Refining the definition of the avian pathogenic *Escherichia coli* (APEC) pathotype through inclusion of high-risk clonal groups

Timothy J. Johnson,[†,*,1] Elizabeth A. Miller,* Cristian Flores-Figueroa,† Jeannette Munoz-Aguayo,†
Carol Cardona,* Katie Fransen,‡ Megan Lighty ®,§ Eric Gonder,# Jill Nezworski,‖ Adam Haag,¶
Michelle Behl,‡ Michelle Kromm,§ Ben Wileman,‡ Marissa Studniski,‡ and Randall S. Singer ®*

*University of Minnesota, Department of Veterinary and Biomedical Sciences, Saint Paul, MN, USA; †University of
Minnesota, Mid-Central Research and Outreach Center, Willmar, MN, USA; ‡Select Genetics LLC, Willmar, MN,
USA; §Jennie-O Turkey Store Inc, Willmar, MN, USA; #Juneau, WI, USA; ‖Blue House Veterinary LLC, Buffalo
Lake, MN, USA; and ¶Pilgrims Pride, Sauk Rapids, MN, USA*

ABSTRACT Colibacillosis in poultry is a unique disease manifestation of *Escherichia coli* in the animal world, as one of the primary routes of entry is via the respiratory tract of birds. Because of this, a novel extraintestinal pathogenic *E. coli* (**ExPEC**) subpathotype coined avian pathogenic *E. coli* (or **APEC**) has been described. Like other ExPEC, this pathotype has been challenging to clearly define, and in the case of APEC, its role as an opportunistic pathogen has further complicated these challenges. Using 3,479 temporally matched genomes of poultry-source isolates, we show that the APEC plasmid, previously considered a defining trait of APEC, is highly prevalent in clinical isolates from diseased turkeys. However, the plasmid is also quite prevalent among cecal *E. coli* isolates from healthy birds, including both turkeys and broilers. In contrast, we identify distinct differences in clonal backgrounds of turkey clinical versus cecal strains, with a subset of sequence types (**ST**s) dominating the clinical landscape (ST23, ST117, ST131, ST355, and ST428), which are rare within the cecal landscape. Because the same clinical STs have also dominated the broiler landscape, we performed lethality assays using strains from dominant STs from clinical or cecal landscapes in embryonated turkey and chicken eggs. We show that, irrespective of plasmid carriage, dominant clinical STs are significantly more virulent than dominant cecal STs. We present a revised APEC screening tool that incorporates APEC plasmid carriage plus markers for dominant clinical STs. This revised APEC pathotyping tool improves the ability to identify high-risk APEC clones within poultry production systems, and identifies STs of interest for mitigation targets.

Key words: poultry, *Escherichia coli*, APEC, colibacillosis, pathotype

INTRODUCTION

Avian colibacillosis has remained one of the most significant bacterial diseases in poultry production for over 50 y (Nolan et al., 2020). The causative agent of this disease is *Escherichia coli*, and colibacillosis manifests itself in a variety of different ways. *E. coli* causing colibacillosis have been studied at the molecular level since the early 1990s, and the use of the terminology "avian pathogenic *E. coli*," or **APEC**, dates back to 1992

(Provence and Curtiss, 1992). In 1999, Dho-Moulin and Fairbrother performed a definitive review on APEC, describing key virulence factors and other characteristics of APEC strains (Dho-Moulin and Fairbrother, 1999). As early as 2004, it was recognized that because APEC-caused diseases are extraintestinal in nature, APEC should fall under the broader *E. coli* pathotype known as extraintestinal pathogenic *E. coli*, or **ExPEC** (Kaper et al., 2004).

Not all *E. coli* from cases of colibacillosis are APEC, and the identification of true APEC can only be confirmed through molecular characterization (Collingwood et al., 2014). Even when APEC are present, colibacillosis is primarily opportunistic in nature. Poultry are stressed by a variety of other challenges, and these challenges are often enough alone to enable any *E. coli*, virulent or not, to cause disease in the bird (Johnson et al., 2008).

Because of this, identifying highly virulent APEC capable of lowering the threshold for opportunistic disease to occur in the bird requires molecular characterization. A number of studies have sought to develop effective APEC typing schemes. Some have focused solely on clinical isolates causing colibacillosis (Janben et al., 2001; Ewers et al., 2004; Yaguchi et al., 2007), whereas others have compared clinical isolates versus isolates from healthy birds (McPeake et al., 2005; Vandekerchove et al., 2005; Kawano et al., 2006). From this, differing opinions exist on what defines the APEC pathotype. A large PCR-based study (Rodriguez-Siek et al., 2005), coupled with later genomic-based studies (Johnson et al., 2006a,b; Tivendale et al., 2009), established that the ColV and ColBM plasmids (referred to hereafter as APEC plasmids) possess a pathogenicity-associated island (**PAI**) containing genes that define APEC and differentiate them from avian commensal *E. coli*. These genes were further refined to a subset of 5 defining APEC genes, based on isolate screens combined with in vivo challenge studies (Johnson et al., 2008). Another approach identified genes correlated with strains of higher virulence, irrespective of their prevalence in clinical populations (Ewers et al., 2004; Ewers et al., 2005). From these findings, two PCR-based screens have been widely used to type APEC (Ewers et al., 2005; Johnson et al., 2008). More recently, a genomics-based study has challenged these approaches, reporting that the APEC plasmid PAI is highly prevalent across isolates from asymptomatic birds, and that other APEC virulence factors are not discriminatory between clinical versus commensal isolates (Mageiros et al., 2021). Moreover, Mehat et al. recently proposed that, despite conservation of some virulence-associated traits, there are multiple distinct clonal backgrounds represented among clinical poultry *E. coli* populations which should be considered when defining APEC (Mehat et al., 2021). This prompted us to reconsider the definition of APEC as it relates to commercial poultry.

MATERIALS AND METHODS

Collection of Clinical Isolates

This work was reviewed by the University of Minnesota Institutional Animal Care and Use Committee, and deemed to be exempt from a need for protocol approval. Samples of convenience were collected from moribund turkeys displaying classical lesions of colibacillosis, including airsacculitis, perihepatitis, and/or pericarditis. Swabs from internal organs of these birds were streaked onto MacConkey agar (BD Difco, Franklin Lakes, NJ) and incubated overnight at 37°C. Following incubation, one suspect *E. coli* colony was selected per sample. These isolates were later confirmed to be *E. coli* through DNA sequencing (see below). Only one sample was taken from a barn experiencing colibacillosis-associated deaths at a given timepoint. Samples were collected between January 2017 and 2018, with 397 total isolates

collected (Dataset S1), referred to throughout as Turkey Clinical. Isolates were collected from 7 major turkey producing companies in the United States, across at least 9 different states and 155 different farms. Ages ranged from day-of-hatch through 55 wk of age.

Bacterial DNA Extraction and Sequencing

All Turkey Clinical isolates were sequenced in this study. DNA was extracted from overnight TS broth (BD Difco) cultures of a single colony using the DNeasy Blood & Tissue kit (Qiagen, Valencia, CA) following the manufacturer's instructions. Genomic DNA libraries were created using the Nextera XT DNA library preparation kit and Nextera XT index kit v3 (Illumina, San Diego, CA), and sequencing was performed using 2 × 300-bp dual-index runs on an Illumina MiSeq.

Procurement of E. coli Database Genomic Data

A search of Enterobase (Zhou et al., 2020) (July 2021) was conducted for *E. coli*. Isolates were retained that 1) were collected as a part of the National Antimicrobial Resistance Monitoring System (**NARMS**) program by US Food and Drug Administration, or collected by the United States Department of Agriculture Food Safety Inspection System; 2) had a known isolation year of 2017 to 2020; 3) had a known isolation source; and 4) included source metadata indicating the isolate was sourced from turkey cecal contents from recently slaughtered turkeys (Turkey Cecal), cecal contents from recently slaughtered broilers (Chicken Cecal), or retail turkey or chicken meat (Turkey Retail and Chicken Retail, respectively). Raw sequencing reads of all identified isolates (N = 1,469 for Turkey Retail, N = 562 for Turkey Cecal, N = 611 for Chicken Retail, N = 440 for Chicken Cecal) were downloaded from the NCBI short-read archive (SRA; https://www.ncbi.nlm.nih.gov/sra) using the SRA Toolkit (v2.8.2) (Dataset S1).

Genome Assembly and Quality Assessment

Raw FASTQ files for each genome were trimmed and quality filtered using Trimmomatic (v0.33) (Bolger et al., 2014), including removal of Illumina adapters, with a sliding window of 4 and average Phred quality score of 20, and 36 as the minimum read length. Assemblies of each genome were performed using Shovill (v1.0.4), specifying the SPAdes assembler (Bankevich et al., 2012), with default parameters (https://github.com/tseemann/shovill). Assembly quality was assessed with QUAST (v5.0.0) (Gurevich et al., 2013).

Serotype and Sequence Type Prediction

In silico (genomic) serotype prediction was performed with ECTyper (v1.0.2) (Bessonov et al., 2021) using a

minimum sequence identity of 50% and minimum hit coverage of 50% searched against a curated database of O and H antigens. Serotypes were reported as O and H antigen types. In silico multilocus sequence typing (**MLST**) was performed using mlst (v2.16.1) (https://github.com/tseemann/mlst), using the seven-gene *E. coli* MLST scheme hosted on the PubMLST website (https://pubmlst.org) (Jolley and Maiden, 2010). Minimum spanning trees based on traditional MLST were generated in GrapeTree (v1.5.0) using the MSTree V2 algorithm (Zhou et al., 2018).

Genetic Feature Identification

ABRicate (v.0.8.13) (https://github.com/tseemann/abricate) was used with a minimum identity of 90% and minimum coverage of 80% to screen isolate genome assemblies for *E. coli* virulence factors using the VirulenceFinder database (Malberg Tetzschner et al., 2020). EZClermont was used to predict *E. coli* phylogenetic group for each isolates using default parameters (Waters et al., 2020).

A custom APEC database consisting of 46 genes was created using a scan of existing literature for genes linked to APEC virulence or fitness through direct evidence or epidemiological association. This database is freely available at https://github.com/JohnsonSinger Lab/APEC_VF_database. ABRicate was used with a minimum identity of 90% and minimum coverage of 80% to determine APEC gene prevalence across all isolates.

Phylogenetic Analyses

Single nucleotide polymorphisms (**SNP**s) were identified in each sample using Snippy (v4.4.0), with a minimum sequencing depth of 8x (https://github.com/tseemann/snippy) and *E. coli* strain APEC O1 used as reference (Johnson et al., 2007a). Separate core SNP alignments were created for all Turkey Clinical isolates (n = 397) and for all ST117 isolates (n = 201). Maximum likelihood trees for both datasets were reconstructed with IQ-TREE (v1.6.10), using 1,000 ultrafast bootstrap iterations (Nguyen et al., 2015). ModelFinder was used to identify the most appropriate substitution models (Kalyaanamoorthy et al., 2017). For the Turkey Clinical isolate tree, TVM+F+R4 was used. For the ST117 isolate tree, TVM+F+ASC+R4 was used. The Interactive Tree of Life was used for tree construction (Letunic and Bork, 2016).

Pan-Genome Analyses

Genome assemblies for each isolate were annotated with Prokka (Seemann, 2014), and core genome alignments were generated using Roary (v3.12.0) (Sitto and Battistuzzi, 2020) with 95% sequence identity. Scoary (v1.6.16) (Brynildsrud et al., 2016) was then used for pan-genome-wide association analysis comparing Turkey Clinical isolates versus Turkey Cecal isolates. A gene was reported as significantly associated with Turkey Clinical isolates if it had a Benjamini-Hochberg (BH)-adjusted P value of ≤ 0.05 and was present in $\geq 50\%$ of isolates in Turkey Clinical isolates and $\leq 50\%$ in Turkey Cecal isolates. Reference sequences of each significant gene were annotated using the top hit from a BLASTX search against the NCBI's nonredundant protein sequence database (Altschul et al., 1990).

Embryo Lethality Assays

A chicken embryo lethality assay was conducted as previously described (Wooley et al., 2000). Embryonated chicken eggs incubated 10 to 11 d were obtained from a local hatchery and transferred to the University of Minnesota Mid-Central Research and Outreach Center. At 12 d of incubation, embryonated eggs were candled for viability and marked around the air sac. An overnight growth in Brain Heart Infusion Broth of each isolate tested was washed twice with sterile phosphate buffered saline (PBS) and diluted to 5,000 colony-forming units (**CFU**)/mL. Eggs were disinfected with 70% ethanol prior to inoculations. Inoculum was delivered at a volume of 0.1 mL (500 CFUs) through the air sac into the allantoic cavity. Eggs were then sealed using glue. Embryos were candled for viability each subsequent day for 5 consecutive days. For each strain tested, 2 biological replicates were performed with 12 eggs per replicate. Negative controls included uninoculated eggs and *E. coli* K-12 strain MG1655.

A turkey embryo lethality assay was also developed based upon previous work studying *Ornithobacterium rhinotracheale* (Walters, 2014). Turkey embryos incubated 12 to 14 d were obtained from a local turkey hatchery and transferred to the University of Minnesota Mid-Central Research and Outreach Center. At 16 d of incubation, embryos were inoculated with 1,000 CFUs contained in 0.2 mL of PBS with a bacterial concentration of 5,000 CFU/mL. All other procedures mimicked the chicken assay.

Development of a Revised APEC Typing Scheme

To identify ST-specific markers, the entire turkey isolate collection (N = 2,428 genomes) was grouped based on ST type, and pangenomic analyses were conducted for each ST on interest versus all other isolates using Roary and Scoary, as described above. From this, distinguishing gene markers were sought for each ST. PCR primers for these gene markers were designed using SeqBuilder Pro (Lasergene, Madison, WI). The revised panel also included 2 markers of the APEC plasmid PAI (Johnson et al., 2008), and an O78 serogroup-specific marker targeting the unique region of the O78 *rfb* gene cluster (Wang et al., 2014) (Table 1). The panel was validated by screening a subset of strains from this study

Table 1. Primers used for a revised multiplex PCR typing scheme for high-risk avian pathogenic *E. coli*.

Name	Sequence	Target	Amplicon Size
ST23_F	TGGAGCTAAATGACCCGAC	Phage holin family protein (ST23)	263
ST23_R	AACCAGACGTGCCACATTG		
ST355_1_F	TCAGAGAAGATTGAAGAGGGCG	Hypothetical protein (ST355)	305
ST355_1_R	AATACTCCACCTGAGAGTCCCG		
ST117_F	GATGCCATAAGGAAGAACGAG	StfH/YfcO family fimbrial adhesin (ST117)	360
ST117_R	CGAGCCATTGGATTGCAAC		
ST428_F	GTGTGCTATGCACTACAGG	DNA-directed RNA polymerase subunit beta (ST428)	400
ST428_R	CTTCAGCAGGTTCAGTCATTC		
HLYF_F	GGCCACAGTCGTTTAGGGTGCTTACC	Avian hemolysin HlyF (APEC plasmid)	450
HLYF_R	GGCGGTTTAGGCATTCCGATACTCAG		
OMPT_F	TCATCCCGGAAGCCTCCCTCACTACTAT	Outer membrane protease (APEC plasmid)	496
OMPT_R	TAGCGTTTGCTGCACTGGCTTCTGATAC		
ST355_2_F	TGCTTGAGAATGTGAAGAACC	DNA cytosine methyltransferase (ST355)	548
ST355_2_R	AGGTGTCAGTCTTCTTGGTC		
ST131_F	GTTCGACAAAATCCTCTCCG	Divalent metal cation transporter (ST131)	578
ST131_R	GCACAACCAGACAAAGCAG		
O78_F	CGATGTTGAGCGCAAGGTTG	Gnd-Wzx (O78 *rfb* gene cluster)	623
O78_R	TAGGTATTCCTGTTGCGGAG		

(n = 18). Amplification of targets was accomplished in a 25 μL reaction volume, prepared as master mix pools prior to the addition of DNA template. Each reaction included 10.675 μL of nuclease-free water, 5.0 μL of 5X Green GoTaq Flexi Buffer, 4.0 μL of 25 mM MgCl$_2$, 2.5 μL of the primer pool which contained 3 μM of each primer, 0.625 μL of 10 mM dNTPs, 0.2 μL of 5 U/μL GoTaq G2 Hot Start DNA Polymerase, and 2.0 μL of template DNA. The reactions were performed using a T100 thermal cycler (BioRad, Hercules, California, USA) using the following cycling parameters: 95°C for 5 min; 30 cycles of 95°C for 35 s, 57°C for 30 s, 72°C for 40 s; and a final cycle of 72°C for 10 min. Samples were subjected to gel electrophoresis in 2% TBE agarose, and amplicons were compared to a 100-bp ladder (New England Biolabs Inc). An isolate was considered to contain a gene of interest if it produced an amplicon of the expected size (Table 1).

A command line tool was also developed for in silico typing of partial or complete *E. coli* genome assemblies based on the revised APEC typing scheme. The Bash shell script, *APECtyper.sh*, wraps the E. coli typing tools ECTyper (v1.0.0) and h (v2.19.0), along with a custom R (Team, 2020) script, into a single pipeline and generates a summary report of serotype, sequence type, and APEC pathotype classification. The tool also uses blastn (Camacho et al., 2009) with user-defined identity and coverage thresholds to scan assemblies for the presence of genes found in the custom APEC virulence and fitness gene database. Further information on APECtyper installation, usage, and outputs can be found at https://github.com/JohnsonSingerLab/APECtyper.

Data Availability

Raw reads from isolates sequenced in this study are available at the NCBI Short Read Archive (**SRA**) under BioProject accession no. PRJNA799011. APECtyper is

freely available at: https://github.com/JohnsonSinger Lab/APECtyper. The APEC virulence and fitness gene database is available at: https://github.com/Johnson SingerLab/APEC_VF_database.

RESULTS

Clinical E. coli from Commercial Turkeys are Dominated by a Subset of Clonal Groups, Irrespective of Geographical Location or Company

Genome sequences of 397 Turkey Clinical *E. coli* representing 7 major turkey-producing companies across 9 US states were first examined for their genetic relatedness using 89,214 core SNP variants (Figure 1). In general, the majority of isolates clustered into 5 dominant
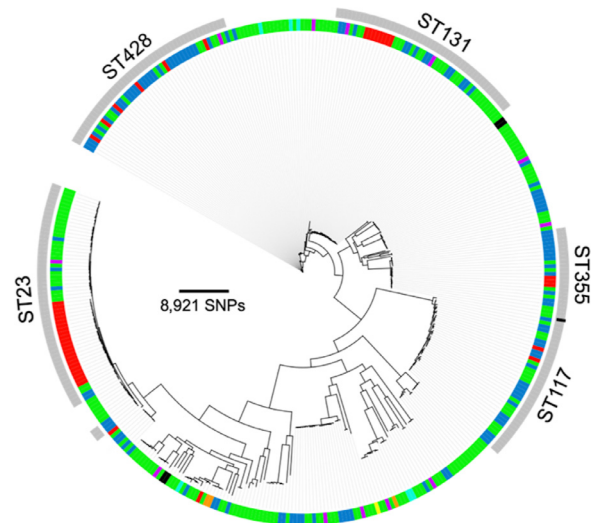

Figure 1. Single nucleotide polymorphism-based phylogenetic tree depicting relationships between Turkey Clinical isolates sequenced in this study (N = 397). The inner ring depicts isolates colored by company (blinded) with different colors representing different companies submitting clinical isolates. The outer ring highlights the five dominant sequence types identified amongst this collection.

clades where isolates within these clades shared close genetic relatedness, and included 54% of the total isolate population. Remaining isolates clustered either into smaller clades or were singletons on the phylogenetic tree. Within the 5 major clades (subsequently typed as ST23, ST117, ST131, ST355, and ST428), multiple companies and geographies were represented.

Clinical and Cecal E. coli From Commercial Turkeys Both Possess the APEC Plasmid Pathogenicity-Associated Island (PAI) at High Proportions

Genomes of isolates (N = 2,642) from 3 turkey-source populations (Clinical, Cecal, and Retail) were examined for their possession of 46 APEC virulence or fitness genes (Dataset S2), including 34 APEC plasmid PAI genes (Johnson et al., 2008) (Figure 2). Among Turkey Clinical isolates, genes within the conserved virulence region of the APEC plasmid PAI (Johnson et al., 2006b), including *etsA* through *sitD* in Figure 2, were found at a rate of 72% to 96%. In Turkey Cecal isolates, genes within this region were found at a rate of 45% to 85%. In Turkey Retail isolates, this region was found at a rate of 56-91%. Comparing Turkey Clinical versus Turkey Cecal, 28/34 of the APEC PAI plasmid genes examined were significantly more prevalent in Turkey Clinical population ($P < 0.05$, Dataset S1). Overall, the patterns of prevalence of APEC plasmid PAI genes among Turkey Clinical *E. coli* were quite similar to that previously found for broiler clinical *E. coli* (Rodriguez-Siek et al., 2005; Johnson et al., 2006b) except that the aerobactin siderophore system was found at lower prevalence. Similar to previous reports (Johnson et al., 2006b), genes outside of the APEC plasmid PAI's conserved region (*aatA* through *eitD*, and *tsh*) were also found at lower prevalence. In general, APEC genes within Turkey Retail isolates were proportionally intermediate to the Turkey Clinical and Turkey Cecal isolates, similar to previous reports

in broilers (Johnson et al., 2007b; Johnson et al., 2009; Danzeisen et al., 2013).

The Genomic Backgrounds of Turkey Clinical and Turkey Cecal E. coli are Substantially Different

To study genomic backgrounds, we examined Clermont phylogenetic group, 7-gene multilocus sequence type (ST), and predicted serogroup among all isolates. Using Clermont phylogenetic group as the highest genomic level depicting *E. coli* chromosomal lineage (Figure 3), it was clear that Turkey Clinical and Turkey Cecal *E. coli* differed substantially. The dominant phylogenetic group among Turkey Clinical isolates was B2 (46%), whereas only 6% of Turkey Cecal isolates belonged to the B2 phylogroup. Conversely, Turkey Cecal isolates belonged primarily to phylogroups B1 (41%) and A (29%), compared to 8% and 2% for Turkey Clinical isolates, respectively. Additionally, phylogroup C was over-represented by Turkey Clinical isolates (19%) compared to Turkey Cecal isolates (2%). In general, Turkey Retail isolates were again proportionally intermediate to the Turkey Clinical and Turkey Cecal isolates with respect to phylogroup.
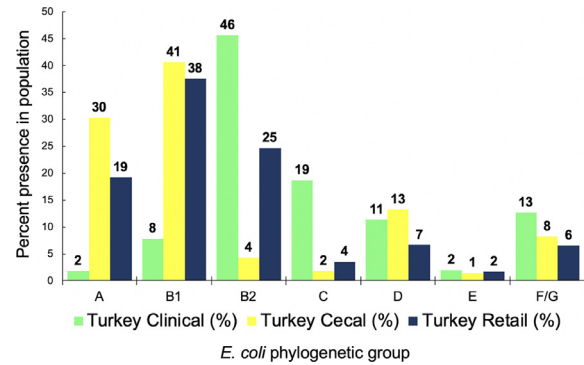

Figure 3. Patterns of prevalence of Clermont phylogenetic groups among Turkey Clinical (N = 397), Turkey Cecal (N = 562), and Turkey Retail (N = 1,468) *E. coli*. Data are displayed using a bar graph depicting population prevalence (%) above each bar.
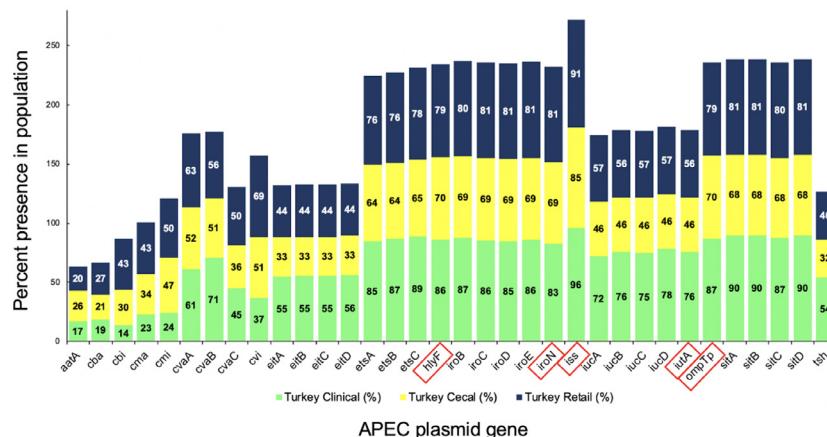

Figure 2. Patterns of prevalence of selected genes of the APEC plasmid PAI among Turkey Clinical (N = 397), Turkey Cecal (N = 562), and Turkey Retail (N = 1,468) *E. coli*. Data are displayed using a stacked bar graph depicting population prevalence (%). Genes of the pentaplex APEC typing scheme (Johnson et al., 2008) are boxed in red.

Table 2. Distribution of *E. coli* sequence types (STs) by isolate source.

Sequence type	Turkey clinical % (N = 397)	Turkey cecal % (N = 562)	Turkey retail % (N = 1,469)	Chicken cecal % (N = 440)	Chicken retail % (N = 611)	Clermont phylotype	P value (Turkey clinical vs. cecal)
10	0.3	8.2	5.0	6.1	3.9	A	<0.0001
23	14.4	0.9	2.6	1.8	2.1	C	<0.0001
58	1.5	11.2	6.3	1.8	1.8	B1	<0.0001
69	2.0	5.5	2.3	1.1	1.0	D	0.0073
117	8.6	4.3	4.3	4.1	10.1	G	0.0085
131	12.6	1.6	3.1	0.5	2.0	B2	<0.0001
155	0.5	6.6	2.9	4.1	3.4	B1	<0.0001
189	0.0	0.0	0.1	12.3	0.2	A/D	1.0000
355	5.5	0.2	3.5	0.2	0.8	B2	<0.0001
428	12.6	0.2	2.9	0.0	0.8	B2	<0.0001
3580	1.5	3.9	6.1	0.0	0.0	B1	0.0321
Total % represented	59.4	42.5	39.1	32.0	26.2		

STs with greater than 5.0% prevalence in any source group are displayed.

Isolates were then classified according to 7-gene multilocus sequence type (**ST**) (Table 2 and Figure 4). The dominant STs from Turkey Clinical isolates included ST23 (14.4%), ST131 (12.6%), ST428 (12.6%), ST117 (8.6%), and ST355 (5.5%). For Turkey Cecal isolates, dominant STs included ST58 (10.5%), ST10 (7.6%), ST155 (6.6%), and ST69 (5.3%). Of the Turkey Clinical-associated STs, all five dominant clinical STs were significantly higher in Turkey Clinical versus Turkey Cecal isolates (Fisher's exact tests, all $P < 0.01$). Of the dominant Turkey Clinical STs, ST131 and ST429 belonged to the B2 phylogroup, ST23 belonged to the C phylogroup, and ST117 belonged to the G phylogroup. The ST23 isolates were predicted to possess the O78 serogroup, ST131 isolates were predicted as O25, ST355 were predicted as O2, and ST428 were predicted as O117. ST117 isolates were predicted as a variety of different serogroups, including O35, O45, O78, O109, and O111. All of the most prevalent Turkey Cecal isolate STs contained multiple predicted serogroups.

Using the previously published 5-gene scheme for typing APEC (Johnson et al., 2008), with criteria of 4 or more genes indicating APEC plasmid possession, APEC plasmid prevalence was examined across ST type (Figure 5). All major STs had isolates carrying genes marking the APEC plasmid; however, the proportions differed by ST type. For example, ST58, ST131, ST355, ST428, and ST3580 had nearly all isolates possessing four of five genes marking the APEC plasmid. Among ST10, ST23, and ST117 isolates, >25% of isolates lacked two or more of these genes indicating absence of the APEC plasmid.

Figure 4. Minimal spanning trees of Turkey Clinical, Turkey Cecal, and Turkey Retail isolates based upon the 7-gene *E. coli* MLST scheme. A: Sequence types (STs) are colored by source and dashed lines indicate membership in Clermont phylogenetic groups. B: STs are colored proportionally by percentage containing the APEC plasmid, and major STs are noted in either green (Cecal-associated) or red (Clinical-associated).

No Clearly Defining Genes Were Identified Between Turkey Clinical and Turkey Cecal Isolates

A comparison of Turkey Clinical versus Turkey Cecal isolates was performed using pan-genome-wide association study to identify genes significantly higher in clinical isolates. A total of 430 genes were identified that were 1) present in ≥50% Turkey Clinical isolates, 2) present in ≤50% Turkey Cecal isolates, and 3) had an adjusted $P < 0.05$ (Dataset S3). This approach did identify some APEC PAI genes, including *eitABCD, iutA, iucABCD*, and *tsh*. Other APEC PAI genes were not identified using this approach because they were present in >50% of Turkey Cecal isolates. Several additional genes of interest were identified ($P < 0.05$, OR presented as risk factor for clinical), including *ccdA* and *ccdB*, a type II toxin-antitoxin system (65%−67% vs. 20%,
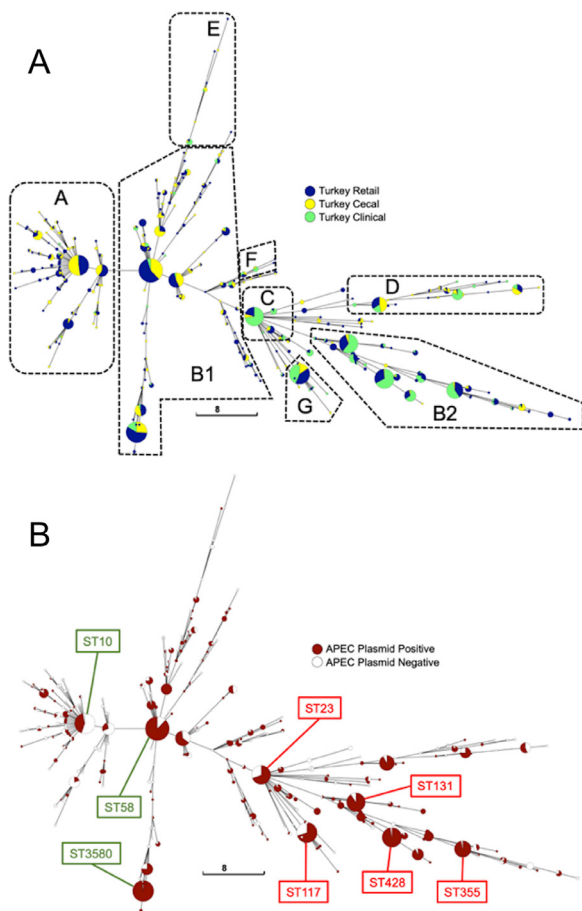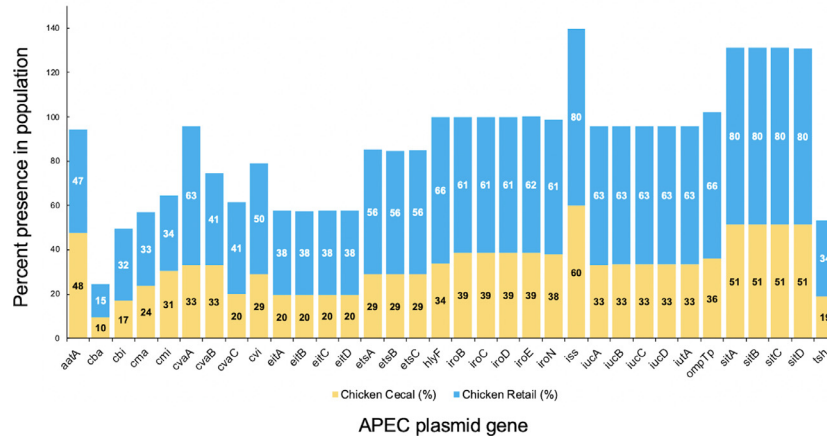
Figure 5. Patterns of prevalence of selected genes of the APEC plasmid PAI among Chicken Cecal (N = 440) and Chicken Retail (N = 611) *E. coli*. Data are displayed using a stacked bar graph depicting population prevalence (%).

OR = 7.4−8); a putative ABC transporter system (71% vs. 28%, OR = 6.1−6.3); a type VI secretion system (54-55% vs. 8−11%, OR = 9.0−13.1); a capsular biosynthesis cluster (57% vs. 12%−13%, OR = 8.7−10.2); a putative iron transport system (58%−68% vs. 17%−21%, OR = 6.9−7.8); and a putative sugar utilization and phosphotransferase system (59%−61% vs. 17%−18%, OR = 6.8−7.4). Notably, no significantly enriched genes were identified, which were >80% present among clinical isolates and <20% present among cecal isolates, and many of the genes found as >75% present among Turkey Clinical isolates were actually allelic variants of gene systems generally conserved among *E. coli*. Collectively, this analysis indicates that a set of genes definitively differentiating between clinical and commensal *E. coli* from poultry likely does not exist.

Chicken Cecal and Chicken Retail E. coli Display Similarities to Turkey E. coli

Using a similar approach, we examined Chicken Cecal (N = 440) and Chicken Retail (N = 611) database isolates for presence of APEC plasmid PAI genes and Clermont phylogenetic type. APEC plasmid PAI gene prevalence was approximately 30% lower in Chicken Cecal versus Turkey Cecal isolates (Figure 5). For example, the conserved region of the APEC plasmid PAI (*etsABC, hlyF, iroBCDEN, iss, ompTp,* and *sitABCD*) ranged from 29% to 60% prevalence in Chicken Cecal isolates, compared to 64%−85% prevalence in Turkey Cecal isolates. Chicken Retail isolates possessed the conserved region of the APEC plasmid PAI at slightly lower prevalence than Turkey Retail isolates (range 61%−80% vs. 76%−91%, respectively). Phylogenetic types of Chicken Cecal and Chicken Retail isolates mostly mimicked those of Turkey Cecal and Turkey Retail isolates (Figure 6), except for lower prevalence of the B2 phylogenetic group in Chicken Retail versus Turkey Retail (8% vs. 25%, respectively), and higher prevalence of the F/G phylogenetic groups in Chicken Retail versus Turkey Retail (19% vs. 6%, respectively). This agrees with previous reports of high occurrence of ST117 in broiler
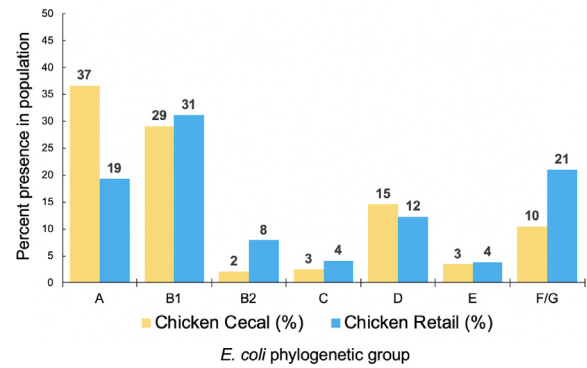

Figure 6. Patterns of prevalence of Clermont phylogenetic groups among Chicken Cecal (N = 440) and Chicken Retail (N = 611) *E. coli*. Data are displayed using a bar graph depicting population prevalence (%) above each bar.

clinical isolates, which belongs to the G phylogenetic group (Braga et al., 2016; Cordoni et al., 2016; Poulsen et al., 2018).

ST117 Exemplifies Evidence of Common Commensal Clones Circulating Between Turkeys and Chickens, with Potential to Cause Disease

We constructed a SNP-based phylogenetic tree of ST117 isolates in the dataset, because of their shared relative isolate abundance across the source populations studied (Figure 7). This analysis illustrates the extensive diversity of ST117, with numerous predicted serogroups represented and 7,810 core SNP sites identified across all isolates. The distribution of serogroups across the tree suggests recombination of O-antigen gene clusters driving serogroup diversity in ST117, as previously documented for *E. coli* (Senchenkova et al., 2016). Turkey and chicken isolates were intermingled throughout the ST117 tree, with evidence of closely related clones colonizing both hosts and differing by <10 SNPs. Furthermore, Turkey Clinical and Turkey Cecal isolates were also intermixed throughout the tree, often genetically indistinguishable (<10 SNPs in some branches). APEC
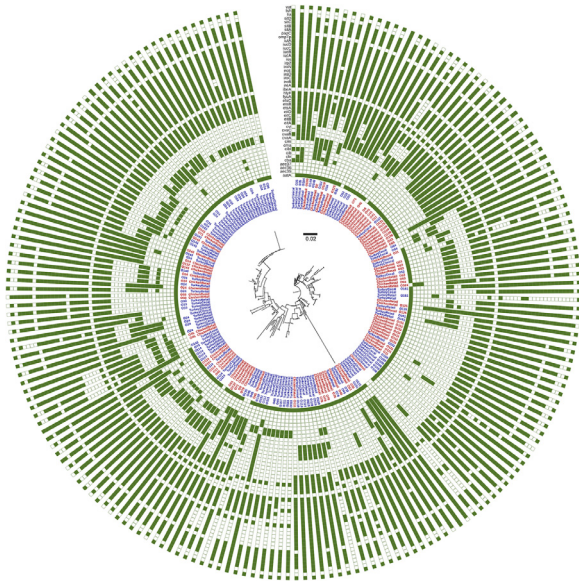
Figure 7. Core single nucleotide polymorphism-based phylogenetic tree depicting relationships between ST117 isolates analyzed in this study (N = 201). The inner ring lists isolate name or NCBI SRA accession number. The second two rings depict isolate source and serogroup, colored by host source (red = chicken and blue = turkey). The outer ring displays presence (green) or absence (white) of 46 APEC-associated fitness or virulence factors.

plasmid PAI genes were highly prevalent across ST117 isolates, irrespective of source. Pangenomic analysis of ST117 Turkey Clinical versus Turkey Cecal isolates (data not shown) again failed to identify genes substantially represented in clinical isolates but lacking from cecal isolates, further demonstrating the lack of genomic differences between isolates from the gut and those causing disease. Together, this highlights the challenges in identifying APEC-defining genes when comparing clinical versus commensal isolates in the context of primarily opportunistic disease, and illustrates that common clones circulate in both broilers and turkeys with the potential to cause disease.

Dominant Clinical STs are More Virulent Than Dominant Cecal STs, Irrespective of APEC Plasmid Carriage

An embryo lethality assay was used with both embryonated turkey and chicken eggs to assess dominant STs found in this study for their virulence potential. This assay has been shown to correlate with in vivo challenge studies (Wooley et al., 2000; Gibbs and Wooley, 2003), and it is amenable towards screening relatively large numbers of isolates compared to live bird challenge models. Isolates were selected from each ST with differential carriage of genes of the APEC plasmid PAI, where possible (Table 3). Two themes emerged from these results. First, isolates from dominant Turkey Clinical STs were clearly more virulent towards turkey embryos than isolates from dominant Turkey Cecal STs. Lethality in the dominant Turkey Cecal STs (ST10 and ST58) ranged from 0.0% to 16.7%, whereas isolates from dominant Turkey Clinical STs (ST23, ST131, ST355, and ST428) ranged from 54.2% to 91.7%. A similar result was observed in embryonated chicken eggs. Second, there were patterns within STs between turkey embryo lethality and APEC plasmid PAI gene content. For example, within ST23 the isolate with only 7 APEC plasmid PAI genes was less lethal toward embryonated turkey eggs than the isolate with 37 APEC plasmid PAI genes (66.7% vs. 91.7%), yet it was still among the most lethal strains tested even without the APEC plasmid PAI genes. In contrast, within ST117, lethality mostly correlated with APEC plasmid PAI gene content for the 4 isolates examined (14−42 genes ranging in lethality from 4.2% to 54.2% in turkey embryos and 16.7%−70.8% in chicken embryos). The most lethal isolates in the turkey ELA were those from ST23 containing 37 APEC plasmid PAI genes, and those from ST131 containing 33 APEC plasmid PAI genes. ST23 isolates were also most lethal in the chicken

Table 3. Virulence of selected avian *E. coli* isolates measured through embryo lethality assay (ELA).

Isolate	Serogroup	Clermont group	ST	APEC plasmid PAI gene count	Turkey ELA Mortality (%)	Chicken ELA Mortality (%)
K-12 MG1655	Rough	A	10	0	0.0	2.1
PP865	O89	A	10	15	0.0	25.0
PP984	O89	A	10	35	8.0	29.2
PP394	O8	B1	58	33	16.7	0.0
PP619	O8	B1	58	35	4.2	0.0
PP234	O36	B1	155	31	50.0	58.3
PP262	O25	B2	131	33	87.5	70.8
PP734	O25	B2	131	33	83.3	70.8
PP348	O2	B2	355	41	75.0	83.3
PP577	O2	B2	355	34	54.2	50.0
PP269	O117	B2	428	35	70.8	75.0
PP417	O117	B2	428	30	54.2	54.2
PP167	O78	C	23	7	66.7	83.3
PP731	O78	C	23	37	91.7	83.3
PP320	O111	G	117	14	4.2	33.3
PP410	O109	G	117	27	29.2	16.7
PP438	O78	G	117	42	50.0	70.8
PP554	O119	G	117	31	54.2	45.8

ELA, along with one ST355 isolate. In general, isolates from the B2 phylogenetic background were more lethal toward turkey embryos than other backgrounds, but this was confounded in some part by number of isolates examined and APEC plasmid PAI gene content.

Development of a Revised APEC Typing Scheme

Pangenome analyses were then conducted to identify genomic markers unique to dominant clinical STs. For the 5 clinical STs targeted, gene markers were successfully identified targeting each ST with high genomic sensitivity and specificity (Dataset S3). From this, a PCR-based scheme was developed to identify these dominant clinical STs (Table 1). Based on the results of the ELA where the APEC plasmid presence slightly enhanced virulence across multiple STs, and O78 isolates were highly virulent irrespective of ST, we included markers for these traits along with markers of dominant clinical STs in a revised typing scheme. In this scheme, detection of an isolate belonging to a known high-virulence ST (ST23, ST131, ST355, and ST428) or serogroup O78 classifies it as high virulence (Figure 8). Presence of the APEC virulence plasmid (via possession of $hlyF$ and $ompTp$) with one of the aforementioned markers classifies an isolate as a high-risk APEC clone. Presence of the APEC plasmid genes in the absence of clonal markers does not rule out that an isolate is APEC or virulent, but does not classify it as a high-risk APEC. Using the primer sets designed, the multiplex PCR was successfully performed on strains characterized in this study, with a representative agarose gel in Figure 9.

This revised APEC typing scheme was also used to develop an in silico typing tool called APECtyper.

For each input partial or complete $E.\ coli$ genome assembly, APECtyper generates a summary report of serotype, sequence type, APEC virulence and fitness genes, and APEC pathotype. This tool, along with user instructions, is freely available at the following GitHub repository: https://github.com/JohnsonSingerLab/APECtyper.
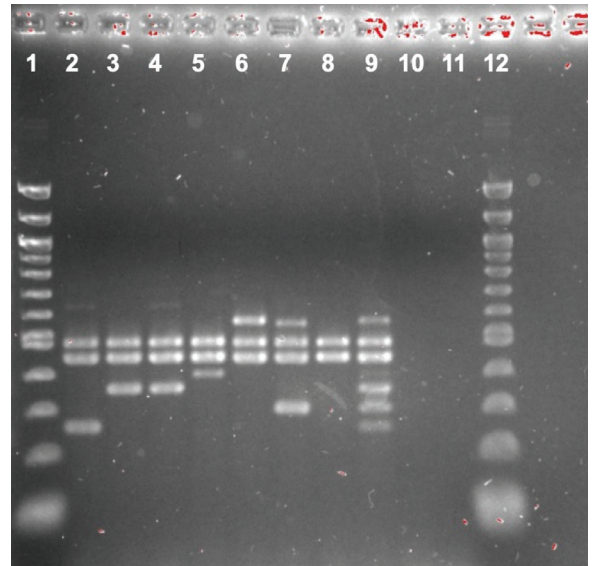

Figure 9. Agarose gel electrophoresis depicting the revised 9-plex PCR that detects high-risk APEC clones. Lane 1 = 100-bp ladder; Lane 2 = PP0878 ($ompT$, $hlyF$, O78, ST23); Lane 3 = PP0507 ($ompT$, $hlyF$, ST117); Lane 4 = PP0293 ($ompT$, $hlyF$, O78, ST117); Lane 5 = PP0178 ($ompT$, $hlyF$, ST428); Lane 6 = PP0171 ($ompT$, $hlyF$, O78, ST131); Lane 7 = PP0209 ($ompT$, $hlyF$, ST355); Lane 8 = PP0902 ($ompT$, $hlyF$); Lane 9 = DNA pool of isolates from Lanes 2-8; Lane 10 = $E.\ coli$ K-12 MG1655; Lane 11 = blank; Lane 12 = 100-bp ladder.
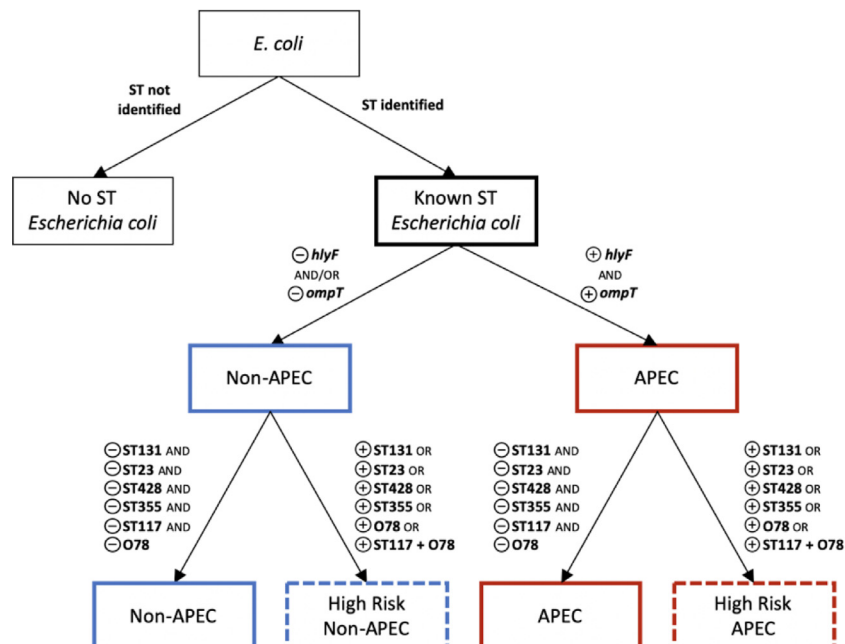

Figure 8. Scheme for revised typing of high-risk avian pathogenic $E.\ coli$ (APEC).

DISCUSSION

The definition of the APEC pathotype has been a subject of debate for many years. This study was prompted by a recent report indicating that a commonly used 5-gene APEC typing scheme may not discriminate virulence potential, based upon the observation that high proportions of gastrointestinal-source avian *E. coli* isolates from healthy birds possessed APEC plasmid PAI genes, contrasting previous findings (Mageiros et al., 2021). Their approach utilized broiler clinical versus "asymptomatic" isolates and determined that, as expected, clinical isolates were identified across a wide range of *E. coli* lineages. In their study, ST117 isolates represented 39% of the total poultry isolates in the dataset, and this seems to be typical for European clinical broiler *E. coli* (Ronco et al., 2017; Mehat et al., 2021). In contrast to previous work (Rodriguez-Siek et al., 2005), Mageiros et al. found that the genes of the APEC plasmid were ubiquitous not only in clinical isolates, but also in asymptomatic isolates. Notably, there did appear to be differences between genes of the APEC plasmid PAI genes in these 2 populations, with some of these genes present in approximately 20% more of clinical than asymptomatic isolates, albeit at high prevalence in both populations. One conclusion from this study was that the APEC pathotype is complex because it relies on combinations of fitness- versus virulence-associated traits, and likely the blend of plasmid-associated traits in multiple, optimal chromosomal backgrounds. This concept was also recently proposed using genomic analyses of clinical isolates, indicating that a subset of clonal groups may dominant the clinical landscape, and that APEC are comprised of multiple lineages which need to be considered in future typing schemes (Mehat et al., 2021).

In this study, we observed that dominant Turkey Clinical STs were largely absent from the Turkey Cecal isolate collection. We utilized a sampling strategy that matched clinical versus cecal isolates both temporally and geographically, in the sense that multiple geographic regions in the USA were represented in both datasets across the same time periods. Also, the datasets used came from unified sampling approaches with clear definition of clinical isolates, and a national sampling program using established protocols for cecal and retail isolates. Importantly, all cecal isolates came from slaughter sampling of cecal pouches of market-aged commercial birds, and likely represented the US turkey and broiler populations due to use of a comprehensive USDA FSIS sampling program.

Similar to Mageiros et al., we found that APEC plasmid genes were indeed highly prevalent not only in clinical isolates, but also in cecal and retail isolates. Using the previously established pentaplex PCR (predicted via WGS data) for APEC (Johnson et al., 2008), the number of APEC − classified as have 4 or more of the genes *iss, iroN, hlyF, ompTp*, and *iutA* − was 81.1% versus 65.5% in Turkey Clinical versus Turkey Cecal isolates, respectively. This appears to be in line with the results obtained from Mageiros et al. (Mageiros et al., 2021). Importantly, these genes and almost all of the other APEC plasmid-associated genes were still of significantly higher prevalence in Turkey Clinical versus Turkey Cecal populations. However, the relatively high prevalence of APEC plasmid genes in Turkey Cecal and Chicken Cecal isolates indicates that they may be less useful as a diagnostic tool for discriminating highly virulent APEC. However, we observed striking differences with respect to phylotypes of Turkey Clinical versus Turkey Cecal isolates, and the data analyzed from chicken-source samples suggests a similar pattern. In contrast to Mageiros et al., we found that Turkey Clinical isolates were heavily dominated by the B2 and C phylotypes, which were relatively rare in Turkey Cecal isolates. This contrasts not only Mageiros et al., but also multiple previous studies where there were less dramatic distinctions between populations with respect to phylotype, and notably proportionally fewer clinical isolates belonging to the B2 phylotype (Rodriguez-Siek et al., 2005). It is unclear if this represents a temporal shift in APEC populations across poultry production in the US, or is simply reflective of inherent differences between true APEC from commercial turkey versus commercial broiler production systems.

Both the present study and the study by Mageiros et al. (Mageiros et al., 2021) contrasts previous reports of a relatively low proportion of *E. coli* isolates from healthy birds possessing APEC plasmid PAI genes (Rodriguez-Siek et al., 2005; Stromberg et al., 2017). The reasons for this discrepancy are currently unknown. One possibility is that the majority of previous studies have used either fecal droppings or environmental samples as a source for such isolates, whereas the current study utilized cecal samples. It is also possible that differences between *E. coli* populations exist across the avian intestinal tract, but this remains to be determined. Another possibility is temporal and geographical fluctuations in *E. coli* populations in poultry − specifically, that *E. coli* in poultry as a whole have changed over time to possess the APEC plasmid at higher prevalence. Again, this remains to be determined, although analysis of the data from this study indicates that populations have not changed substantially over the past 5 years (data not shown).

The results of the challenge experiments in turkey and chicken embryos indicate that the simple use of the presence of APEC plasmid PAI genes alone is not sufficient to fully discriminate between higher versus lower virulence *E. coli* clones. This is supported by 2 lines of evidence. First, in STs such as ST58, isolates which contained a high number of APEC plasmid PAI genes (33−35) were still not lethal in the ELA models. This, and our genomic screens, indicate that dominant gut strains commonly possess genes of the APEC plasmid PAI but are not particularly virulent toward birds. Second, within dominant clinical STs, isolates that would all be classified as APEC solely using the 5-gene APEC plasmid PAI scheme differed in their lethality toward embryos (e.g., ST117 and ST428). With that said, our

data and the data of others clearly shows that the presence of the APEC plasmid or its genes enhances virulence in certain clonal backgrounds (Skyberg et al., 2006; Skyberg et al., 2008; Tivendale et al., 2009). However, some clones are highly virulent even when they lack APEC plasmid PAI genes, such as ST23. Together, the evidence indicates that the best means for poultry producers to identify clones of higher virulence, and thus higher risk, needs to include the presence of the APEC plasmid in combination with clone-specific markers.

Using the information gleaned from the genomic and phenotypic experiments in this study, we propose here the use of a revised approach to APEC typing in poultry. Rather than focusing solely on specific sets of virulence and fitness factors, we propose the use of markers of the APEC plasmid combined with ST- or serogroup-specific genomic background as an improved tool. The presence of the 2 APEC plasmid markers (*hlyF* and *ompTp*), which are among the most highly conserved of the plasmid, plus the presence of one of the clonal background-specific markers (ST23, ST117+O78, ST131, ST355, ST428, or O78), would indicate presence of a high-risk APEC and provide additional data on clonal type. It is clear from this and other studies that the APEC plasmid is highly prevalent in both broiler and turkey clinical *E. coli* isolates, and previous work has demonstrated a clear role in fitness and virulence in birds and persistence within poultry barns. However, this alone may not identify highly virulent and thus "high-risk" APEC clones. Combining the presence of the APEC plasmid with clinically-dominating STs and the O78 serogroup, based on our results, provides more definitive proof that a strain can be classified as APEC with high virulence potential.

This study is not without limitations. While we performed a comprehensive screen of clinical isolates from commercial turkeys, we relied on previously published work on broiler clinical isolates to compare and contrast dominant clonal groups in broiler colibacillosis. The broiler-focused studies ranged in dates of isolation, geography, syndrome, and methods. Therefore, they were not as controlled as the current study and may not be completely reflective of the broiler landscape, particularly in the USA. However, while we did not include broiler clinical isolates in this study, a wealth of literature exists for these *E. coli* populations. Multiple studies have reported the same dominant ST types in broiler clinical populations as those reported here for Turkey Clinical isolates, including ST23, ST117, ST131, ST155, ST355, and ST428 (Danzeisen et al., 2013; Hussein et al., 2013; Pires-Dos-Santos et al., 2013; Maluta et al., 2014; Braga et al., 2016; Heidemann Olsen et al., 2016; Poulsen et al., 2018; Chen et al., 2021; Mageiros et al., 2021; Mehat et al., 2021). Furthermore, the broiler cecal and retail populations analyzed here support the presence of these same high-risk clones within broiler production, and ELA data indicates that they are similarly virulent in broilers compared to commercial turkeys. The ELA model is a good model for rapid screen of *E. coli* virulence potential in birds, and

has previously shown correlation with live bird challenge models (Gibbs and Wooley, 2003). With that said, the embryo is an innately different challenge environment than the live bird. Also, the ELA work done here was intended as descriptive and, as such, the relatively low number of isolates used for each ST prevented us from running formal statistical analyses. Therefore, additional future work should be performed with larger numbers of isolates, and to determine if these dominant clones behave differently using a live bird respiratory challenge. Finally, while this work identifies an improved method by which to identify high-risk APEC, it does not explain the underlying reasons as to why these clones are of higher virulence. Such work is critically important, and the present study adds to the body of literature that can be used to further characterize those factors.

While the focus of this revised APEC typing scheme is based on comprehensive data identifying dominant high-risk clones of interest, it will undoubtedly not identify all highly virulent APEC. We have focused on dominant APEC STs in developing this revised scheme, as they are the most likely to pose problems in poultry production and our lethality data confirms their enhanced virulence. Other STs will likely arise in the future, and some strains within STs may become more virulent or successful than others. Thus, this scheme will likely need revisions as the landscape of APEC evolves. However, comprehensive analyses of the current APEC landscape indicate that this revised scheme will detect strains of highest risk to poultry health.

ACKNOWLEDGMENTS

DISCLOSURES

The authors declare no conflicts of interest.

SUPPLEMENTARY MATERIALS

Supplementary material associated with this article can be found in the online version at doi:10.1016/j. psj.2022.102009.

REFERENCES

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. J. Mol. Biol. 215:403–410.

Bankevich, A., S. Nurk, D. Antipov, A. A. Gurevich, M. Dvorkin, A. S. Kulikov, V. M. Lesin, S. I. Nikolenko, S. Pham, A. D. Prjibelski, A. V. Pyshkin, A. V. Sirotkin, N. Vyahhi, G. Tesler, M. A. Alekseyev, and P. A. Pevzner. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19:455–477.

Bessonov, K., C. Laing, J. Robertson, I. Yong, K. Ziebell, V. P. J. Gannon, A. Nichani, G. Arya, J. H. E. Nash, and S. Christianson. 2021. ECTyper: in silico Escherichia coli serotype and species prediction from raw and assembled whole-genome sequence data. Microb. Genom. 7:000728.

Bolger, A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120.

Braga, J. F. V., N. K. Chanteloup, A. Trotereau, S. Baucheron, R. Guabiraba, R. Ecco, and C. Schouler. 2016. Diversity of Escherichia coli strains involved in vertebral osteomyelitis and arthritis in broilers in Brazil. BMC Vet. Res. 12:140.

Brynildsrud, O., J. Bohlin, L. Scheffer, and V. Eldholm. 2016. Rapid scoring of genes in microbial pan-genome-wide association studies with Scoary. Genome Biol. 17:238.

Camacho, C., G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer, and T. L. Madden. 2009. BLAST+: architecture and applications. BMC Bioinform. 10:421.

Chen, X., W. Liu, H. Li, S. Yan, F. Jiang, W. Cai, and G. Li. 2021. Whole genome sequencing analysis of avian pathogenic Escherichia coli from China. Vet. Microbiol. 259:109158.

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

Cordoni, G., M. J. Woodward, H. Wu, M. Alanazi, T. Wallis, and R. M. La Ragione. 2016. Comparative genomics of European avian pathogenic E. coli (APEC). BMC Genomics 17:960.

Danzeisen, J. L., Y. Wannemuehler, L. K. Nolan, and T. J. Johnson. 2013. Comparison of multilocus sequence analysis and virulence genotyping of Escherichia coli from live birds, retail poultry meat, and human extraintestinal infection. Avian Dis. 57:104–108.

Dho-Moulin, M., and J. M. Fairbrother. 1999. Avian pathogenic Escherichia coli (APEC). Vet. Res. 30:299–316.

Ewers, C., T. Janssen, S. Kiessling, H. C. Philipp, and L. H. Wieler. 2004. Molecular epidemiology of avian pathogenic Escherichia coli (APEC) isolated from colisepticemia in poultry. Vet. Microbiol. 104:91–101.

Ewers, C., T. Janssen, S. Kiessling, H. C. Philipp, and L. H. Wieler. 2005. Rapid detection of virulence-associated genes in avian pathogenic Escherichia coli by multiplex polymerase chain reaction. Avian Dis. 49:269–273.

Gibbs, P. S., and R. E. Wooley. 2003. Comparison of the intravenous chicken challenge method with the embryo lethality assay for studies in avian colibacillosis. Avian Dis. 47:672–680.

Gurevich, A., V. Saveliev, N. Vyahhi, and G. Tesler. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075.

Heidemann Olsen, R., M. Bisgaard, J. P. Christensen, S. Kabell, and H. Christensen. 2016. Pathology and molecular characterization of Escherichia coli associated with the avian salpingitis-peritonitis disease syndrome. Avian Dis. 60:1–7.

Hussein, A. H., I. A. Ghanem, A. A. Eid, M. A. Ali, J. S. Sherwood, G. Li, L. K. Nolan, and C. M. Logue. 2013. Molecular and phenotypic characterization of Escherichia coli isolated from broiler chicken flocks in Egypt. Avian Dis. 57:602–611.

Janben, T., C. Schwarz, P. Preikschat, M. Voss, H. C. Philipp, and L. H. Wieler. 2001. Virulence-associated genes in avian pathogenic Escherichia coli (APEC) isolated from internal organs of poultry having died from colibacillosis. Int. J. Med. Microbiol. 291:371–378.

Johnson, T. J., S. J. Johnson, and L. K. Nolan. 2006a. Complete DNA sequence of a ColBM plasmid from avian pathogenic Escherichia coli suggests that it evolved from closely related ColV virulence plasmids. J. Bacteriol. 188:5975–5983.

Johnson, T. J., S. Kariyawasam, Y. Wannemuehler, P. Mangiamele, S. J. Johnson, C. Doetkott, J. A. Skyberg, A. M. Lynne, J. R. Johnson, and L. K. Nolan. 2007a. 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.

Johnson, T. J., C. M. Logue, Y. Wannemuehler, S. Kariyawasam, C. Doetkott, C. DebRoy, D. G. White, and L. K. Nolan. 2009. Examination of the source and extended virulence genotypes of Escherichia coli contaminating retail poultry meat. Foodborne Pathog. Dis. 6:657–667.

Johnson, T. J., K. E. Siek, S. J. Johnson, and L. K. Nolan. 2006b. DNA sequence of a ColV plasmid and prevalence of selected plasmid-encoded virulence genes among avian Escherichia coli strains. J. Bacteriol. 188:745–758.

Johnson, T. J., Y. Wannemuehler, C. Doetkott, S. J. Johnson, S. C. Rosenberger, and L. K. Nolan. 2008. Identification of minimal predictors of avian pathogenic Escherichia coli virulence for use as a rapid diagnostic tool. J. Clin. Microbiol. 46:3987–3996.

Johnson, T. J., Y. M. Wannemuehler, S. J. Johnson, C. M. Logue, D. G. White, C. Doetkott, and L. K. Nolan. 2007b. Plasmid replicon typing of commensal and pathogenic Escherichia coli isolates. Appl. Environ. Microbiol. 73:1976–1983.

Jolley, K. A., and M. C. Maiden. 2010. BIGSdb: Scalable analysis of bacterial genome variation at the population level. BMC Bioinform. 11:595.

Kalyaanamoorthy, S., B. Q. Minh, T. K. F. Wong, A. von Haeseler, and L. S. Jermiin. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Meth. 14:587–589.

Kaper, J. B., J. P. Nataro, and H. L. Mobley. 2004. Pathogenic Escherichia coli. Nat. Rev. Microbiol. 2:123–140.

Kawano, M., K. Yaguchi, and R. Osawa. 2006. Genotypic analyses of Escherichia coli isolated from chickens with colibacillosis and apparently healthy chickens in Japan. Microbiol. Immunol. 50:961–966.

Letunic, I., and P. Bork. 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res. 44:W242–W245.

Mageiros, L., G. Meric, S. C. Bayliss, J. Pensar, B. Pascoe, E. Mourkas, J. K. Calland, K. Yahara, S. Murray, T. S. Wilkinson, L. K. Williams, M. D. Hitchings, J. Porter, K. Kemmett, E. J. Feil, K. A. Jolley, N. J. Williams, J. Corander, and S. K. Sheppard. 2021. Genome evolution and the emergence of pathogenicity in avian Escherichia coli. Nat. Commun. 12:765.

Malberg Tetzschner, A. M., J. R. Johnson, B. D. Johnston, O. Lund, and F. Scheutz. 2020. In silico genotyping of Escherichia coli isolates for extraintestinal virulence genes by use of whole-genome sequencing data. J. Clin. Microbiol. 58:e01269-20.

Maluta, R. P., C. M. Logue, M. R. Casas, T. Meng, E. A. Guastalli, T. C. Rojas, A. C. Montelli, T. Sadatsune, M. de Carvalho Ramos, L. K. Nolan, and W. D. da Silveira. 2014. Overlapped sequence types (STs) and serogroups of avian pathogenic (APEC) and human extra-intestinal pathogenic (ExPEC) Escherichia coli isolated in Brazil. PLoS One 9:e105016.

McPeake, S. J., J. A. Smyth, and H. J. Ball. 2005. Characterisation of avian pathogenic Escherichia coli (APEC) associated with colisepticaemia compared to faecal isolates from healthy birds. Vet. Microbiol. 110:245–253.

Mehat, J. W., A. H. M. van Vliet, and R. M. La Ragione. 2021. The avian pathogenic Escherichia coli (APEC) pathotype is comprised of multiple distinct, independent genotypes. Avian Pathol. 50:402–416.

Nguyen, L. T., H. A. Schmidt, A. von Haeseler, and B. Q. Minh. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32:268–274.

Nolan, L., J. P. Vaillancourt, N. L. Barbieri, and C. M. Logue. 2020. Colibacillosis. Diseases of Poultry. D. E. Swayne, ed. Wiley, Ames, IA.

Pires-Dos-Santos, T., M. Bisgaard, and H. Christensen. 2013. Genetic diversity and virulence profiles of Escherichia coli causing salpingitis and peritonitis in broiler breeders. Vet. Microbiol. 162:873–880.

Poulsen, L. L., M. Bisgaard, S. L. Jorgensen, T. Dideriksen, J. R. Pedersen, and H. Christensen. 2018. Characterization of *Escherichia coli* causing cellulitis in broilers. Vet. Microbiol. 225:72–78.

Provence, D. L., and R. Curtiss 3rd. 1992. Role of *crl* in avian pathogenic *Escherichia coli*: a knockout mutation of *crl* does not affect hemagglutination activity, fibronectin binding, or Curli production. Infect. Immun. 60:4460–4467.

Rodriguez-Siek, K. E., C. W. Giddings, C. Doetkott, T. J. Johnson, and L. K. Nolan. 2005. Characterizing the APEC pathotype. Vet. Res. 36:241–256.

Ronco, T., M. Stegger, R. H. Olsen, C. Sekse, A. B. Nordstoga, T. Pohjanvirta, B. Lilje, U. Lyhs, P. S. Andersen, and K. Pedersen. 2017. Spread of avian pathogenic *Escherichia coli* ST117 O78:H4 in Nordic broiler production. BMC Genom. 18:13.

Seemann, T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069.

Senchenkova, S. N., X. Guo, O. I. Naumenko, A. S. Shashkov, A. V. Perepelov, B. Liu, and Y. A. Knirel. 2016. Structure and genetics of the O-antigens of *Escherichia coli* O182-O187. Carbohydr. Res. 435:58–67.

Sitto, F., and F. U. Battistuzzi. 2020. Estimating pangenomes with Roary. Mol. Biol. Evol. 37:933–939.

Skyberg, J. A., T. J. Johnson, J. R. Johnson, C. Clabots, C. M. Logue, and L. K. Nolan. 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.

Skyberg, J. A., T. J. Johnson, and L. K. Nolan. 2008. Mutational and transcriptional analyses of an avian pathogenic *Escherichia coli* ColV plasmid. BMC Microbiol. 8:24.

Stromberg, Z. R., J. R. Johnson, J. M. Fairbrother, J. Kilbourne, A. Van Goor, R. R. Curtiss, and M. Mellata. 2017. Evaluation of *Escherichia coli* isolates from healthy chickens to determine their potential risk to poultry and human health. PLoS One 12: e0180599.

Team, R. C. 2020. R: A language and environment for statistical computing. https://www.R-project.org/2022.

Tivendale, K. A., J. L. Allen, and G. F. Browning. 2009. Plasmid-borne virulence-associated genes have a conserved organization in virulent strains of avian pathogenic *Escherichia coli*. J. Clin. Microbiol. 47:2513–2519.

Vandekerchove, D., F. Vandemaele, C. Adriaensen, M. Zaleska, J. P. Hernalsteens, L. De Baets, P. Butaye, F. Van Immerseel, P. Wattiau, H. Laevens, J. Mast, B. Goddeeris, and F. Pasmans. 2005. Virulence-associated traits in avian *Escherichia coli*: comparison between isolates from colibacillosis-affected and clinically healthy layer flocks. Vet. Microbiol. 108:75–87.

Walters, J. 2014. Characterization of Atypical Hemolytic *Ornithobacterium rhinotracheale* Isolates and Comparison With the Normal Non-Hemolytic Phenotype. Virginia Polytechnic Institute and State University, Blacksburg, VA PhD.

Wang, S., Q. Meng, J. Dai, X. Han, Y. Han, C. Ding, H. Liu, and S. Yu. 2014. Development of an allele-specific PCR assay for simultaneous sero-typing of avian pathogenic *Escherichia coli* predominant O1, O2, O18 and O78 strains. PLoS One 9:e96904.

Waters, N. R., F. Abram, F. Brennan, A. Holmes, and L. Pritchard. 2020. Easy phylotyping of *Escherichia coli* via the EzClermont web app and command-line tool. Access Microbiol. 2: acmi000143.

Wooley, R. E., P. S. Gibbs, T. P. Brown, and J. J. Maurer. 2000. Chicken embryo lethality assay for determining the virulence of avian *Escherichia coli* isolates. Avian Dis 44:318–324.

Yaguchi, K., T. Ogitani, R. Osawa, M. Kawano, N. Kokumai, T. Kaneshige, T. Noro, K. Masubuchi, and Y. Shimizu. 2007. Virulence factors of avian pathogenic *Escherichia coli* strains isolated from chickens with colisepticemia in Japan. Avian Dis. 51:656–662.

Zhou, Z., N. F. Alikhan, K. Mohamed, Y. Fan, G. Agama Study, and M. Achtman. 2020. The EnteroBase user's guide, with case studies on *Salmonella* transmissions, *Yersinia pestis* phylogeny, and *Escherichia* core genomic diversity. Genome Res. 30:138–152.

Zhou, Z., N. F. Alikhan, M. J. Sergeant, N. Luhmann, C. Vaz, A. P. Francisco, J. A. Carrico, and M. Achtman. 2018. GrapeTree: visualization of core genomic relationships among 100,000 bacterial pathogens. Genome Res. 28:1395–1404.