**). No plasmid replicons identified in either *S.* Kentucky or *S.* Enteritidis isolates were lost over the multiple flock cohorts. These data suggest that VF and plasmid replicons remained constant within and across serovars. VFs and plasmid replicons harbored by *S.* Kentucky and *S.* Enteritidis isolates may be under specific environmental pressure from the litter environment.

## 5.5 Discussion

The purpose of this study was to characterize the genomes of *Salmonella* isolates recovered from peanut hull-based broiler litter during the grow-out of 3 consecutive flocks of broiler chickens. Our objective was to identify ARGs, VFs and metal resistance genes harbored by the *Salmonella* isolates as well as understand how management and environmental factors can lead to genomic changes or persistence over multiple flock cohorts. Other studies, such as the one performed by Roll et al (Roll, Dai Prá and Roll, 2011) have demonstrated that *Salmonella* could persist over multiple broiler flock cohorts but no studies to the authors' knowledge have characterized the genome of *Salmonella* isolates from flock to flock.

We isolated *S.* Kentucky, *S.* Enteritidis and *S.* Senftenberg serovars in the litter of the broiler houses studied (**Table 5.1**). Serovar Kentucky had the highest prevalence and persisted from flock 1 to 3. It is not surprising that *S.* Kentucky was the most dominant serovar in our study as it is one of the most commonly isolated poultry serovars from domestic chicken samples in the United States (Dunn *et al.*, 2022). *S.* Kentucky was also the major serovar in breeder flocks from 2016 to 2020 (Siceloff, Waltman and Shariat, 2022). Unique to serovar Kentucky isolates were the VFs of the K88, F4, fimbrial *fae* operon and the iron uptake-associated gene *entE*. Also, unique to serovar Kentucky isolates were the IncX1and IncI1 plasmid replicons, of which the IncI1 contigs harbored copper and silver resistance operons in addition to conjugation related genes (**Table S5.2**). Our results suggest that *S.* Kentucky has evolved with VFs and plasmids which allow it to colonize broiler chickens and persist in the broiler house environment. Kentucky isolates harbored multiple metal resistance operons which may have conferred a competitive advantage to surviving within the production environment, resulting in their persistence in the 3 flock cohorts. It is possible that the waterers and water lines acted as reservoirs for these isolates as copper sulfate was added to the drinking water. Addition of acidified copper sulfate is a common practice to sanitize water lines and waterers (Scott *et al.*, 2018). However, while Senftenberg isolates harbored these operons on the chromosome, no Senftenberg isolates were observed during the second and final flock cohorts. It is plausible that metal resistance genes harbored on the chromosome poses a higher fitness cost on the bacterial host (e.g., *S.* Seftenberg) compared to when they are carried on plasmids (e.g., *S.* Kentucky). Our results suggest that *S.* Kentucky has evolved with VFs and plasmids which allow it to colonize broiler chickens and persist in the broiler house environment.

Virulence factor differences were explained by serovars (**Figure 5.3**), nevertheless all serovars harbored VFs with functions relating to secretion systems, fimbrial adherence determinants, invasion, iron uptake, macrophage inducible genes, magnesium uptake and nonfimbrial adhesion determinants (**Table S5.3**). *Salmonella* Enteritidis isolates harbored additional set of 13 VFs that were not present in either Kentucky or Senftenberg isolates. These VFs included plasmid genes (*spvABCDR* operon, *pefABCD* operon and *mig-5*) located on IncFIB and IncFII plasmid replicon contigs (**Figure S5.2)**. These VFs have been identified in human cases associated with serovars Typhimurium and Enteritidis (Kuijpers *et al.*, 2019; Seribelli *et al.*, 2020). As these VFs encode for functions conferring increased fitness and survival within host tissues, is it possible that the chickens acted as a reservoir for these strains over multiple flock cohorts, and through the copraphagic nature of chickens, were continually deposited in and re-ingested from litter. Interestingly, house 4 had no detectable *S.* Enteritidis in the first flock however an Enteritidis isolate was obtained during the grow-out of the second flock. Similarly, no *S.* Kentucky isolates were obtained during the first 2 flocks in house 4 however during the third flock cohort 8 Kentucky isolates were obtained. This suggests that human or rodent transmission between houses and/or failed detection by the methods used are possible reasons for this observation (Backhans and Fellström, 2012). Additionally, it is possible that these isolates were introduced from the hatchery upon placement of the flock 3 cohort of chicks.

We found that while all *Salmonella* isolates harbored a core set of ARGs and VFs (95% and 72%, respectively) (**Figure 5.4, Figure 5.3**), each set of isolates grouped by serovar harbored a distinct set of ARGs, VFs, and plasmids (**Table 5.3**). While ARG identification resulted in 42 chromosome encoded ARGs, AST revealed that only one isolate was resistant to streptomycin (**Table S5.5**). This low prevalence of AMR is not surprising since the integrated

farm adopted a “No Antibiotics Ever” program during grow-out (Oladeinde *et al.*, 2023). Therefore, it is plausible that no antibiotic selective pressure existed in the litter.

We have provided new data on the genomic characteristics of *Salmonella* serovars Kentucky, Enteritidis and Senftenberg found in peanut hull-based litter during pre-harvest broiler chicken grow-out. We demonstrated that AMR was similar across all serovars while VF and plasmid profiles varied with respect to each serovar. Serovar Enteritidis harbored IncF plasmid associated replicons that were located on the same contigs as virulence factors associated with pathogenicity in humans, while serovar Kentucky isolates harbored an IncI1 plasmid harboring copper and silver resistance proteins. While some *S.* Kentucky isolates persisted from flock 1 to 3, no *S.* Enteritidis isolates did. Therefore, these data suggest that some *Salmonella* serovars and strains are equipped with genetic factors that allow them to persist in broiler house environments under specific management practices including “NAE” programs that administer copper sulfate via drinking water. Nonetheless, there are several limitations of the study which could have biased our interpretation of the results including the small number of flocks monitored and the limit of detection of the method used for *Salmonella* isolation. Similarly, the initial broiler house cleaning procedures may have failed to remove all residual contaminants from previous flocks and subsequently resulted in these strains colonizing the litter environment. Thus, it is plausible the cleaning procedures resulted in differential results across each broiler house. Lastly, the results described were from one broiler farm that reused peanut hull-based litter and may not be representative of all farms that reuse litter for poultry production.

## 5.6 Materials and Methods

Methods involving litter management on-farm, sample collection of litter as well as microbiological methods for bacterial isolation have been previously described (Oladeinde *et al.*, 2023). Materials and methods used in this study, and others re-described from the previous work, is presented below.

### 5.6.1 Study design

Three cohorts of broiler flocks were raised in succession in 4 co-located integrated commercial broiler houses in South Georgia between February and August 2018. Each of the broiler houses contained 22,000 to 24,000 broilers per flock. Each of the four co-located broiler houses underwent a full cleanout before fresh peanut hull-based litter placement. Each cohort was raised on the previous flock's litter without any cleanout between and only mechanical conditioning was performed to remove caked portions. A commercial litter acidifier was applied during the downtime between flock cohorts for ammonium control (typically 1 week before sampling). Half house brooding was performed for the first 14 days of each flock. This practice restricted the birds to the front section of each broiler house. Additionally, copper sulfate was given via drinking water. Management practices and procedures employed are within the scope of routine industry practices.

### 5.6.2 Litter sampling

Litter samples (n = 288) were collected across four broiler houses during early (chick age between 4-14 days old) and late grow-out (32-38 days old) for three consecutive flocks. This amounted to 6 sampling times per broiler house; each flock was sampled at two separate time points. Each house was divided into 12 subsections; four sections from front to back and each

section was divided into three subsections from left to right (**Figure 5.1**). During sampling, three litter grab samples were taken from each of the 12 subsections and pooled into one Whirl Pak bag. Pooled litter samples were then transported on ice to the United States National Poultry Research Center for further processing.

### 5.6.3 *Salmonella* isolation and identification

All samples were processed within 24 h of collection. Thirty grams from each pooled litter grab sample was mixed with 120 ml phosphate buffered saline and shook with a hand wrist shaker (Boekel Scientific, Model 401000) for 10 min. Litter eluate (100 µl) was direct plated to both Brilliant Green Sulfur (BGS) and Xylose Lysine Tergitol-4 (XLT-4) agars. Plates were incubated 18-24 h at 37°C. Additionally, aliquots (1 ml) of the eluate were enriched in buffered peptone water (9 ml) for 18-24 h at 37°C. Enrichments were plated to BGS and XLT-4 agars and transferred to GN (Gram Negative) Hajna and Tetrathionate broths and incubated 24 h and 48 h, respectively at 37°C. Afterwards, 100 µl of GN Hajna and Tetrathionate broth Broths were then transferred to Rappaport-Vassiliadis R10 (RV) broth (BD; Franklin Lakes, NJ) and incubated at 37ºC for 18-24 h. Thereafter, 10 µl of RV broth was plated to both BGS and XLT-4 agars. Isolated colonies characteristic of *Salmonella* were struck onto triple sugar iron and lysine iron agar slants and incubated at 37ºC for 18-24 h. Presumptive *Salmonella* isolates (n = 55) were serogrouped with antisera (Becton Dickinson) and then cryopreserved.

### 5.6.4 Antibiotic susceptibility testing

Antimicrobial susceptibility of *Salmonella* isolates was determined using the Sensititre™ semi-automated system (Thermo Fisher Scientific, Kansas City, KS) according to manufacturer's instructions. Briefly, bacterial suspensions equivalent to a 0.5 McFarland standard were

prepared, aliquoted into a CMV4AGNF panel and incubated at 37°C for 18 h. Minimum inhibitory concentrations were determined and categorized as resistant according to Clinical and Laboratory Standards Institute (CLSI) guidelines when available (CLSI, 2019); otherwise, breakpoints established by the National Antimicrobial Resistance Monitoring System (NARMS) were used (https://www.fda.gov/media/108180/download). Clinical and Laboratory Standards Institute. 2019. Performance standards for antimicrobial susceptibility testing, 30th ed. CLSI Document M100-Ed30. Clinical and Laboratory Standards Institute, Wayne, PA.

### 5.6.5 Whole genome sequencing, processing and taxonomic classification

Select *Salmonella* isolates recovered from litter underwent Illumina short read sequencing and Oxford Nanopore long read sequencing. All DNA extraction was performed using the FastDNA spin kit for soil. Short read libraries were prepared using Nextera XT DNA library preparation kits (Illumina, Inc., San Diego, CA) following the manufacturers protocol. Libraries were sequenced on the MiSeq platform with 250-bp paired end reads. Four isolates underwent sequencing at Novogene (Novogene Co., Ltd., Tianjin) using the Illumina NovaSeq 6000 platform with 150-bp paired end reads. Long read sequencing was conducted at Novogene using the GridION platform (Oxford Nanopore Technology). Sequencing read quality control, adaptor and quality trimming, genome assembly, antimicrobial resistance gene identification, virulence factor identification, plasmid replicon identification, phage region identification and genome annotation were done using Reads2Resistome pipeline v0.0.2 (Woyda, Oladeinde and Abdo, 2023). Isolates selected for both short read and long read sequencing were assembled using both short and long reads using the --hybrid option in the Reads2Resistome pipeline. ResFinder

(Bortolaia *et al.*, 2020) was utilized for annotation of acquired resistance genes. Antimicrobial resistance gene identification was performed using the Resistance Gene Identifier v5.1.1 (RGI) which relies on the Comprehensive Antimicrobial Resistance Database and antibacterial biocide and metal resistance genes were identified with BacMet. MLST was determined using the mlst software (Seemann T, mlst Github https://github.com/tseemann/mlst), which utilizes the PubMLST website (https://pubmlst.org/). Serovar identification was performed using SISTR (Yoshida *et al.*, 2016) via Reads2Resistome. The minimum identity match used for all reference database query hits was 85%. Verification of genes, as well as suspect plasmid replicon-containing contigs was done using Megablast implemented in Geneious Prime version 2022.2.2. *Escherichia coli* UMNK88 plasmid (Genbank accession CP002730.1) was used as a reference for the *fae* operon. Plasmid replicons were identified using the PlasmidFinder database via ABRICATE (Carattoli *et al.*, 2014; Seemann, no date). Additional gene annotation, with available subsystem annotation, was performed using the Rapid Annotation using Subsystem Technology (RAST) (Aziz *et al.*, 2008). Plasmid maps were generated using BLAST Ring Image Generator (Alikhan *et al.*, 2011).

### 5.6.6 Statistical analysis

Antimicrobial resistance genes identified through RGI were filtered to include hits which matched >=95% to the reference database sequence. Genes identified through BacMet and ABRICATE (virulence factors) were filtered to include those matching >=80% to the reference database. A table was generated based on the presence/absence of the identified genes in each isolate from each database. Correspondence analysis on the presence/absence table of antimicrobial resistance genes and virulence factors was conducted in R using the following packages: factoextra v1.0.7, FactoMineR v2.4 and corrplot v0.2-0. Heatmaps were generated

using the pheatmap package in R. A distance matrix was generated using the jaccard metric via the vegdist function from the vegan v2.6-4 package (Jari Oksanen *et al.*, 2022). The distance matrix was used to determine the optimal number of clusters and was implemented using the silhouette method from the fviz_nbclust function from the factoextra v1.0.7 package. Via the pheatmap function, hclust() from the stats v3.6.2 package was then utilized to perform hierarchical clustering using the 'average' (UPGMA) method. All analyses were done in R v4.0.4 utilizing RStudio v1.2.1106. Using the Illumina quality-controlled short-read sequences, taxonomic classification was performed using Kraken2.

## 5.9 References

ADAS (2016) *Comparison of the regulatory framework and key practices in the poultry meat supply chain in the EU and USA: a study for the Association of Poultry Processors and Poultry Trade in the EU Countries.* ADAS, Ltd. Available at: https://britishpoultry.org.uk/identity-cms/wp-content/uploads/2018/05/2016-ADAS-EU-US-comparison.pdf.

Alcock, B.P. *et al.* (2020) 'CARD 2020: Antibiotic resistome surveillance with the comprehensive antibiotic resistance database', *Nucleic Acids Research*, 48(D1), pp. D517–D525. Available at: https://doi.org/10.1093/nar/gkz935.

Alikhan, N.F. *et al.* (2011) 'BLAST Ring Image Generator (BRIG): Simple prokaryote genome comparisons', *BMC Genomics*, 12. Available at: https://doi.org/10.1186/1471-2164-12-402.

Andrews, S.C., Robinson, A.K. and Rodríguez-Quiñones, F. (2003) 'Bacterial iron homeostasis', *FEMS Microbiology Reviews*, 27(2–3), pp. 215–237. Available at: https://doi.org/10.1016/S0168-6445(03)00055-X.

Antunes, P. *et al.* (2016) 'Salmonellosis: the role of poultry meat', *Clinical Microbiology and Infection*, 22(2), pp. 110–121. Available at: https://doi.org/10.1016/j.cmi.2015.12.004.

Aziz, R.K. *et al.* (2008) 'The RAST Server: Rapid Annotations using Subsystems Technology', *BMC Genomics*, 9(1), p. 75. Available at: https://doi.org/10.1186/1471-2164-9-75.

Backhans, A. and Fellström, C. (2012) ‘Rodents on pig and chicken farms – a potential threat to human and animal health’, *Infection Ecology & Epidemiology*, 2(1), p. 17093. Available at: https://doi.org/10.3402/iee.v2i0.17093.

Bearson, B.L. *et al.* (2008) ‘Iron regulated genes of Salmonella enterica serovar Typhimurium in response to norepinephrine and the requirement of fepDGC for norepinephrine-enhanced growth’, *Microbes and Infection*, 10(7), pp. 807–816. Available at: https://doi.org/10.1016/j.micinf.2008.04.011.

Bearson, B.L. *et al.* (2019) ‘Complete Genome Sequence of Multidrug-Resistant Salmonella enterica Serovar I 4,[5],12:i:− 2015 U.S. Pork Outbreak Isolate USDA15WA-1’, *Microbiology Resource Announcements*. Edited by J.C. Dunning Hotopp, 8(40), pp. e00791-19. Available at: https://doi.org/10.1128/MRA.00791-19.

Bortolaia, V. *et al.* (2020) ‘ResFinder 4.0 for predictions of phenotypes from genotypes’, *The Journal of antimicrobial chemotherapy*, 75(12), pp. 3491–3500. Available at: https://doi.org/10.1093/JAC/DKAA345.

Božić, A. *et al.* (2020) ‘Inhibition of Salmonella Binding to Porcine Intestinal Cells by a Wood-Derived Prebiotic’, *Microorganisms*, 8(7), p. 1051. Available at: https://doi.org/10.3390/microorganisms8071051.

Brooks, J.P. *et al.* (2010) ‘Microbial and antibiotic resistant constituents associated with biological aerosols and poultry litter within a commercial poultry house’, *Science of the Total Environment*, 408(20), pp. 4770–4777. Available at: https://doi.org/10.1016/j.scitotenv.2010.06.038.

Bukowski, M. *et al.* (2019) ‘Prevalence of Antibiotic and Heavy Metal Resistance Determinants and Virulence-Related Genetic Elements in Plasmids of Staphylococcus aureus’, *Frontiers in Microbiology*, 10, p. 805. Available at: https://doi.org/10.3389/fmicb.2019.00805.

Carattoli, A. *et al.* (2014) ‘In Silico detection and typing of plasmids using plasmidfinder and plasmid multilocus sequence typing’, *Antimicrobial Agents and Chemotherapy*, 58(7), pp. 3895–3903. Available at: https://doi.org/10.1128/AAC.02412-14.

Dunn, L.L. *et al.* (2022) ‘The prevalence and concentration of Salmonella enterica in poultry litter in the southern United States’, *PLoS ONE*. Edited by S.H. Park, 17(5), p. e0268231. Available at: https://doi.org/10.1371/JOURNAL.PONE.0268231.

Dziewit, L. *et al.* (2015) ‘Diversity and role of plasmids in adaptation of bacteria inhabiting the Lubin copper mine in Poland, an environment rich in heavy metals’, *Frontiers in Microbiology*, 6. Available at: https://doi.org/10.3389/fmicb.2015.00152.

Feng, Y. *et al.* (2020) ‘Characterization and Source Investigation of Multidrug-Resistant *Salmonella* Anatum from a Sustained Outbreak, Taiwan’, *Emerging Infectious Diseases*, 26(12), pp. 2951–2955. Available at: https://doi.org/10.3201/eid2612.200147.

Ferrari, R.G. *et al.* (2019) ‘Worldwide Epidemiology of *Salmonella* Serovars in Animal-Based Foods: a Meta-analysis’, *Applied and Environmental Microbiology*. Edited by J. Björkroth, 85(14), pp. e00591-19. Available at: https://doi.org/10.1128/AEM.00591-19.

Freeman, Z.N., Dorus, S. and Waterfield, N.R. (2013) ‘The KdpD/KdpE Two-Component System: Integrating K+ Homeostasis and Virulence’, *PLOS Pathogens*, 9(3), p. e1003201. Available at: https://doi.org/10.1371/JOURNAL.PPAT.1003201.

Fricke, W.F. *et al.* (2009) ‘Antimicrobial Resistance-Conferring Plasmids with Similarity to Virulence Plasmids from Avian Pathogenic *Escherichia coli* Strains in *Salmonella enterica* Serovar Kentucky Isolates from Poultry’, *Applied and Environmental Microbiology*, 75(18), pp. 5963–5971. Available at: https://doi.org/10.1128/AEM.00786-09.

Gao, R., Wang, L. and Ogunremi, D. (2020) ‘Virulence Determinants of Non-typhoidal *Salmonellae*’, in M. Blumenberg, M. Shaaban, and A. Elgaml (eds) *Microorganisms*. IntechOpen. Available at: https://doi.org/10.5772/intechopen.88904.

Gieraltowski, L. *et al.* (2016) ‘National Outbreak of Multidrug Resistant Salmonella Heidelberg Infections Linked to a Single Poultry Company’, *PLOS ONE*. Edited by P.R. Cleary, 11(9), p. e0162369. Available at: https://doi.org/10.1371/journal.pone.0162369.

Heffernan, E.J. *et al.* (1994) ‘Specificity of the complement resistance and cell association phenotypes encoded by the outer membrane protein genes rck from Salmonella typhimurium and ail from Yersinia enterocolitica’, *Infection and Immunity*, 62(11), pp. 5183–5186. Available at: https://doi.org/10.1128/iai.62.11.5183-5186.1994.

Hughes, D.T. *et al.* (2009) ‘The QseC Adrenergic Signaling Cascade in Enterohemorrhagic E. coli (EHEC)’, *PLOS Pathogens*, 5(8), p. e1000553. Available at: https://doi.org/10.1371/JOURNAL.PPAT.1000553.

Jajere, S.M. (2019) ‘A review of Salmonella enterica with particular focus on the pathogenicity and virulence factors, host specificity and antimicrobial resistance including multidrug resistance’, *Veterinary World*, 12(4), p. 504. Available at: https://doi.org/10.14202/VETWORLD.2019.504-521.

Janakiraman, A. and Slauch, J.M. (2000) ‘The putative iron transport system SitABCD encoded on SPI1 is required for full virulence of Salmonella typhimurium’, *Molecular Microbiology*, 35(5), pp. 1146–1155. Available at: https://doi.org/10.1046/J.1365-2958.2000.01783.X.

Jones, F.T. *et al.* (1991) ‘A Survey of Salmonella Contamination in Modern Broiler Production’, *Journal of Food Protection*, 54(7), pp. 502–507. Available at: https://doi.org/10.4315/0362-028X-54.7.502.

Kelley, T.R. *et al.* (1995) ‘Bacterial Pathogens and Indicators in Poultry Litter during Re-Utilization’, *Journal of Applied Poultry Research*, 4(4), pp. 366–373. Available at: https://doi.org/10.1093/japr/4.4.366.

Kuijpers, A.F.A. *et al.* (2019) ‘Phenotypic prediction: Linking in vitro virulence to the genomics of 59 Salmonella enteric strains’, *Frontiers in Microbiology*, 10(JAN), p. 3182. Available at: https://doi.org/10.3389/FMICB.2018.03182/BIBTEX.

Line, J. (2002) ‘Campylobacter and Salmonella populations associated with chickens raised on acidified litter’, *Poultry Science*, 81(10), pp. 1473–1477. Available at: https://doi.org/10.1093/ps/81.10.1473.

Mezal, E.H. *et al.* (2014) ‘Isolation and molecular characterization of Salmonella enterica serovar Enteritidis from poultry house and clinical samples during 2010’, *Food Microbiology*, 38, pp. 67–74. Available at: https://doi.org/10.1016/j.fm.2013.08.003.

Nairz, M. *et al.* (2015) ‘Iron Regulatory Proteins Mediate Host Resistance to Salmonella Infection’, *Cell Host & Microbe*, 18(2), pp. 254–261. Available at: https://doi.org/10.1016/j.chom.2015.06.017.

Neuert, S. *et al.* (2018) ‘Prediction of Phenotypic Antimicrobial Resistance Profiles From Whole Genome Sequences of Non-typhoidal Salmonella enterica’, *Frontiers in Microbiology*, 9, p. 592. Available at: https://doi.org/10.3389/fmicb.2018.00592.

Oladeinde, A. *et al.* (2019) ‘Horizontal Gene Transfer and Acquired Antibiotic Resistance in Salmonella enterica Serovar Heidelberg following In Vitro Incubation in Broiler Ceca’, *Applied and Environmental Microbiology*, 85(22). Available at: https://doi.org/10.1128/aem.01903-19.

Oladeinde, A. *et al.* (2023) ‘Management and environmental factors influence the prevalence and abundance of food-borne pathogens and commensal bacteria in peanut hull-based broiler litter’, *Poultry Science*, p. 102313. Available at: https://doi.org/10.1016/j.psj.2022.102313.

Pal, C. *et al.* (2014) ‘BacMet: antibacterial biocide and metal resistance genes database’, *Nucleic Acids Research*, 42(D1), pp. D737–D743. Available at: https://doi.org/10.1093/nar/gkt1252.

Pal, C. *et al.* (2015) ‘Co-occurrence of resistance genes to antibiotics, biocides and metals reveals novel insights into their co-selection potential’, *BMC Genomics*, 16(1), p. 964. Available at: https://doi.org/10.1186/s12864-015-2153-5.

Pezza, A. *et al.* (2016) ‘Compartment and signal-specific codependence in the transcriptional control of *Salmonella* periplasmic copper homeostasis’, *Proceedings of the National Academy of Sciences*, 113(41), pp. 11573–11578. Available at: https://doi.org/10.1073/pnas.1603192113.

Pope, M.J. and Cherry, T.E. (2000) ‘An evaluation of the presence of pathogens on broilers raised on poultry litter treatment®-treated litter’, *Poultry Science*, 79(9), pp. 1351–1355. Available at: https://doi.org/10.1093/ps/79.9.1351.

Rensing, C. and Grass, G. (2003) '*Escherichia coli* mechanisms of copper homeostasis in a changing environment', *FEMS Microbiology Reviews*, 27(2–3), pp. 197–213. Available at: https://doi.org/10.1016/S0168-6445(03)00049-4.

Roberts, B.N. *et al.* (2013) 'Spatial and temporal analysis of microbial populations in production broiler house litter in the southeastern United States', *Journal of Applied Poultry Research*, 22(4), pp. 759–770. Available at: https://doi.org/10.3382/japr.2012-00688.

Roll, V.F.B., Dai Prá, M.A. and Roll, A.P. (2011) 'Research on salmonella in broiler litter reused for up to 14 consecutive flocks', *Poultry Science*, 90(10), pp. 2257–2262. Available at: https://doi.org/10.3382/ps.2011-01583.

Rome, K. *et al.* (2018) 'The Two-Component System ZraPSR Is a Novel ESR that Contributes to Intrinsic Antibiotic Tolerance in Escherichia coli', *Journal of Molecular Biology*, 430(24), pp. 4971–4985. Available at: https://doi.org/10.1016/j.jmb.2018.10.021.

Rychlik, I., Gregorova, D. and Hradecka, H. (2006) 'Distribution and function of plasmids in Salmonella enterica', *Veterinary Microbiology*, 112(1), pp. 1–10. Available at: https://doi.org/10.1016/j.vetmic.2005.10.030.

Scallan, E. *et al.* (2011) 'Foodborne illness acquired in the United States–major pathogens', *Emerging infectious diseases*, 17(1), pp. 7–15. Available at: https://doi.org/10.3201/EID1701.P11101.

Scott, A. *et al.* (2018) 'Effect of different levels of copper nanoparticles and copper sulphate on performance, metabolism and blood biochemical profiles in broiler chicken', *Journal of Animal*

*Physiology and Animal Nutrition*, 102(1), pp. e364–e373. Available at: https://doi.org/10.1111/jpn.12754.

Seemann, T. (no date) 'ABRICATE', *Github* [Preprint]. Available at: https://github.com/tseemann/abricate.

Seribelli, A.A. *et al.* (2020) 'Correction: Phenotypic and genotypic characterization of Salmonella Typhimurium isolates from humans and foods in Brazil', *PLOS ONE*, 15(9), p. e0240055. Available at: https://doi.org/10.1371/journal.pone.0240055.

Shang, K., Wei, B. and Kang, M. (2018) 'Distribution and dissemination of antimicrobial-resistant Salmonella in broiler farms with or without enrofloxacin use', *BMC Veterinary Research*, 14(1), pp. 1–14. Available at: https://doi.org/10.1186/s12917-018-1590-1.

Shepherd, M.W. *et al.* (2010) 'Microbiological analysis of composts produced on South Carolina poultry farms', *Journal of applied microbiology*, 108(6), pp. 2067–2076. Available at: https://doi.org/10.1111/J.1365-2672.2009.04610.X.

Siceloff, A.T., Waltman, D. and Shariat, N.W. (2022) 'Regional Salmonella Differences in United States Broiler Production from 2016 to 2020 and the Contribution of Multiserovar Populations to Salmonella Surveillance', *Applied and Environmental Microbiology*, 88(8). Available at: https://doi.org/10.1128/aem.00204-22.

Souza, S.S.R. *et al.* (2022) 'Population analysis of heavy metal and biocide resistance genes in Salmonella enterica from human clinical cases in New Hampshire, United States', *Frontiers in Microbiology*, 13, p. 983083. Available at: https://doi.org/10.3389/fmicb.2022.983083.

Staehlin, B.M. *et al.* (2016) ‘Evolution of a heavy metal homeostasis/resistance island reflects increasing copper stress in Enterobacteria’, *Genome Biology and Evolution*, p. evw031. Available at: https://doi.org/10.1093/gbe/evw031.

United States Food and Drug Administration, Center for Veterinary Medicine (2022) ‘2021 Summary Report On Antimicrobials Sold or Distributed for Use in Food-Producing Animals’, pp. 1–49.

USDA (2022) *Quarterly Sampling Reports on Salmonella and Campylobacter, Quarterly Sampling Reports on Salmonella and Campylobacter*. Available at: https://www.fsis.usda.gov/science-data/data-sets-visualizations/microbiology/microbiological-testing-program-rte-meat-and-7.

USDA, Economic Research Service (2022) *Per capita availability of chicken higher than that of beef since 2010, Per capita availability of chicken higher than that of beef since 2010*.

Vaz, C.S.L. *et al.* (2017) ‘Interventions to reduce the bacterial load in recycled broiler litter’, *Poultry Science*, 96(8), pp. 2587–2594. Available at: https://doi.org/10.3382/ps/pex063.

Velasquez, C.G. *et al.* (2018) ‘Prevalence and antimicrobial resistance patterns of Salmonella isolated from poultry farms in southeastern United States’, *Poultry Science*, 97(6), pp. 2144–2152. Available at: https://doi.org/10.3382/PS/PEX449.

Wernicki, A., Nowaczek, A. and Urban-Chmiel, R. (2017) ‘Bacteriophage therapy to combat bacterial infections in poultry’, *Virology Journal*, 14(1), pp. 1–13. Available at: https://doi.org/10.1186/S12985-017-0849-7/TABLES/2.

World Health Organization, G.W.H. (2019) *WHO | WHO list of Critically Important Antimicrobials (CIA)*. Available at: http://www.who.int/foodborne_disease/resistance/cia/en/#.UiMEZ7zmSDA.mendeley.

Woyda, R., Oladeinde, A. and Abdo, Z. (2023) 'Chicken Production and Human Clinical *Escherichia coli* Isolates Differ in Their Carriage of Antimicrobial Resistance and Virulence Factors', *Applied and Environmental Microbiology*. Edited by M. Vives, pp. e01167-22. Available at: https://doi.org/10.1128/aem.01167-22.

Yoshida, C.E. *et al.* (2016) 'The salmonella in silico typing resource (SISTR): An open web-accessible tool for rapidly typing and subtyping draft salmonella genome assemblies', *PLoS ONE*, 11(1). Available at: https://doi.org/10.1371/journal.pone.0147101.

Zakaria, Z. *et al.* (2022) 'Virulence Gene Profile, Antimicrobial Resistance and Multilocus Sequence Typing of Salmonella enterica Subsp. enterica Serovar Enteritidis from Chickens and Chicken Products', *Animals*, 12(1), p. 97. Available at: https://doi.org/10.3390/ani12010097.

Zhao, L. *et al.* (2010) 'Staphylococcus aureus AI-2 quorum sensing associates with the KdpDE two-component system to regulate capsular polysaccharide synthesis and virulence', *Infection and Immunity*, 78(8), pp. 3506–3515. Available at: https://doi.org/10.1128/IAI.00131-10/SUPPL_FILE/SUPPLEMENTARY_LEGEND.DOC.

# CHAPTER 6: CONCLUSION

This work has successfully evaluated the presence of pathogenic bacterial species in broiler chicken production in the United States and has provided in-depth genomic characterization regarding antimicrobial resistance, virulence and metal resistance. The novel contributions of this work include: a robust bioinformatics pipeline which provides quality control of input sequencing reads, genome assembly, genome annotation and antimicrobial resistance gene, virulence gene, and prophage characterization in a high-throughput manner (Chapter 2); an evaluation and comparison of the current distribution of antimicrobial resistance genes and virulence factors present in *E. coli* isolates from chicken production and human clinical settings in the United States (Chapter 3); and genomic characterization of the pathogenic bacteria, *Campylobacter* and *Salmonella* (Chapter 4 and Chapter 5, respectively), in commercial scale broiler chicken production as well as elucidation of the influence management practices and environmental factors impose on their identified genetic elements. Overall, this work has shown management practices, such as the addition of copper sulfate to drinking water, may be selecting, and providing a reservoir, for bacteria harboring copper resistance on plasmids. Additionally, this work has led to the proposal of the hypothesis that the reused litter environment, including both microbial and physiochemical factors, may be capable of imposing selective pressures on virulence factors relating to host colonization.

Much of the work characterizing pathogenic bacterial species in poultry production has been focused on post-harvest isolates and those studying pre-harvest isolates have lacked in-depth genomic characterization. Enumeration of bacterial species through microbiological

methods give insight into the presence and persistence of species in environmental samples. The impact of management practices and environmental factors on bacterial populations can be elucidated through methods such as direct plating tools however, these methods lack the ability to ask what genetic elements are contributing to bacterial fitness. Next generation sequencing and high-performance computing have enabled researchers to rapidly characterize bacterial genomes at scale. Chapter 2 presents a bioinformatics pipeline which performs assembly of bacterial genomes and focuses on antimicrobial resistance and virulence factor identification.

Antimicrobial resistance is a growing threat to human health and the solution to this problem is still not clear. The removal of antibiotics as growth promotors in broiler chicken production and the rise of programs limiting antibiotic usage, such at the “No Antibiotics Ever” program, have been aimed at reducing the incidence of resistant food-borne pathogens. However, the incidence of drug-resistant food-borne pathogens has remained relatively unchanged in the last 10 years. Thus, alternative management practices need to be, and have been, implemented to reduce the presence of drug-resistant food-borne pathogens. The addition of copper sulfate to drinking water is one such alternative management practice. This practice is often used as a growth promotion method for chickens, or as a method of sanitization of water lines. However, while this practice is meant to reduce bacterial populations, in Chapter 5 we observed the persistence of *Salmonella* isolates, over multiple flock cycles, which harbored plasmids relating to survival in copper-rich environments. Additionally, the practice of litter reuse, which is practiced widely in the United States, is banned in Europe and Canada due to the presence of pathogenic bacterial species. However, in Chapter 4, our data resulted in the hypothesis that the litter environment was selecting for *Campylobacter* strains which are less likely to infect chickens and humans. Thus, selective pressures imposed by management practices may not

always align with initial expectations and we must further understand the influence of these practices so we may appropriately use them to our advantage in reducing the incidence of pathogenic, virulent and drug-resistant bacterial species in the food production space.

Overall, this dissertation contributes to the field of food and environmental microbiology by providing methods for robust and high-throughput analysis of bacterial genomes and providing knowledge of an understudied litter medium and insights into the selective pressures imposed by modern-day broiler chicken production management practices. In closing, I hope this work serves as a foundation for further work into elucidating selective pressures on bacterial pathogens imposed by management practices.

# APPENDIX

**Table S2.1.** Bandage graphs for isolates under various assembly conditions.

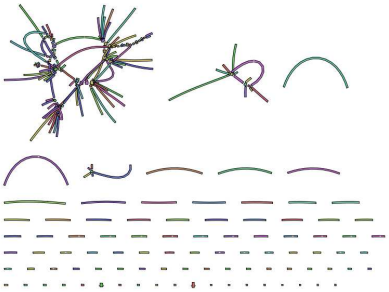
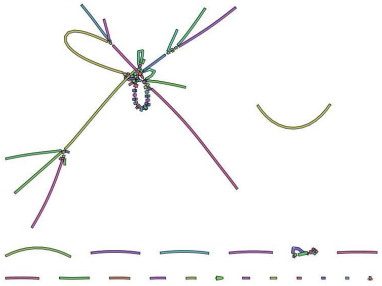
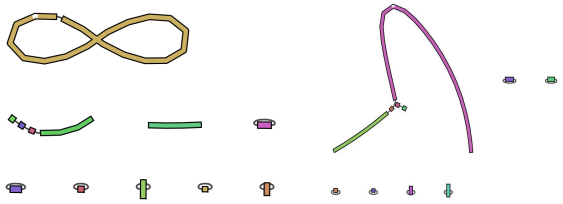
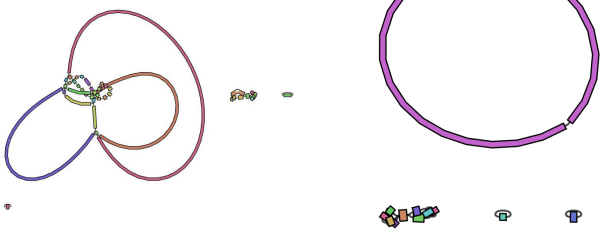
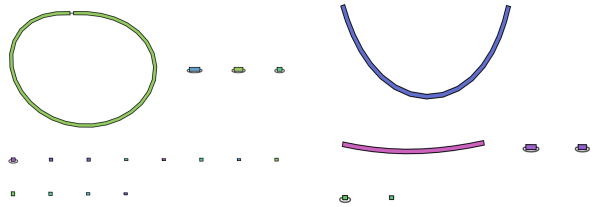
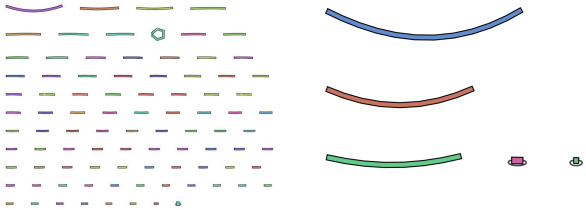
| EC-IC | | SH-IC | |
|---|---|---|---|
| Short read assembly | | | |
|  | |  | |
| Hybrid assembly | | | |
| MinION | PacBio | MinION | PacBio |
|  | |  | |
| Long read assembly | | | |
| MinION | PacBio | MinION | PacBio |
|  | |  | |

**Table S2.2.** Summary of assembly resources and run-time under various assembly conditions.

| Isolates included in run | Assembly method | Elapsed time | --threads (option) | CPU-hours |
|---|---|---|---|---|
| EC-IC Illumina; SH-IC Illumina | Short Read | 16m 13s | 64 | 0.6 |
| EC-IC MinION; EC-IC PacBio; SH-IC MinION; SH-ICPacBio | Hybrid | 4h 31m 45s | 64 | 5.3 |
| EC-IC MinION; EC-IC PacBio; SH-IC MinION; SH-IC PacBio | Long Read | 23m 46s | 64 | 0.9 |

**Table S2.3.** Summary of resistome characterization for assembled isolates under various assembly conditions.

| | | Short Read | Hybrid | | Long Read | |
|---|---|---|---|---|---|---|
| | | Illumina | MinION | PacBio | MinION | PacBio |
| Isolate | Database | Number of identified elements | | | | |
| EC-IC | ARG-ANNOT | 8 | 8 | 8 | 7 | 8 |
| | CARD | 48 | 48 | 48 | 44 | 43 |
| | MEGARes | 58 | 58 | 58 | 53 | 52 |
| | NCBI | 4 | 4 | 4 | 3 | 4 |
| | PlasmidFinder | 7 | 7 | 7 | 5 | 5 |
| | ResFinder | 4 | 4 | 4 | 3 | 4 |
| | VirulenceFinder | 70 | 72 | 72 | 55 | 55 |
| SH-IC | ARG-ANNOT | 8 | 8 | 8 | 6 | 5 |
| | CARD | 29 | 30 | 30 | 6 | 9 |
| | MEGARes | 36 | 37 | 37 | 6 | 16 |
| | NCBI | 4 | 4 | 4 | 4 | 3 |
| | PlasmidFinder | 3 | 3 | 3 | 2 | 2 |
| | ResFinder | 5 | 5 | 5 | 5 | 4 |
| | VirulenceFind | 105 | 105 | 106 | 44 | 99 |

| | |
|---|---|
| | er |

**Table S2.4.** Summary of identified bacteriophages under various assembly conditions.

| | Short Read | Hybrid | | Long Read | |
|---|---|---|---|---|---|
| | Illumina | MinION | PacBio | MinION | PacBio |
| Isolate | Bacteriophage family regions identified | | | | |
| EC-IC | *Siphoviridae*<br>*Myoviridae*<br>*Myoviridae* | *Siphoviridae*<br>*Siphoviridae*<br>*Siphoviridae*<br>*Myoviridae*<br>*Myoviridae*<br>Unknown | *Siphoviridae*<br>*Siphoviridae*<br>*Siphoviridae*<br>*Myoviridae*<br>*Myoviridae*<br>Unknown | *Siphoviridae* | None |
| SH-IC | *Siphoviridae*<br>*Podoviridae*<br>*Myoviridae*<br>Unknown | *Siphoviridae*<br>*Podoviridae*<br>*Myoviridae*<br>Unknown | *Siphoviridae*<br>*Podoviridae*<br>*Myoviridae*<br>Unknown | None | None |

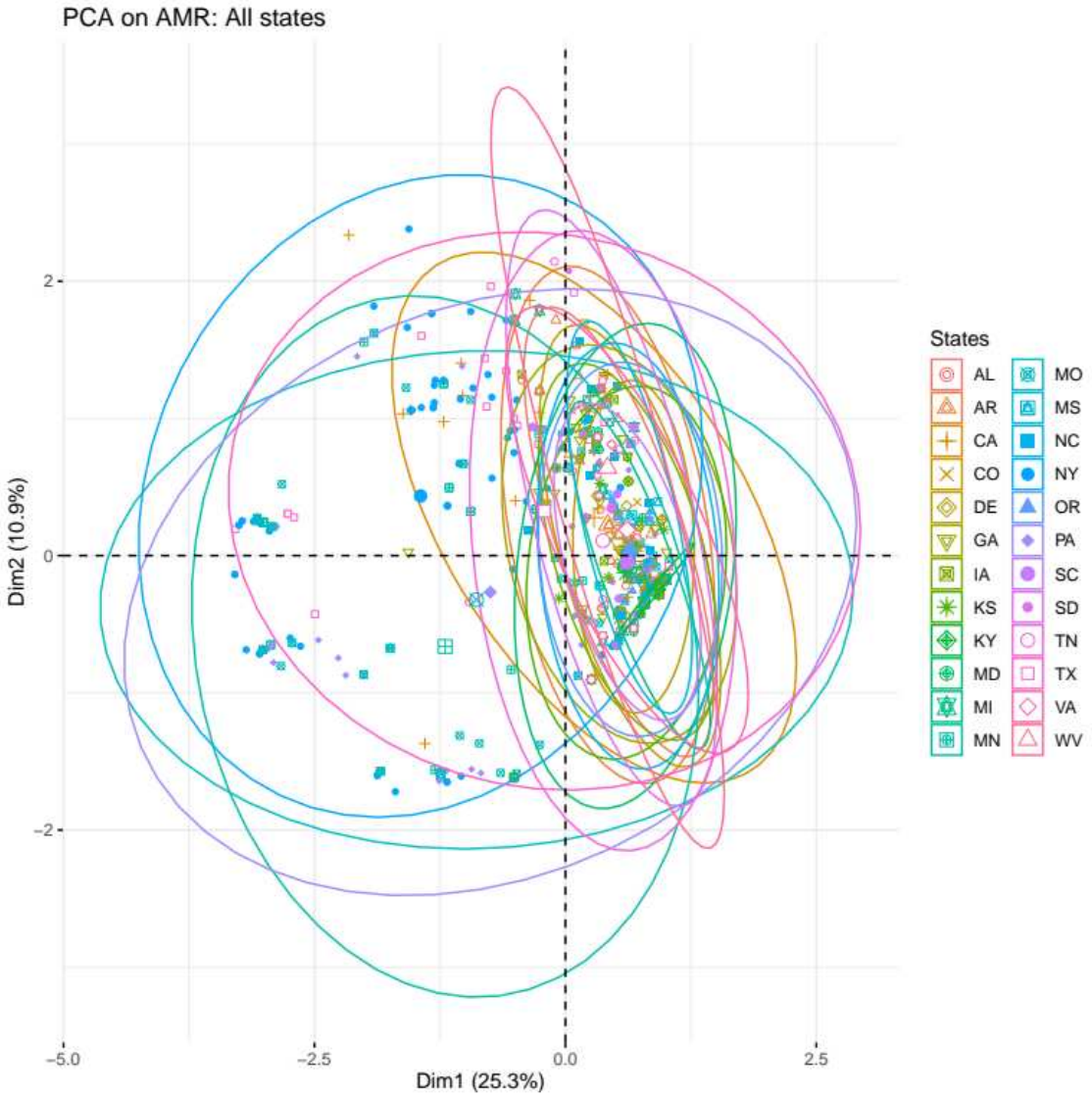



**Figure S3.1.** Principal component analysis of identified antimicrobial resistance genes from RGI and ResFinder. Isolates are labeled corresponding to their state origin.

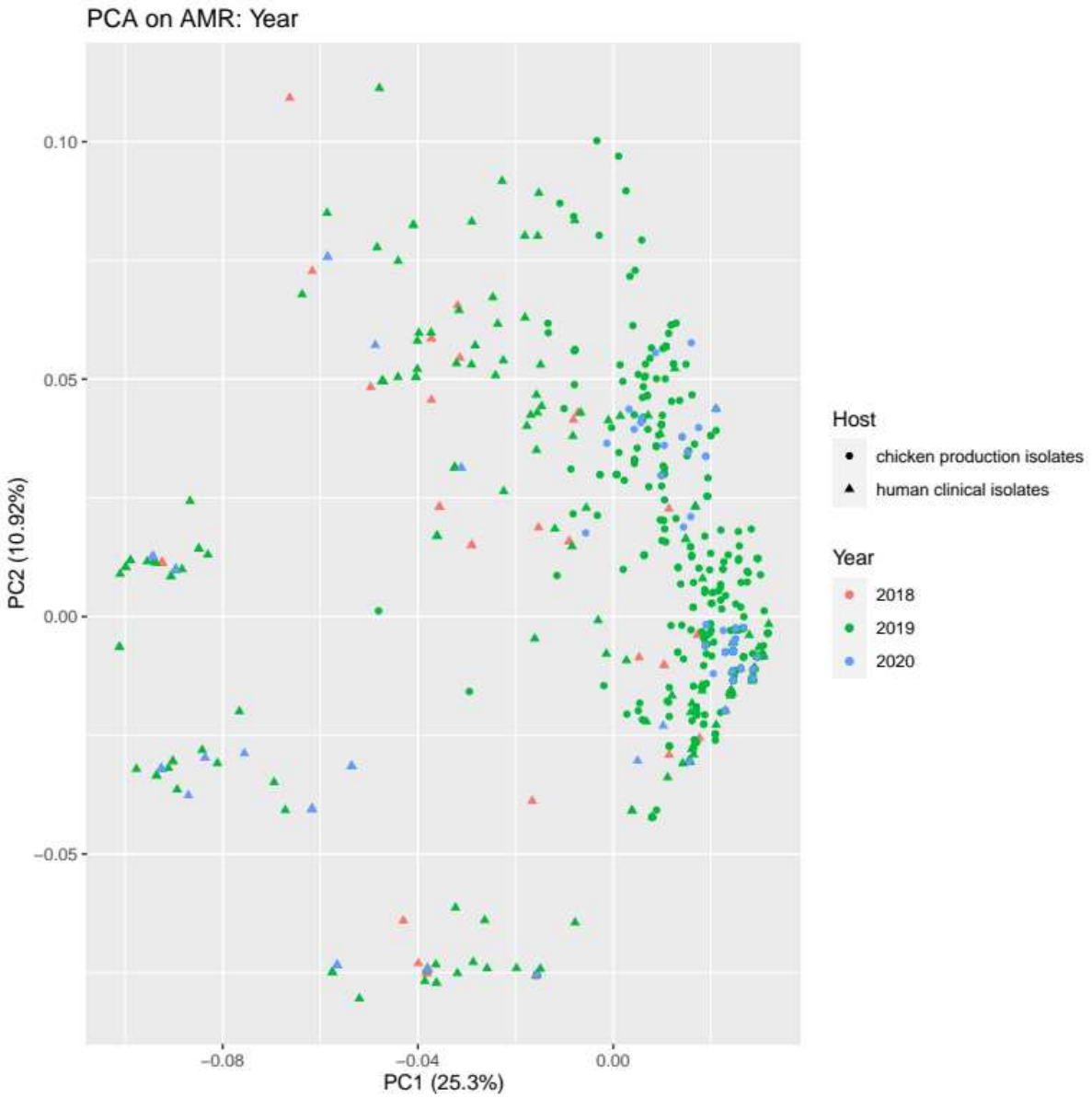



**Figure S3.2.** Principal component analysis of identified antimicrobial resistance genes from RGI and ResFinder. Isolates are labeled with the corresponding sampling year.

**Table S3.2.** Drug classes, identified through RGI (CARD) output of AMR genes, which significantly differed in proportion between isolates from the various isolation years as determined by the Wilcoxon rank-sum test. Genes conferring resistance to drug classes were enumerated for each isolate and a proportion was calculated using the total number of genes in the study population conferring resistance to a given drug class. P value adjustment performed by the Benjamini-Hochberg false discovery rate correction.

| **2018 vs 2019** | | **2018 vs 2020** | | **2019 vs 2020** | |
|---|---|---|---|---|---|
| **Drug class** | **Adjusted p value** | **Drug class** | **Adjusted p value** | **Drug class** | **Adjusted p value** |
| fluoroquinolone | 3E-12 | aminoglycoside | 5.7E-07 | cephamycin | 3.6E-08 |
| macrolide | 7.3E-07 | sulfonamide | 0.0002 | fosfomycin | 2.7E-07 |
| cephalosporin | 0.00027 | acridinedye | 0.00042 | aminoglycoside | 3.6E-06 |
| cephamycin | 0.003 | penam | 0.0012 | diaminopyrimidine | 6.1E-05 |
| penem | 0.013 | tetracycline | 0.002 | macrolide | 0.00028 |
| monobactam | 0.056 | fosfomycin | 0.11 | fluoroquinolone | 0.00041 |
| tetracycline | 0.071 | diaminopyrimidine | 0.46 | penam | 0.0057 |
| peptide | 0.21 | fluoroquinolone | 1 | sulfonamide | 0.0077 |
| acridinedye | 0.42 | cephalosporin | 1 | acridinedye | 0.012 |
| sulfonamide | 0.52 | glycylcycline | 1 | cephalosporin | 0.18 |
| oxazolidinone | 0.74 | phenicol | 1 | glycylcycline | 1 |

| | | | | | |
|---|---|---|---|---|---|
| aminoglycoside | 0.85 | rifamycin | 1 | phenicol | 1 |
| glycylcycline | 1 | triclosan | 1 | rifamycin | 1 |
| penam | 1 | cephamycin | 1 | tetracycline | 1 |
| phenicol | 1 | peptide | 1 | triclosan | 1 |
| rifamycin | 1 | aminocoumarin | 1 | peptide | 1 |
| triclosan | 1 | glycopeptide | 1 | aminocoumarin | 1 |
| aminocoumarin | 1 | macrolide | 1 | glycopeptide | 1 |
| glycopeptide | 1 | lincosamide | 1 | lincosamide | 1 |
| diaminopyrimidine | 1 | streptogramin | 1 | streptogramin | 1 |
| lincosamide | 1 | nitrofuran | 1 | elfamycin | 1 |
| streptogramin | 1 | carbapenem | 1 | benzalkoniumchloride | 1 |
| fosfomycin | 1 | monobactam | 1 | rhodamine | 1 |
| elfamycin | 1 | penem | 1 | nitrofuran | 1 |
| benzalkoniumchloride | 1 | nucleoside | 1 | carbapenem | 1 |
| rhodamine | 1 | oxazolidinone | 1 | monobactam | 1 |
| nitrofuran | 1 | pleuromutilin | 1 | penem | 1 |
| carbapenem | 1 | elfamycin | NaN | nucleoside | 1 |

| | | | | | |
|---|---|---|---|---|---|
| nucleoside | 1 | benzalkoniumchloride | NaN | pleuromutilin | 1 |
| pleuromutilin | 1 | rhodamine | NaN | nitroimidazole | NaN |
| nitroimidazole | NaN | nitroimidazole | NaN | oxazolidinone | NaN |

**Table S3.3.** Drug classes, identified through the ResFinder output of acquired AMR genes, which significantly differed in proportion between human clinical isolates and chicken production isolates as determined by the Wilcoxon rank-sum test. Chicken production isolate average proportions were compared against human clinical isolates. Genes conferring resistance to drug classes were enumerated for each isolate and a proportion was calculated using the total number of genes in the study population conferring resistance to a given drug class. P value adjustment performed by the Benjamini-Hochberg false discovery rate correction.

| **Drug Class** | **Adjusted p value** |
|---|---:|
| cephalosporin | 1.99E-20 |
| penam | 6.33E-20 |
| penem | 5.85E-17 |
| monobactam | 5.85E-17 |
| diaminopyrimidine antibiotic | 3.83E-12 |
| macrolide antibiotic | 6.94E-09 |
| disinfecting agents and intercalating dyes | 5.70E-08 |
| sulfonamide antibiotic | 0.00033 |
| carbapenem | 0.0019 |
| fluoroquinolone antibiotic | 0.015 |
| phenicol | 0.042 |
| aminoglycoside antibiotic | 0.49 |
| tetracycline antibiotic | 0.85 |
| cephamycin | 1 |
| streptogramin antibiotic | 1 |
| lincosamide antibiotic | 1 |
| disinfecting agents | 1 |
| fosfomycin | 1 |
| peptide antibiotic | 1 |
| benzalkonium chloride | 1 |
| rhodamine | 1 |

**Table S3.4.** Drug classes, identified through RGI (CARD) output of AMR genes, which significantly differed in proportion between human clinical isolates and chicken production isolates as determined by the Wilcoxon rank-sum test. Chicken production isolate average proportions were compared against human clinical isolates. Genes conferring resistance to drug classes were enumerated for each isolate and a proportion was calculated using the total number of genes in the study population conferring resistance to a given drug class. P value adjustment performed by the Benjamini-Hochberg false discovery rate correction.

| **Drug Class** | **Adjusted p value** |
|---|---|
| diaminopyrimidine | 1.53E-42 |
| fluoroquinolone | 1.75E-41 |
| macrolide | 8.25E-32 |
| cephamycin | 1.53E-27 |
| monobactam | 1.49E-18 |
| penem | 9.27E-18 |
| fosfomycin | 6.42E-17 |
| aminoglycoside | 6.76E-11 |
| sulfonamide | 2.84E-09 |
| elfamycin | 1.10E-05 |
| phenicol | 0.00027 |
| peptide | 0.00029 |
| nucleoside | 0.0069 |
| streptogramin | 0.015 |
| rifamycin | 0.21 |

| | |
|---|---|
| lincosamide | 0.24 |
| tetracycline | 0.39 |
| aminocoumarin | 0.44 |
| cephalosporin | 1 |
| glycylcycline | 1 |
| penam | 1 |
| triclosan | 1 |
| glycopeptide | 1 |
| benzalkoniumchloride | 1 |
| rhodamine | 1 |
| nitrofuran | 1 |
| carbapenem | 1 |
| acridinedye | 1 |
| oxazolidinone | 1 |
| pleuromutilin | 1 |
| nitroimidazole | NaN |

**Table S3.6.** Confusion matrix resulting from Random Forests classification on filtered presence/absence table of identified AMR and virulence factors (**Table S3.1**). All 791 chicken production and human clinical isolates were reclassified into their respective clusters which were identified through hierarchical clustering. Rows indicate reclassified isolates and columns original hierarchical cluster.

| Hierarchical Cluster | A | B | Error |
|---|---|---|---|
| **A** | **644** | **0** | **0** |
| **B** | **0** | **130** | **0** |

**Table S3.7.** RGI-identified antimicrobial resistance genes with mutations conferring resistance.

| Gene w/mutation | Mutation |
|---|---|
| Escherichia coli EF-Tu mutants conferring resistance to Pulvomycin | R234F |
| Escherichia coli parC conferring resistance to fluoroquinolones | S80I |
| Escherichia coli UhpT with mutation conferring resistance to fosfomycin | E350Q |
| Escherichia coli PtsI with mutation conferring resistance to fosfomycin | V25 I |
| Escherichia coli gyrA conferring resistance to fluoroquinolones | S83L (30 isolates)<br>S84L (2 isolates)<br>D87Y (275 isolates) |
| Escherichia coli GlpT with mutation conferring resistance to fosfomycin | E448K |
| Haemophilus influenzae PBP3 conferring resistance to beta-lactam antibiotics | D350N |
| Escherichia coli gyrA conferring resistance to triclosan | D87G |
| Escherichia coli cyaA with mutation conferring resistance to fosfomycin | S352T |
| Escherichia coli marR mutant conferring antibiotic resistance (fluoroquinolone antibiotic; cephalosporin; glycylcycline; penam; tetracycline antibiotic; rifamycin antibiotic; phenicol antibiotic; triclosan) | Y137H |
| Escherichia coli soxR with mutation conferring antibiotic resistance (fluoroquinolone antibiotic; cephalosporin; glycylcycline; penam; tetracycline antibiotic; rifamycin antibiotic; phenicol antibiotic; triclosan) | R20H and G121D |
| Escherichia coli soxS with mutation conferring antibiotic resistance (fluoroquinolone antibiotic; monobactam; carbapenem; cephalosporin; | n/a |

| | |
|---|---|
| glycylcycline; cephamycin; penam; tetracycline antibiotic; rifamycin antibiotic; phenicol antibiotic; triclosan; penem) | |
| Escherichia coli acrR with mutation conferring multidrug antibiotic resistance (fluoroquinolone antibiotic; cephalosporin; glycylcycline; penam; tetracycline antibiotic; rifamycin antibiotic; phenicol antibiotic; triclosan) | n/a |

**Table S3.8.** Confusion matrix resulting from Random Forests classification on filtered presence/absence table of identified AMR and virulence factors (**Table S3.1**). All 791 chicken production and human clinical isolates were reclassified into their respective host sources. Rows indicate reclassified isolates, columns original host, and red numbers indicate misclassified.

| Host | Chicken Production | Human Clinical | Error |
|---|---|---|---|
| **Chicken Production** | **447** | **5** | **0.011** |
| **Human Clinical** | **19** | **323** | **0.0556** |

**Table S3.9.** Random Forests classification misclassified host results. All 791 chicken production and human clinical isolates were reclassified into their respective Host sources. Random Forests classification was performed on the filtered presence/absence table of identified AMR and virulence factors (**Table S3.1**).

| **Random Forests Host Predictions** | | | | | |
|---|---|---|---|---|---|
| **Isolate** | **Predicted Host** | **Actual Host** | **Isolation source** | **Phylogroup** | **Hierarchical Cluster** |
| SRR10687714 | Human | Chicken Pro. | Chicken Thighs | F | A |
| SRR10687982 | Human | Chicken Pro. | Chicken Wings | A | A |
| SRR8590847 | Human | Chicken Pro. | Chicken | B2 | A |
| SRR9984631 | Human | Chicken Pro. | Chicken Wings | B1 | A |
| SRR9984720 | Human | Chicken Pro. | Chicken Wings | B1 | A |
| SAMN09981238 | Chicken Pro. | Human | Feces | A | A |
| SAMN10620159 | Chicken Pro. | Human | Sepsis | A | A |
| SAMN10722958 | Chicken Pro. | Human | Rectal | A | A |
| SAMN10722963 | Chicken Pro. | Human | Rectal | A | A |
| SAMN10722966 | Chicken Pro. | Human | Rectal | A | A |
| SAMN09981239 | Chicken Pro. | Human | Feces | B1 | A |
| SAMN09981247 | Chicken Pro. | Human | Feces | B1 | A |

| | | | | | |
|---|---|---|---|---|---|
| SAMN09981249 | Chicken Pro. | Human | Feces | B1 | A |
| SAMN09981263 | Chicken Pro. | Human | Feces | B1 | A |
| SAMN09981265 | Chicken Pro. | Human | Feces | B1 | A |
| SRR10728218 | Chicken Pro. | Human | Blood | B1 | A |
| SAMN09981305 | Chicken Pro. | Human | Feces | B1 | A |
| SAMN09981307 | Chicken Pro. | Human | Feces | B1 | A |
| SAMN09981302 | Chicken Pro. | Human | Feces | B2 | A |
| SAMN10722961 | Chicken Pro. | Human | Rectal | B2 | A |
| SAMN10620143 | Chicken Pro. | Human | Sepsis | D | A |
| SAMN09981237 | Chicken Pro. | Human | Feces | D | A |
| SRR6892704 | Chicken Pro. | Human | Feces | D | A |
| SRR6892705 | Chicken Pro. | Human | Feces | D | A |

**Table S3.10.** Confusion matrix resulting from Random Forests classification on filtered presence/absence table of identified AMR and virulence factors (**Table S3.1**). All 791 chicken production and human clinical isolates were reclassified into their respective ClermonTyper identified phylogroups sources. Rows indicate reclassified isolates, columns original host, and red numbers indicate misclassified.

| Phylogroup | A | B1 | B2 | C | clade I | D | E | E or clade I | F | G | Unknown | Error |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **A** | **117** | **14** | **0** | **0** | **0** | **1** | **0** | **0** | **0** | **0** | **0** | **0.11** |
| **B1** | **5** | **177** | **0** | **1** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0.03** |
| **B2** | **0** | **0** | **184** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** |
| **C** | **0** | **11** | **0** | **12** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0.48** |
| **clade I** | **0** | **0** | **0** | **0** | **8** | **0** | **0** | **0** | **0** | **0** | **0** | **0** |
| **D** | **0** | **0** | **0** | **0** | **0** | **103** | **0** | **0** | **0** | **0** | **0** | **0** |
| **E** | **0** | **0** | **0** | **0** | **0** | **4** | **71** | **0** | **0** | **0** | **0** | **0.05** |
| **E or clade I** | **0** | **0** | **0** | **0** | **0** | **3** | **1** | **0** | **0** | **0** | **0** | **1** |
| **F** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **25** | **1** | **0** | **0.04** |
| **G** | **0** | **0** | **0** | **0** | **0** | **1** | **0** | **0** | **0** | **54** | **0** | **0.02** |
| **Unknown** | **1** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **1** |

**Table S3.11.** Virulence factor-associated functions which significantly differed in proportion between human clinical isolates and chicken production isolates as determined by the Wilcoxon rank-sum test. Virulence genes associated with each function were enumerated for each isolate and a proportion was calculated using the total number of genes in the study population with the given function. P value adjustment performed by the Benjamini-Hochberg false discovery rate correction.

| **VFDB Fcn** | **Adjusted p value** |
|---|---|
| Autotransporter | 7.95E-42 |
| Toxin | 2.54E-28 |
| Adherence | 1.08E-23 |
| Invasion | 6.34E-23 |
| LEE-encoded-TTSS-effectors | 1.79E-12 |
| Iron.uptake | 5.33E-11 |
| Non-LEE-encoded-TTSS-effectors | 6.53E-09 |
| Secretion system | 0.00029 |
| Protease | 0.051 |
| Immune evasion | 0.067 |
| Fimbrial adherence determinants | 0.078 |

**Table S3.12.** Proportion (%) of virulence factor-associated functions across identified phylogroups. The set of functions for each gene was counted and summed for all isolates in a given phylogroup. Virulence genes associated with each function were enumerated for each isolate and a proportion was calculated using the total number of genes in the study population with the given function.

| Phylogroup | Adherence | Protease | Autotransporter | Invasion | Iron uptake | LEE-encoded T3SS effectors | Non-LEE encoded T3SS effectors | Secretion system | Toxin | Immune evasion | Fimbrial adherence determinants |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **A** | 13.41 | 1.22 | 0.04 | 11.45 | 22.31 | 8.72 | 10.31 | 6.47 | 1.59 | 5.6 | 55.94 |
| **B1** | 14.93 | 3.06 | 0.11 | 0.36 | 24.15 | 6.32 | 10.22 | 7.34 | 3.57 | 2.37 | 57.14 |
| **B2** | 19.69 | 0.55 | 4.4 | 35.59 | 45.98 | 0.86 | 0.19 | 5.42 | 5.75 | 2.2 | 56.99 |
| **C** | 16.73 | 0 | 0 | 0 | 33.75 | 0 | 7 | 4.93 | 0.18 | 0 | 57.14 |
| **clade I** | 16.63 | 0 | 0 | 36.11 | 32.5 | 0 | 15.51 | 6.32 | 3.13 | 0 | 57.14 |
| **D** | 16.82 | 0 | 0.83 | 24.49 | 38.57 | 6.66 | 17.76 | 6.76 | 2.59 | 3.56 | 54.51 |
| **E** | 17.81 | 16 | 3.93 | 12.3 | 27.75 | 67.05 | 47.06 | 14.58 | 24.22 | 0.89 | 56.95 |
| **E or clade I** | 15.63 | 0 | 0 | 11.11 | 31.25 | 32.14 | 20.37 | 15.99 | 4.17 | 0 | 53.57 |
| **F** | 16.83 | 0.77 | 0.77 | 33.76 | 42.16 | 0 | 11.61 | 6.13 | 1.6 | 0 | 57.14 |
| **G** | 16.83 | 0.77 | 0.77 | 33.76 | 42.16 | 0 | 11.61 | 6.13 | 1.6 | 0 | 57.14 |
| **Unknown** | 7.14 | 0 | 0 | 11.11 | 36.25 | 0 | 12.96 | 6.4 | 4.17 | 0 | 57.14 |

**Table S4.2.** Virulence factor-associated functions across all *C. jejuni* and *C. coli* isolated. Virulence factor identification was performed with ABRICATE which utilized the Virulence Factor Database (VFDB). For each identified virulence gene, the associated function(s) were enumerated.

| Virulence Factor Function | Count |
|---|---|
| Motility and export apparatus | 1623 |
| Colonization and Immune evasion | 779 |
| Glycosylation system | 436 |
| Immune evasion | 340 |
| Toxin | 312 |
| Adherence | 302 |
| Invasion | 295 |
| Chemotaxis and motility | 220 |
| Secretion system | 44 |
| capsule | 44 |

**Table S4.3.** Comparison of average proportions of virulence factor-associated functions for *C. jejuni* isolates which significantly differed in proportion between broiler houses 1 and 2 as determined by the Wilcoxon rank-sum test. Virulence genes associated with each function were enumerated for each isolate and a proportion was calculated using the total number of genes in the study population with the given function. Adjusted p value adjustment was performed by the Benjamini-Hochberg false discovery rate correction method. 'NaN' values are due to the inability to compute p values due to average proportion values being identical for all isolates.

| **Wilcoxon rank-sum test:house 1 vs house 2 (*C. jejuni*)** | |
|---|---|
| **Function** | **adjusted p value** |
| Adherence | 1 |
| Colonization and Immune evasion | 1 |
| Capsule | NaN |
| Chemotaxis and motility | NaN |
| Glycosylation system | NaN |
| Immune evasion | NaN |
| Invasion | NaN |
| Motility and export apparatus | NaN |
| Secretion system | NaN |
| Toxin | NaN |

**Table S4.4.** Comparison of average proportions of virulence factor-associated functions for *C. jejuni* isolates which significantly differed in proportion between broiler houses 2 and 4 as determined by the Wilcoxon rank-sum test. Virulence genes associated with each function were enumerated for each isolate and a proportion was calculated using the total number of genes in the study population with the given function. Adjusted p value adjustment was performed by the Benjamini-Hochberg false discovery rate correction method. 'NaN' values are due to the inability to compute p values due to average proportion values being identical for all isolates.

| **Wilcoxon rank-sum test: house 2 vs house 4 (*C. jejuni*)** | |
|---|---|
| **Function** | **adjusted p value** |
| Immune evasion | 0.16 |
| Adherence | 0.30 |
| Colonization and Immune evasion | 0.37 |
| Glycosylation system | 0.49 |
| Invasion | 1 |
| Motility and export apparatus | 1 |
| Toxin | 1 |
| Capsule | NaN |
| Chemotaxis and motility | NaN |
| Secretion system | NaN |

**Table S4.5.** Comparison of average proportions of virulence factor-associated functions for *C. jejuni* isolates which significantly differed in proportion between broiler houses 1 and 4 as determined by the Wilcoxon rank-sum test. Virulence genes associated with each function were enumerated for each isolate and a proportion was calculated using the total number of genes in the study population with the given function. Adjusted p value adjustment was performed by the Benjamini-Hochberg false discovery rate correction method. 'NaN' values are due to the inability to compute p values due to average proportion values being identical for all isolates.

| **Wilcoxon rank-sum test: house 1 vs house 4 (*C. jejuni*)** | |
|---|---|
| **Function** | **adjusted p value** |
| Immune evasion | 0.17 |
| Colonization and Immune evasion | 0.39 |
| Glycosylation system | 0.87 |
| Invasion | 1 |
| Motility and export apparatus | 1 |
| Toxin | 1 |
| Adherence | 1 |
| Capsule | NaN |
| Chemotaxis and motility | NaN |
| Secretion system | NaN |

**Table S4.6.** Comparison of average proportions of virulence factor-associated functions for *C. jejuni* isolates which significantly differed in proportion between broiler houses 3 and 4 as determined by the Wilcoxon rank-sum test. Virulence genes associated with each function were enumerated for each isolate and a proportion was calculated using the total number of genes in the study population with the given function. Adjusted p value adjustment was performed by the Benjamini-Hochberg false discovery rate correction method. 'NaN' values are due to the inability to compute p values due to average proportion values being identical for all isolates.

| **Wilcoxon rank-sum test: house 3 vs house 4 (*C. jejuni*)** | |
|---|---|
| **Function** | **adjusted p value** |
| Colonization and Immune evasion | 1 |
| Glycosylation system | 1 |
| Immune evasion | 1 |
| Invasion | 1 |
| Motility and export apparatus | 1 |
| Toxin | 1 |
| Adherence | NaN |
| Capsule | NaN |
| Chemotaxis and motility | NaN |
| Secretion system | NaN |

**Table S4.7.** Comparison of virulence factor-associated functions for *C. coli* isolates from flock 1 and flock 2. Comparison was performed by the Wilcoxon rank-sum test. Virulence genes associated with each function were enumerated for each isolate and a proportion was calculated using the total number of genes in the study population with the given function. Adjusted p value adjustment was performed by the Benjamini-Hochberg false discovery rate correction method. 'NaN' values are due to the inability to compute p values due to average proportion values being identical for all isolates.

| **Wilcoxon rank-sum test: flock 1 vs flock 2 (*C. coli:* houses 2,3,4)** | |
|---|---|
| **Function** | **adjusted p value** |
| Colonization and Immune evasion | 1 |
| Glycosylation system | 1 |
| Immune evasion | 1 |
| Invasion | 1 |
| Toxin | 1 |
| Adherence | NaN |
| Capsule | NaN |
| Chemotaxis and motility | NaN |
| Motility and export apparatus | NaN |
| Secretion system | NaN |

**Table S4.8.** Comparison of virulence factor-associated functions for *C. jejuni* isolates from flock 2 and flock 3. Comparison was performed by the Wilcoxon rank-sum test. Virulence genes associated with each function were enumerated for each isolate and a proportion was calculated using the total number of genes in the study population with the given function. Adjusted p value adjustment was performed by the Benjamini-Hochberg false discovery rate correction method. 'NaN' values are due to the inability to compute p values due to average proportion values being identical for all isolates.

| **Wilcoxon by flock 2 vs flock 3 C. jejuni (house 1, 2 and 4)** | |
|---|---|
| **Function** | **adjusted p value** |
| Immune evasion | 0.63 |
| Colonization and Immune evasion | 0.63 |
| Adherence | 1 |
| Invasion | 1 |
| Glycosylation system | 1 |
| Secretion system | NaN |
| Toxin | NaN |
| Capsule | NaN |
| Motility and export apparatus | NaN |
| Chemotaxis and motility | NaN |

**Table S4.9:** Virulence factors and functions absent from grow-out cycle 3 isolates.

| Virulence Factor | Function | Function |
|---|---|---|
| ArsP | arsenic resistance | arsenic resistance |
| acr3 | arsenic resistance | arsenic resistance |
| fcl | GDP-L-fucose synthetase | related to capsule - virulence |
| kfiD | UDP-glucose 6-dehydrogenase | related to capsule - virulence |
| PseE.maf5 | Glycosylation system | N-linked protein glycosylation, mutation results in non-motile phenotype |
| glf | UDP-galactopyranose mutase | Galactose metabolism, Amino sugar and nucleotide sugar metabolism |
| Cj1440c | Colonization and Immune evasion | Capsule biosynthesis and transport |
| Cj1432c | Colonization and Immune evasion | Capsule biosynthesis and transport |
| Cj1421c | Colonization and Immune evasion | Capsule biosynthesis and transport |
| Cj1438c* | Colonization and Immune evasion | Capsule biosynthesis and transport |
| Cj1422c* | Colonization and Immune evasion | Capsule biosynthesis and transport |
| Cj1437c* | Colonization and Immune evasion | Capsule biosynthesis and transport |
| Cj1436c* | Colonization and Immune evasion | Capsule biosynthesis and transport |
| Cj1435c* | Colonization and Immune evasion | Capsule biosynthesis and transport |

* indicates genes absent from grow-out cycle 2

**Table S4.10.** Megablast results for *Campyloabcter coli* isolates against the NZ_CP014743 T4SS.

| Isolate ID | Flock cohort | % Pairwise Identity | Query coverage | E Value | Max Sequence Length |
|---|---|---|---|---|---|
| CC32 | 2 | 97.5% | 61.78% | 0 | 12354 |
| CC34 | 2 | 97.5% | 61.45% | 0 | 12288 |
| CC36 | 2 | 96.5% | 37.07% | 0 | 7414 |
| CC32 | 2 | 99.5% | 34.87% | 0 | 6974 |
| CC36 | 2 | 99.5% | 34.87% | 0 | 6974 |
| CC33 | 2 | 95.1% | 24.70% | 0 | 4940 |
| CC36 | 2 | 99.1% | 24.05% | 0 | 4810 |
| CC34 | 2 | 99.7% | 17.77% | 0 | 3553 |
| CC34 | 2 | 99.3% | 16.47% | 0 | 3294 |
| CC33 | 2 | 99.2% | 11.48% | 0 | 2296 |

*T4SS query length 21,526 bp from the *Campylobacter* megaplasmid NZ_CP014743 (50,831bp - 72,356bp). Results with a query coverage cutoff <10% were excluded. Isolates with multiple row entries are representative of hits on separate contigs.

**Table S4.11.** Megablast results for *Campyloabcter jejuni* isolates against the NZ_CP014743 T6SS.

| Isolate ID | Flock cohort | % Pairwise Identity | Query coverage | E Value | Max Sequence Length |
|---|---|---|---|---|---|
| CJ15 | 1 | 99.3% | 95.13% | 0 | 16171 |
| CJ23 | 1 | 99.3% | 95.13% | 0 | 16171 |
| CJ29 | 1 | 99.3% | 95.13% | 0 | 16171 |
| CJ4 | 1 | 99.3% | 95.13% | 0 | 16171 |
| CJ6 | 1 | 99.3% | 95.13% | 0 | 16171 |
| CJ7 | 1 | 99.3% | 95.13% | 0 | 16171 |
| CH22 | 1 | 99.4% | 85.14% | 0 | 14461 |
| CJ12 | 1 | 99.4% | 85.14% | 0 | 14461 |
| CJ16 | 1 | 99.4% | 85.14% | 0 | 14461 |
| CJ2 | 1 | 99.4% | 85.14% | 0 | 14461 |
| CJ20 | 1 | 99.4% | 85.14% | 0 | 14461 |
| CJ21 | 1 | 99.4% | 85.14% | 0 | 14461 |
| CJ3 | 1 | 99.4% | 85.14% | 0 | 14461 |
| CJ35 | 2 | 99.4% | 85.14% | 0 | 14461 |
| CJ10 | 1 | 99.3% | 76.20% | 0 | 12958 |
| CJ25 | 1 | 98.3% | 74.05% | 0 | 12593 |
| CJ26 | 1 | 99.2% | 64.11% | 0 | 10910 |
| CJ5 | 1 | 99.4% | 37.08% | 0 | 6306 |
| CJ13 | 1 | 99.7% | 30.72% | 0 | 5214 |
| CJ5 | 1 | 99.7% | 30.67% | 0 | 5207 |
| CJ26 | 1 | 99.7% | 30.39% | 0 | 5159 |
| CJ13 | 1 | 99.4% | 30.25% | 0 | 5147 |
| CJ25 | 1 | 99.7% | 25.58% | 0 | 4342 |
| CJ19 | 1 | 99.6% | 21.38% | 0 | 3630 |
| CJ17 | 1 | 98.8% | 20.82% | 0 | 3551 |
| CJ17 | 1 | 99.6% | 20.37% | 0 | 3457 |
| CJ27 | 1 | 99.7% | 15.37% | 0 | 2609 |
| CJ30 | 1 | 97.8% | 14.49% | 0 | 2462 |
| CJ27 | 1 | 98.4% | 13.36% | 0 | 2284 |
| CJ9 | 1 | 99.7% | 12.97% | 0 | 2201 |
| CJ17 | 1 | 99.1% | 12.48% | 0 | 2118 |
| CJ9 | 1 | 98.4% | 12.16% | 0 | 2081 |

*T6SS query length 16,975 from the *Campylobacter* megaplasmid NZ_CP014743 (13,397bp to 30,371bp). Results with a query coverage cutoff <10% were excluded. Isolates with multiple row entries are representative of hits on separate contigs.

**Table S4.12.** Antibiotic susceptibility testing of *Campylobacter* isolates (separate file).

**Table S4.13.** Antimicrobial resistance genes identified in *Campylobacter* isolates.

| ARG | Drug class | Present in |
|---|---|---|
| cmeABC | cephalosporin, fusidane antibiotic, macrolide antibiotic, fluoroquinolone antibiotic | C. coli |
| cmeDEF | cephalosporin, fusidane antibiotic, macrolide antibiotic, fluoroquinolone antibiotic | C. coli, C. jejuni |
| cmeR | cephalosporin, fusidane antibiotic, macrolide antibiotic, fluoroquinolone antibiotic | C. jejuni |
| acr3 | arsenic | C. jejuni |
| arsP | arsenic | C. jejuni |
| blaOXA-61 | cephalosporin, penam, carbapenem | C. coli, C. jejuni |
| tetO | tetracycline | C. coli |

**Table S4.14:** Roary presence absence spreadsheet (separate file).

**Figure S5.1.** Annotated *Salmonella* Kentucky IncI1 plasmid.

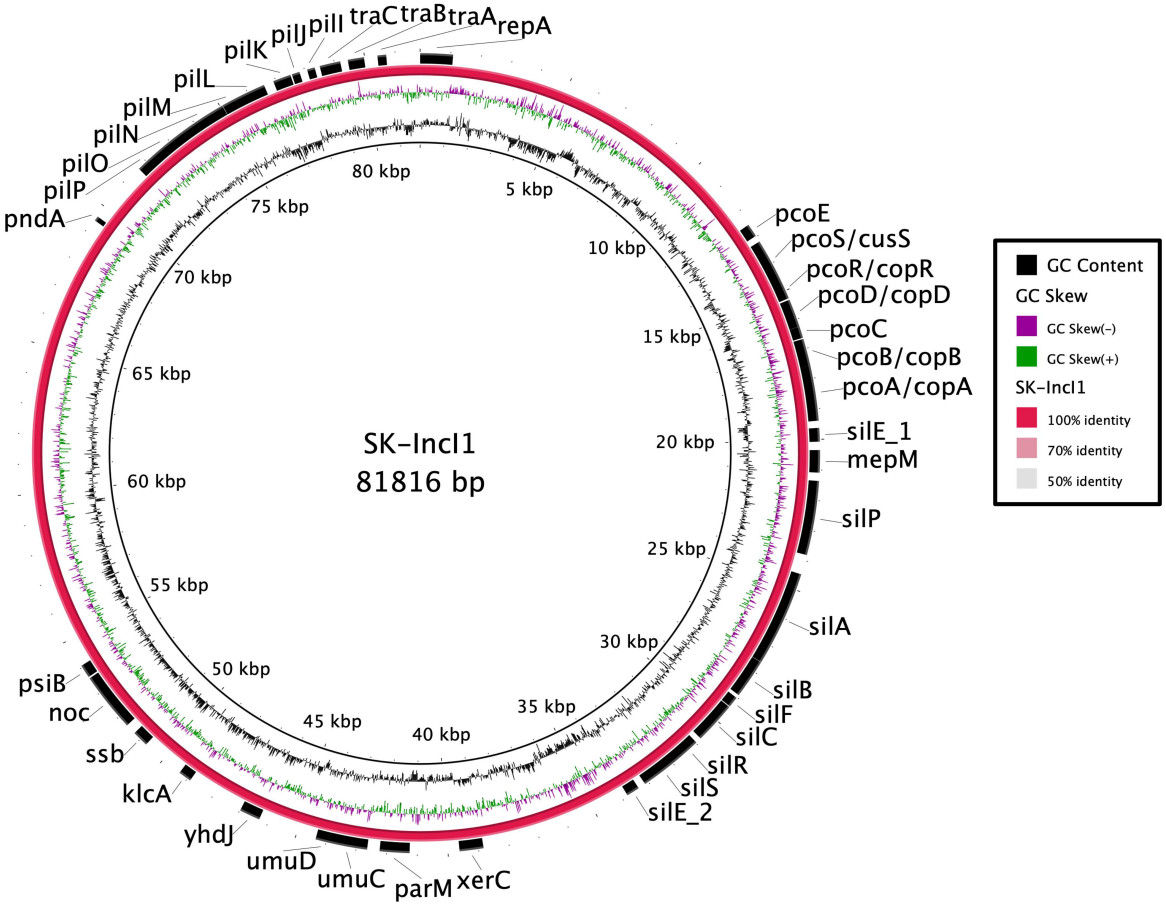


**Figure S5.2.** BLASTn alignment of the *Salmonella* Enteritidis IncF plasmid to the plasmid pSJTUF10978 (GenBank accession # CP015525).

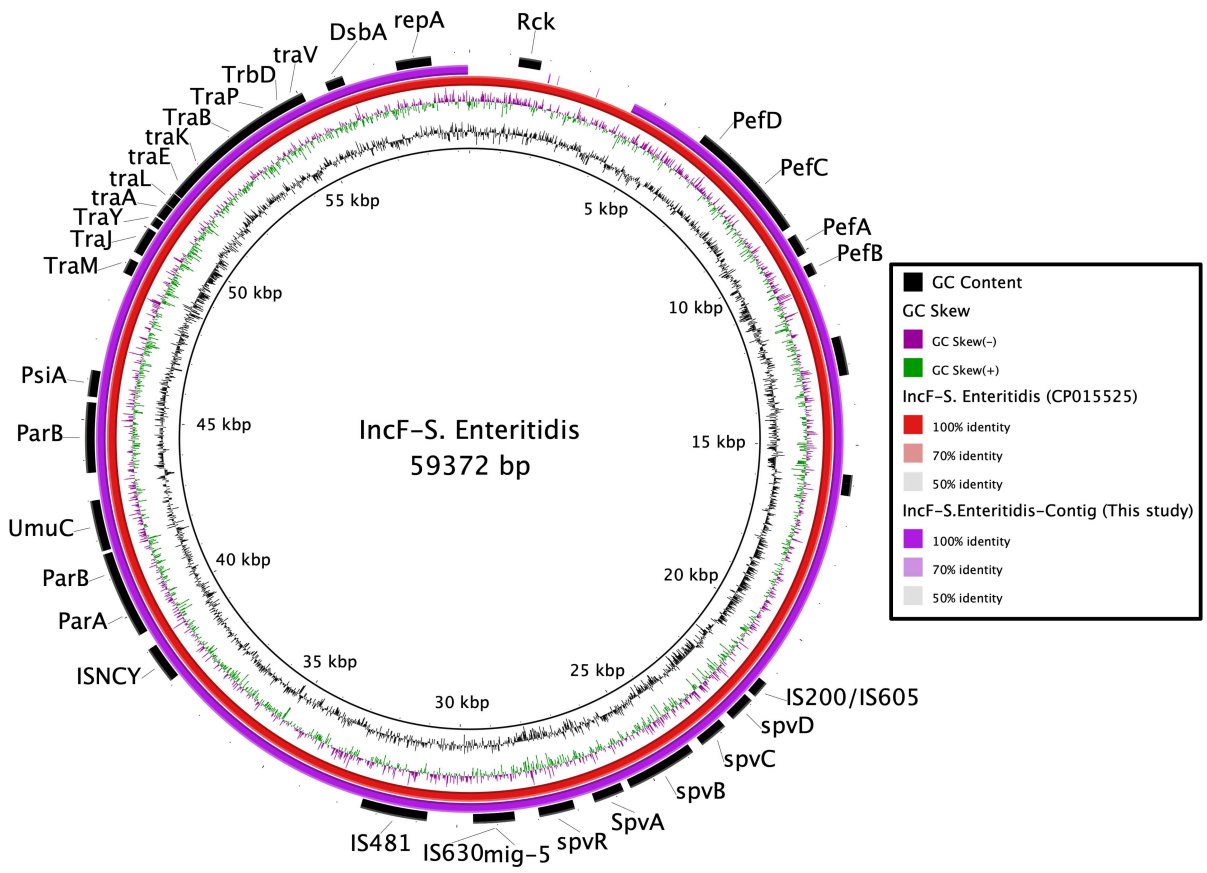



**Table S5.1:** Metadata and ARG and VF presence absence spreadsheet (separate file).

**Table S5.2:** BacMet identified biocide and metal resistance genes (separate file).

**Table S5.3.** Virulence factor functional categories of virulence factors identified from the Virulence Factor Database (VFDB) (via ABRICATE (Reads2Resistome) (separate file).

**Table S5.4.** Drug classes conferred by genes identified by the Resistance Gene Identifier which utilizes the Comprehensive Antibiotic Resistance Database (separate file).

**Table S5.5.** Antibiotic susceptibility testing results for *Salmonella* isolates (separate file).

**Table S5.6.** *Salmonella* isolate-associated plasmid replicons (separate file).

**Table S5.7.** RAST plasmid annotations (separate file).