



Article Tracing the Evolutionary Pathways of Serogroup O78 Avian Pathogenic Escherichia coli

Eun-Jin Ha^{1,2}, Seung-Min Hong^{1,2}, Seung-Ji Kim^{1,2}, Sun-Min Ahn², Ho-Won Kim², Kang-Seuk Choi^{1,2,*} and Hyuk-Joon Kwon^{2,3,4,5,*}

- ¹ Laboratory of Avian Diseases, Department of Farm Animal Medicine, College of Veterinary Medicine and BK21 PLUS for Veterinary Science, Seoul National University, Seoul 088026, Republic of Korea; flower678@snu.ac.kr (E.-J.H.); topkin@snu.ac.kr (S.-M.H.); seungji910@snu.ac.kr (S.-J.K.)
- ² Research Institute for Veterinary Science, College of Veterinary Medicine, Seoul 08826, Republic of Korea; vicky.ahn@snu.ac.kr (S.-M.A.); iamkhw52@snu.ac.kr (H.-W.K.)
- ³ Laboratory of Poultry Medicine, Department of Farm Animal Medicine, College of Veterinary Medicine and BK21 PLUS for Veterinary Science, Seoul National University, Seoul 088026, Republic of Korea
- ⁴ Farm Animal Clinical Training and Research Center (FACTRC), GBST, Seoul National University, Pyeongchang 25354, Republic of Korea
- ⁵ GeNiner Inc., Seoul 08826, Republic of Korea
- * Correspondence: kchoi0608@snu.ac.kr (K.-S.C.); kwonhj01@snu.ac.kr (H.-J.K.); Tel.: +82-2-880-1266 (H.J.-K.)

Abstract: Avian pathogenic E. coli (APEC) causes severe economic losses in the poultry industry, and O78 serogroup APEC strains are prevalent in chickens. In this study, we aimed to understand the evolutionary pathways and relationships between O78 APEC and other E. coli strains. To trace these evolutionary pathways, we classified 3101 E. coli strains into 306 subgenotypes according to the numbers and types of single nucleotide polymorphisms (RST0 to RST63-1) relative to the consensus sequence (RST0) of the RNA polymerase beta subunit gene and performed network analysis. The E. coli strains showed four apparently different evolutionary pathways (I-1, I-2, I-3, and II). The thirtytwo Korean O78 APEC strains tested in this study were classified into RST4-4 (45.2%), RST3-1 (32.3%), RST21-1 (12.9%), RST4-5 (3.2%), RST5-1 (3.2%), and RST12-6 (3.2%), and all RSTs except RST21-1 (I-2) may have evolved through the same evolutionary pathway (I-1). A comparative genomic study revealed the highest relatedness between O78 strains of the same RST in terms of genome sequence coverage/identity and the spacer sequences of CRISPRs. The early-appearing RST3-1 and RST4-4 prevalence among O78 APEC strains may reflect the early settlement of O78 E. coli in chickens, after which these bacteria accumulated virulence and antibiotic resistance genes to become APEC strains. The zoonotic risk of the conventional O78 APEC strains is low at present, but the appearance of genetically distinct and multiple virulence gene-bearing RST21-1 O78 APEC strains may alert us to a need to evaluate their virulence in chickens as well as their zoonotic risk.

Keywords: *Escherichia* evolution; avian pathogenic *E. coli; rpoB* sequence typing; network analysis; comparative genomics; CRISPR spacer; molecular prophage typing; virulence and antibiotic resistance genes; bacteriocin

1. Introduction

Avian pathogenic *E. coli* (APEC) is an extraintestinal pathogenic *E. coli* (ExPEC) that causes severe economic loss in the poultry industry and increases public health concerns [1,2]. The different compositions and linkages of carbohydrates determine the O serogroups of *E. coli*, and to date, 185 O serogroups have been identified [3]. The high diversity of O serogroups in *E. coli* strains may be partially due to escape mutations and selection resulting from the detrimental infection of lytic bacteriophages, with O serogroup changes occurring through various mechanisms, including the homologous recombination of newly introduced O-antigen gene clusters (O-AGCs) [3]. Although the frequencies of O



Citation: Ha, E.-J.; Hong, S.-M.; Kim, S.-J.; Ahn, S.-M.; Kim, H.-W.; Choi, K.-S.; Kwon, H.-J. Tracing the Evolutionary Pathways of Serogroup O78 Avian Pathogenic *Escherichia coli*. *Antibiotics* **2023**, *12*, 1714. https:// doi.org/10.3390/antibiotics12121714

Academic Editor: Teresa V. Nogueira

Received: 27 September 2023 Revised: 17 November 2023 Accepted: 4 December 2023 Published: 9 December 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). serogroups of APEC strains depend on farms and countries, O78 is one of the most frequent serogroups in the world [4,5].

Various molecular methods, such as multilocus sequence typing (MLST), pulsed-field gel electrophoresis (PFGE), phylogrouping, *rpoB* sequence typing (RSTing), molecular prophage typing, and profiling spacer sequences of the CRISPR–Cas (clustered regularly interspaced short palindromic repeats and CRISPR-associated genes) system, have been applied to differentiate *E. coli* strains and elucidate the molecular epidemiology of outbreaks of pathogenic *E. coli* infections [6–9]. Core genome-based MLST (cgMLST) is applied to elucidate evolutionary relationships between *E. coli* strains, other *Escherichia* species, and *Salmonella enterica* serovars, but the strategy of using multiple genes with different evolutionary statuses may complicate the analysis of the results [10]. Pangenome analyses have provided new insights into the evolution of species, antibiotic resistance, and the pathogenicity of bacteria [11]. However, simpler and more cost-effective first-line methods are needed.

The RNA polymerase beta subunit (RpoB) is an enzymatic component of RNA polymerase. RNA polymerase plays biologically and evolutionarily important roles, including pivotal roles in decoding gene information into mRNAs for protein translation and the transcription of other important RNAs. Therefore, RpoB is one of the key molecules in which the evolutionary histories of the present united DNA/RNA/protein worlds are embedded. The applicability of the *rpoB* gene as a chronological molecule to understand the progenitor-progeny relationships of bacteria was hinted at in *Staphylococcus aureus* (S. aureus) strains [12,13]. Humans are natural hosts of S. aureus, and the bacterium spills over into other animals via adaptation based on missense mutations in essential genes and the acquisition/loss of genes [14–17]. The major human strains have only two (rpoB sequence type, RST 2-1) or four (RST4-1) mutations compared to the consensus *rpoB* sequence, which is a hypothetical progenitor sequence. Although the *rpoB* genes of the RST2-1 and RST4-1 strains have not changed considerably from the consensus sequence (RST0), they have acquired various prophages and genomic islands to increase genome sizes. Therefore, the commensal life of *S. aureus* in humans may not cause the accumulation of multiple mutations in *rpoB*. In this context, the presence and predominance of early-appearing RSTs in certain hosts may reflect the early settlement of bacterial species in terms of their evolution. Currently, Shigella species are considered toxigenic variants of E. coli, and E. fergusonii and E. albertii are pathogenic and are not easily differentiated from E. coli due to biological similarities [18–20]. Additionally, Salmonella enterica is genetically closely related to E. coli [21]. While E. coli genomes are subject to changes in size and gene contents, some host-adapted Salmonella genomes are suppressed to such modifications by other genetic circuits and by the robustness and redundancy of the genomic network, making them less variable [22–24]. Accordingly, a trial to understand the progenitor–progeny relationship between E. coli and other Escherichia and Salmonella species may be valuable. Therefore, a *rpoB*-based approach to unravel the genetic relationships among them may be of interest.

Due to the similarity between APEC and human ExPEC, such as uropathogenic *E. coli* and neonatal meningitis *E. coli*, the zoonotic risks of APEC need to be addressed. Genomic relatedness between O1 APEC and human ExPEC has been suspected, but the zoonotic risk of O78 APEC still needs to be further determined [25,26]. The cumulative acquisition of virulence and antibiotic-resistance genes in pathogenic *E. coli* increases public health concerns and needs to be elucidated more clearly.

In this study, we modified our previous RSTing approach for *E. coli* similarly to the *S. aureus* RSTing scheme [6,12]. We performed network analysis with single nucleotide polymorphisms (SNPs) of representative RSTs, including *Shigella* strains and selected strains of other *Escherichia* species and *Salmonella enterica*. We typed the RSTs of O78 APEC strains recorded in Korea during 2012–2020 and characterized their virulence, antibiotic resistance, and microcin/colicin gene contents. Furthermore, we performed a comparative genomics analysis of the identified major RSTs of O78 APEC strains to demonstrate the evolutionary relationships between them and other *E. coli* strains of the same and different RSTs.

2. Results and Discussion

2.1. O78 APEC Strains Identified during 2012–2020

A total of 31 of the O78 APEC strains accounted for 19.0% of the 163 APEC strains identified during 2012–2020, and they were isolated from birds of various ages and poultry types: broiler breeders (45.2%), commercial layers (32.3%), broilers (19.4%), and Korean native chickens (3.2%). Phylogenetic group B1 was most predominant (51.6%), followed by C (25.8%), E (9.7%), F (9.7%), and A (3.2%) (Table 1).

2.2. Modified rpoB Sequence Typing (RSTing) of E. coli Strains

In this study, we modified our previous RSTing scheme by reclassifying RSTs according to the number of nucleotide differences from the consensus sequence (RST0) [6,12]. A total of 3101 *E. coli* strains, comprising 3029 sourced from the GenBank database and an additional 32 O78 APEC strains (including the previously reported E123 strain) along with 40 APEC strains of other serotypes (in manuscript preparation), were classified into 306 RSTs, spanning from RST2-1 to RST63-1 [6]. The predominant RSTs identified were RST3-1, accounting for 17.4%, followed by RST6-1 at 12.7%, among others (Table S1). The thirty-one O78 strains were classified into RST4-4 (45.2%), RST3-1 (32.3%), RST21-1 (12.9%), RST4-5 (3.2%), RST5-1 (3.2%), and RST12-6 (3.2%), and E123 was classified as RST3-1 (Table 1). We collected O78 reference strains using BLAST searches with the nucleotide sequence of the O-AGC of the E123 strain, and their RSTs and serotypes are summarized in Table S2. The twenty O78 reference strains were classified into RST3-1 (70%), RST6-1 (15%), RST4-30 (5%), RST7-3 (5%), and RST21-1 (5%), and including our 32 O78 strains, RST3-1 (48.1%, 25/52) was most prevalent, followed by RST4-4 (26.9%, 14/52) and RST21-1 (9.6%, 5/52).

Shigella species are now considered *E. coli* sensu lato and are thought to have evolved in a convergent manner by acquiring Shiga toxin genes [21]. Some *Shigella* strains are classified into *E. coli* RSTs, including RST3-1, but others are classified into the unique *Shigella* RSTs listed in Table S3. The results may reflect the early and continuous acquisition of *Shigella* phenotypes during *E. coli* evolution. Among 306 RSTs, 68 RSTs (22.2%) possessed missense mutations, and O78 strains of RST4-30 (Y47C) and RST12-6 (T141S-F195L-V196A-L204P-L213V-K422R) possessed different mutations from rifampin resistance-related mutations (500-575 and V146F/W) (Table S1) [27]. The majority of *E. coli* strains among the early RSTs, RST2-1, and RST3-1, originated from humans, e.g., 67.1% of RST2-1 strains (Tables S4 and S5).

2.3. Hypothetical Evolutionary Pathways of E. coli and O78 APEC Strains

Using the SNPs of RSTs, network analysis was performed to understand the evolutionary pathways of E. coli, E. fergusonii, E. albertii, E. marmotae, and S. enterica serovar Indiana (S. Indiana) (Figure 1). Two branches, I and II, came from RST0, and branch I was divided into four subbranches: I-1, I-2, I-3, and I-4. Subbranch I-1 contained prevalent RST3-1, RST6-1, and all of the O78 RSTs except for RST21-1. The farthest RST from RST0 in subbranch I-1 was RST13-7. Branch I-2 contained the O78 strain containing RST21-1, and the farthest RST was RST62-1. Interestingly, an E. marmotae strain (RST106) emerged from branch I-2. Branch I-3 contained RST27-1, and the farthest RST was RST48-1. Branch I-4 was short and contained only RST9-1 and RST13-1. Branch II contained prevalent RST24-1 and RST20-1, and the farthest RST was RST63-1. Interestingly, E. albertii (RST69, RST100, and RST104), E. fergusonii (RST40, RST53, and RST60), and S. Indiana (RST177) came from branch II. Considering the low RST numbers and the identity percentages (98.5-99.1%) of E. fergusonii strains, they may be more closely related to E. coli strains than other *Escherichia* species. *E. fergusonii* and *E. albertii* are emerging pathogens and are difficult to differentiate from E. coli [18]. Therefore, the present RSTing scheme may help to identify them. Salmonella enterica is known to be most closely related among the different bacterial genera, and our result showing the evolutionary pathway sharing of S. Indiana and some E. coli strains may be interesting to study further [21,28]. We also performed phylogenetic analysis using nucleotide sequences of *rpoB* utilizing Bayesian inference (Figure S1). The clustering pattern and topology are similar to network analysis, but network analysis intuitively shows progenitor and progeny relationships between RST0 and other RSTs.

Strain	Year of Isolation	Chicken Type ^a	Age (Day-Old)	Phylo- Group	RST ^b	No. of Virulence Genes	Strain	Year of Isolation	Chicken Type ^a	Age (Day-Old)	Phylo- Group	RST	No. of Virulence Genes
E12049	2012	KNC	40	F	21-1	20	E16015	2016	В	24	B1	5-1	15
E14033	2014	L	76	F	21-1	19	E17039	2017	BB	263	B1	12-6	15
E12090	2012	L	124	F	21-1	18	E19033	2019	BB	134	B1	4-4	15
E18005	2018	BB	209	B1	4-4	17	E19057	2019	L	148	С	3-1	14
E16018	2016	L	451	B1	4-4	16	E16023	2016	BB	203	Е	3-1	14
E18016	2018	BB	163	B1	4-4	16	E16039	2016	L	252	B1	4-4	14
E18027	2018	В	2	Е	21-1	16	E19024	2019	L	315	С	3-1	14
E19014	2019	BB	7	B1	4-4	16	E20019	2020	L	310	Е	3-1	14
E19025	2019	В	7	B1	4-4	16	E17001	2017	BB	264	С	3-1	13
E19034	2019	BB	235	B1	4-4	16	E12050	2012	В	26	С	3-1	12
E19045	2019	L	19	B1	4-4	16	E13028	2013	BB	29	С	3-1	12
E19070	2019	BB	170	B1	4-4	16	E20042	2020	L	161	С	3-1	11
E12091	2012	BB	236	B1	4-4	15	E19013	2019	BB	77	С	4-5	9
E15016	2015	В	30	B1	4-4	15	E19040	2019	BB	195	С	3-1	9
E15026	2015	BB	273	B1	4-4	15	E19056	2019	В	3	А	3-1	9
E16011	2016	L	136	B1	4-4	15	E123	2003	NA ^c	NA	С	3-1	17

^a L: layer chicken; B: broiler chicken; BB: broiler breeder chicken; KNC: Korean native chicken. ^b RST: *rpoB* sequence type. ^c NA, not available.



Figure 1. Median-Joining Network of *E. coli* RSTs. The evolutionary tracing of a total of 3029 *E. coli* strains was conducted based on *rpoB* SNP haplotypes. This study also includes representatives of other species within the genus *Escherichia* as well as the genus *Salmonella*, which are labeled using orange text. The median-joining network method was employed for this analysis, and the network was constructed using popART software (v.1.7) [29,30]. RST numbers were derived from the consensus sequence (RST0), ranging from RST2-1 to RST63-1. Two main branches were identified: branch I, comprising subbranches 1-4, and branch II. RST3-1 and RST4-4 were located within subbranch I-1, while RST21-1 was classified under subbranch I-2. The frequency of RSTs was reflected by variations in the colors and sizes of the circles within the network. A blue border indicates that the hosts of the RSTs are birds, a red border signifies human hosts, and a yellow border represents a mixed host category that includes the environment, humans, and other animals. Black dots are median vectors that can be biologically interpreted as extant unsampled sequences or extinct ancestral sequences [29].

The *rpoB* gene sizes of major gram-positive and gram-negative bacteria were compared, and the RSTs of additional bacteria with the same size of *rpoB* as *E. coli* are summarized in Table 2. The *rpoB* gene sizes of the gram-positive bacteria are smaller than those of the gram-negative bacteria and increase in order from 3531 bp to 3717 bp. In the case of gram-negative bacteria, the lengths of *rpoB* varied from shorter (4026 bp) to longer (4239 bp) than those of *E. coli*. Thus, the differences in the nucleotide sequences and the length of *rpoB* may provide a hint of the evolutionary direction and order of important pathogenic bacteria.

Pathogenic Bacteria	Strain	Accession No.	Size of rpoB	RST	Identity (%)
Corynebacterium diphtheriae	31A	CP003206.1	3531	ND	ND
Bacillus anthracis	A2084	NC_007530.2	3534	ND	ND
Staphylococcus aureus	PMB81	CP03444.1	3552	ND	ND
Listeria monocytogenes	EGD-e	NC_003210.1	3555	ND	ND
Streptococcus pneumoniae	Hu17	NZ_CP020549.1	3612	ND	ND
Enterococcus faecalis	T5	NZ_KB944666.1	3615	ND	ND
Clostridium perfringens	CPI 18-6	NZ_CP075979.1	3705	ND	ND
Clostridium difficile	s-0253	NZ_CP076401.1	3717	ND	ND
Vibrio cholerae	RFB16	CP043554.1	4026	ND	ND
Escherichia coli	RST0	ND ^a	4029	0	100
Escherichia fergusonii	RHB10-C04	CP057918.1	4029	40	99.1
Escherichia fergusonii	ATCC 35471	CP042945.1	4029	53	98.7
Escherichia fergusonii	RHB02-C14	CP055872.1	4029	60	98.5
Escherichia albertii	RM10507	CP043258.1	4029	69	98.3
Escherichia albertii	ChinaSP140150	CP025676.1	4029	100	97.5
Escherichia albertii	BIA_36	CP117590.1	4029	104	97.4
Escherichia marmotae	W49-2	CP093239.1	4029	106	97.4
Salmonella enterica serovar Indiana	SI67	CP050783.1	4029	177	95.6
Salmonella bongori	N268-08	CP006608.1	4029	251	93.8
Klebsiella pneumoniae	XH210	CP052761.1	4029	259	93.6
Enterobacter cloacae	colR/S	CP010512.1	4029	264	93.4
Serratia marcescens	SmUNAM836	CP012685.1	4029	491	87.8
Yersinia pseudotuberculosis	IP2666pIB1	CP032566.1	4029	588	85.4
Yersinia pestis	FDAARGOS_603	CP033690.1	4029	590	85.4
Proteus mirabilis	1035	CP072779.1	4029	758	81.1
Pasteurella multocida	FDAARGOS_218	CP020405.2	4029	1027	74.5
Haemophilus influenzae	Hi375	CP009610.1	4032	ND	ND
Pseudomonas aeruginosa	PAO1	AE004091.2	4074	ND	ND
Acinetobacter baumannii	Ab421_GEIH-2010	CP014266.1	4074	ND	ND
Legionella pneumophila	C9	CP015941.1	4107	ND	ND
Bordetella bronchiseptica	NCTC10543	LR134326.1	4113	ND	ND
Neisseria gonorrheae	AT159	CP097846.1	4179	ND	ND
Moraxella catarrhalis	CCRI-195ME	CP018059.1	4239	ND	ND

Table 2. Comparison of *rpoB* gene size, identity, and *rpoB* sequence typing.

^a ND, not determined.

2.4. General Genomic Information of the Six Representative O78 APEC Strains from RST3-1, RST4-4, and RST21-1

The *E. coli* pangenome is highly variable, and the reference genome sequences of O79 APEC strains are not sufficient for resequencing. For these reasons, the complete genome sequences of the six representative O78 APEC strains classified into the three major RSTs, aRST3-1 (E19057 and E123), RST4-4 (E18005 and E19025), and RST21-1 (E12049 and E14033), were determined using de novo sequencing and compared with the corresponding reference strains, which were selected on the basis of an identical serogroup (PSUO78), high genome coverage/identity (NCTC11129), or both (APEC O78). Basic information on the genomes, such as their lengths, GC contents, numbers of CDSs, rRNAs, tRNAs and

CRISPRs, coding ratios (%), plasmids, and numbers of genomic islands, is summarized in Table 3. The determined genome sizes ranged from 4,885,187 bp (E19057) to 5,170,367 bp (E12049), and the GC ratio was 50.60–50.79%. The H antigen type, sequence type (ST) of MLST, and molecular prophage type (mPPT) of O78 strains were determined, and within the six strains, two CRISPRs, between 1 and 4 plasmids, and 69 and 94 genomic islands were identified (Table 3).

2.5. Correlation of Genomic Sequence Coverage/Identity and CRISPR Contents between the Same RSTs Containing Major O78 Strains

The whole-genome sequences of O78 strains were compared, and the coverage and nucleotide identities were measured. The coverage/identity between the same RSTs of the O78 APEC strains was highest, and the coverage/identity values of RST3-1, RST4-4, and RST21-1 were 97%/99.99%, 98%/99.99%, and 96%/99.98%, respectively (Figure 2). It is reasonable to deduce that the higher the genomic coverage/identity is, the closer the evolutionary distance.

Not all E. coli strains within the same RST exhibited a high degree of genomic similarity; some strains exhibited higher genetic congruence, while others showed less or much less genetic similarity. The early-appearing RSTs, such as RST2-1 and RST3-1, encompassed more than three (RST2-1_G1, G2, G3, etc.) and four (RST3-1_G1, G2, G3, G4, etc.) distinct genome coverage/identity groups, respectively (Figure 2; Tables S4 and S5). However, some strains showed relatively lower genome coverages, ranging from 87 to 93%, with high genome sequence identities ranging from 99.91 to 99.97% (RST2-1 G1) (Table S4). As the number of compared *E. coli* genomes increased, the number of core genes decreased to 753 [11]. The pangenome of *E. coli* is regarded as open and still evolving by gene acquisition and diversification [31]. The lower genome coverage may be due to the different accessory and unique genes that were acquired from different metagenomic environments surrounding the habitats of different E. coli. Therefore, the comparison of genome coverages/identities may be a simple way to evaluate evolutionary relatedness between strains before core and pangenome analyses. Interestingly, the RST21-1 O78 APEC strains revealed elevated genome identities (99.72-99.89%) exclusively with RST4-4 APEC strains but not with other RST4-4 strains from different countries (Figure 2). Therefore, recent genetic exchanges between them are suspected (Figure S2).

		1 0	1						
	E19057	E123	APEC O78	E18005	E19025	NCTC11129	E12049	E14033	PSUO78
Accession no.	CP126934.1	CP126955.1	CP004009.1	CP126946.1	CP126931.1	LR134222.1	CP126952.1	CP126948.1	CP012112.1
Origin	Chicken	Chicken	Turkey	Chicken	Chicken	unknown	Chicken	Chicken	Chicken
RST	3-1	3-1	3-1	4-4	4-4	4-4	21-1	21-1	21-1
MLST (Achman)	ST23	ST369	ST23	ST155	ST155	ST155	ST11749	ST48	ST5940
mPPT	2-5	2-3-4	3-4	1-2-3-5(2)-6	1-2-3(2)-4-5-6	2-3-4-5	1-2-5	1-2-4-5	2-4-5-7
Serotype	O78:H4	O78:H4	O78:H9	O78:H51	O78:H51	O8, O60:H51	O78:H4	O78:H4	O78:H4
Total Length (bp)	4,885,187	4,998,625	4,798,435	5,129,598	5,086,337	5,008,027	5,170,367	5,021,334	4,988,493
GC Content (%)	50.60%	50.64%	50.70%	50.64%	50.70%	50.78%	50.74%	50.79%	50.80%
No. of CDSs	4586	4690	4504	4771	4751	4697	4788	4658	4791
No. of rRNA	22	22	19	21	22	22	22	22	22
No. of tRNA	86	82	88	92	95	89	97	90	88
No. of CRISPRS	2	2	2	2	2	2	2	2	2
Coding ratio (%)	87.30%	87.10%	87.60%	87.00%	87.20%	87.10%	86.80%	87.20%	86.60%
Plasmid 1 ª	pEND_Eco19057-1 (CP126935.1) (IncI1-1 (alpha), 100%; 110,406 bp)	pEND_Eco 123-1 (CP126956.1) (IncFIB, 98.39%; 121,788 bp)	217,830 bp	pEND_Eco 18005 (CP126947.1) (IncFIB, 98.39%; 204,838 bp)	pEND_Eco 19025-1 (CP126932.1) (IncFIB, 98.39%; 160,439 bp)	No data	pEND_Eco 12049-1 (CP126953.1) (IncFIB, 99.12%; 155,857 bp)	pEND_Eco 14033-1 (CP126949.1) (IncFIB, 98.39%; 145,212 bp)	pPSUO78_1 (CP012113.1) (132,464 bp)
Plasmid 2	pEND_Eco 19057-2 (CP126936.1) (IncFIB, 98.39%; 104,962 bp)	pEND_Eco 123-2 (CP126957.1) (No hit; 68,426 bp)	113,260 bp		pEND_Eco 19025-2 (CP126933.1) (IncY, 99.38%; 90,722 bp)		pEND_Eco 12049-2 (CP126954.1) (Inc11-1(alpha), 100%; 114.229 bp)	pEND_Eco 14033-2 (CP126950.1) (IncI1-1 (alpha), 100%; 115,173 bp)	pPSUO78_2 (CP012114.1) (109,613 bp)
Plasmid 3 Plasmid 4							>	pEND_Eco 14033-3 (CP126951.1) (IncFIB (H89-PhagePlasmid), 98.7%; 108,542 bp) pEND_Eco 14033-4 (CP130557.1) (Incl2(delta), 97.15%; 61,041 bp)	
No. Genomic islands (GI) ^b	69	70	47	94	88	61	94	77	80

Table 5. Comparative genomics of the representative 070 Ar EC sha	Table 3.	Comparative	genomics o	of the rep	resentative	078 A	PEC strai
--	----------	-------------	------------	------------	-------------	-------	-----------

^a Plasmid classification according to [32,33]. ^b Prediction was performed using the Islandviewer 4 (including IslandPick, IslandPath-DIMOB, SIGI-HMM, and Islander).

9	of	23
---	----	----

					RST3-1				RST4-4				RST21-1					
RSTs	Strain	Accession no.	In silico	Host	E19057	Identity	E123	Identity	E18005	Identity	E19025	Identity	E12049	Identity	E14033	Identity		
			scrotype		coverage	Identity	Coverage	ruentity	Coverage	Identity								
RST3-1	E19057	CP126934.1	O78:H4	Chicken	100%	100.00%	95%	99.99%	89%	99.39%	89%	99.39%	85%	98.10%	88%	98.10%	100% T	100%
GI	E123	CP126955.1	O78:H4	Chicken	97%	99.99%	100%	100.00%	90%	99.40%	91%	99.40%	86%	98.10%	87%	98.10%		
	APEC O78	CP004009.1	O78:H9	Chicken	95%	99.99%	94%	99.98%	88%	99.39%	89%	99.39%	83%	98.10%	84%	98.10%	(N	
	NC19	CP076654.1	O78:H9	Duck	95%	99.98%	94%	99.98%	88%	99.39%	88%	99.39%	84%	98.10%	85%	98.10%	(N	
	FDAARGOS_433	CP023899.1	O8:H21	Human	91%	99.96%	89%	99.97%	86%	99.40%	86%	99.40%	81%	98.11%	82%	98.11%		
	RUT3575	CP048010.1	O8:H9	Human	90%	99.98%	90%	99.98%	86%	99.27%	87%	99.27%	82%	98.13%	83%	98.12%		
1	RHB13-C24	CP055964.1	O8:H9	Pig	93%	99.98%	92%	99.99%	86%	99.40%	87%	99.40%	82%	98.10%	83%	98.11%	(I	
1	WCHEC025943	CP027205.2	O8:H9	Wastewater	92%	99.97%	91%	99.98%	86%	99.27%	87%	99.27%	82%	98.12%	83%	98.12%		
RST3-1	268.2	CP059947.1	O8:H10	Turkey	92%	99.39%	91%	99.39%	91%	99.98%	91%	99.98%	84%	98.16%	85%	98.15%	90.50% -	- 99.03%
G2	RHBSTW-00046	CP056894.1	O8:H30	Wastewater	91%	99.30%	90%	99.30%	90%	99.92%	91%	99.97%	84%	98.23%	85%	98.23%		
	RHB23-C24	CP057512.1	ONT:H21	Sheep	90%	99.44%	89%	99.46%	90%	99.94%	90%	99.94%	83%	98.12%	84%	98.11%		
	E-T654	CP090293.1	O9:H4	Human	89%	99.42%	89%	99.43%	89%	99.91%	90%	99.94%	81%	98.16%	82%	98.16%		
	WS3294A	CP035819.1	O78:H17	Human	90%	99.23%	89%	99.34%	90%	99.79%	93%	99,73%	86%	98.16%	85%	98.15%		
	FWSEC0011	CP031892.1	O113:H21	Bovine	91%	99.42%	89%	99.40%	91%	99.89%	90%	99.75%	85%	98.16%	85%	98,16%		
RST4-4	E18005	CP126946.1	O78:H51	Chicken	93%	99.39%	92%	99.40%	100%	100.00%	98%	99.99%	87%	99.72%	87%	98.13%		
	E19025	CP126931.1	O78:H51	Chicken	92%	99.39%	92%	99.40%	97%	99.99%	100%	100.00%	87%	99.89%	88%	99.79%		
	PK8566	CP080172.1	ONT:H9	Chicken	91%	99.44%	90%	99.45%	89%	99.99%	90%	99.96%	83%	98.16%	84%	98.15%	81.00%	98.07%
	13C1079T	CP019267.1	O8/O179:H51	Chicken	91%	99.39%	89%	99.40%	90%	99.95%	91%	99.98%	82%	98.16%	83%	98.16%		
	EF7-18-58	CP068591.1	O8:H25	Chicken	90%	99 39%	89%	99.40%	89%	99.95%	90%	99 98%	82%	98.16%	84%	98.16%		
	NCTC11129	LR134222.1	O8:H51	NA	92%	99.40%	91%	99.40%	92%	99.98%	93%	99.98%	82%	98.16%	86%	98.16%		
	E16EC0792	CP088890.1	O123/O186:H40	Human	90%	99.78%	89%	99.43%	90%	99.69%	91%	99.98%	82%	98 16%	84%	98.16%		
	EC71	CP085622.1	O9/O160:H51	Human	91%	99.28%	90%	00 38%	91%	99.0976	91%	99.97%	84%	98 16%	85%	98.16%		
	RHBSTW-00905	CP056249.1	O9a/H10	Wastewater	929/.	99.40%	96%	99.41%	97%	99.95%	9704	99.97%	82%	98 16%	92%	98.16%		
	AMSCJX04	CP058308.1	O21:H51	Panda	0070	99.4076	019/	99.41%	019/	99.9376	07/0	00.079/	929/	90.1076	05/0	38.1076 09.169/		
RST21-1	E12049	CP126952.1	O78:H4	Chicken	92.76	99.4076	9170	99.41%	9170	00.729/	9270	00.909/	100%	30.1076	0570	96.1076		
	E14033	CP126948.1	O78:H4	Chicken	0970	90.1070	0070	98.1070	0/70	99.7270	0170	99.8976	0694	100.00 %	9576	99.00 %		
	APEC 02-211	CP006834.2	02:H4	Chicken	80%	98.10%	8/70	98.10%	8/70	98.13%	870	99.79%	90%	99.80%	100%	100.00%		
	PSU078	CP012112.1	078·H4	Chicken	88%	98.10%	88%	98.10%	80%	98.14%	8/%	98.07%	94%	99.98%	96%	99.98%		
	56.2	CP059922.1	ONT-H4	Turkey	8/%	98.09%	8/%	98.10%	85%	98.13%	8/%	98.13%	95%	99.84%	96%	99.98%		
	D10	CP010157 1	0143-114	Dog	86%	98.09%	85%	98.09%	84%	98.13%	85%	98.13%	92%	99.95%	94%	99.97%		
	KKa000	CP001673 1	08/0160-04	Dug Divor water	87%	98.11%	87%	98.11%	85%	98.13%	86%	98.14%	93%	99.98%	94%	99.96%		
	MCUC 122	CP041204.1	024.114	Humon	87%	98.10%	86%	98.10%	85%	98.13%	85%	98.14%	91%	99.80%	94%	99.96%		
	2000/C 2122	CP012025 1	024:114	Human	88%	98.10%	87%	98.10%	85%	98.14%	86%	98.14%	93%	99.93%	94%	99.98%		
	2009C-5155	CP015025.1	062.019	Dia	88%	98.10%	88%	98.10%	86%	98.11%	87%	98.07%	95%	99.85%	97%	99.87%		
	KIID14-C14	Cr05/818.1	005.1118	rig	86%	98 10%	86%	98 10%	84%	98 13%	84%	98 14%	90%	99 99%	92%	99 96%	1	

Figure 2. Comparison of genomic coverages (%) and identities (%) of O78 *E. coli* strains. A heatmap was constructed to represent the average nucleotide identity (ANI) and alignment coverage between O78 APEC and other RST strains. The values are visually depicted based on color intensity, with the alignment coverage ranging from red to green and the nucleotide identity represented by a color gradient from red to blue.

The contents of CRISPR spacers reflect the recent invasion–defense history of E. coli strains, and we can obtain information about evolutionary relatedness between E. coli strains and the identities of present invaders by comparing the contents and numbers of spacers [34]. The compositions of CRISPR spacers tend to be conserved among E. coli strains with identical RSTs, especially when their genome sequence coverages/identities are notably high (Tables S5 and S6). Typically, a new spacer is appended to the 5'-end to the most recent spacers, and some unnecessary spacers are lost randomly [8]. To elucidate the evolutionary trajectory of spacers, considering both the nucleotide sequences and length, we collected all the spacers of the closely related E. coli strains of each RST and inferred the evolutionary progression based on the presence and absence of specific spacers. As a result, the stepwise deletion and/or integration of spacers among related E. coli strains within RST3-1, RST4-4, and RST21-1 could be delineated by considering the spacer profiles of putative progenitors (Figure 3). For RST3-1, E19057 and E123 may have lost spacers 4-6 and 1-5 in CRT2, respectively. In the case of RST4-4, there may be a hypothetical progenitor harboring all spacers, and the RST4-4 strains may have lost varying spacers. However, in E19025, E18005, and NCTC11129, a new spacer was added. The evolutionary steps of RST21-1 spacers were more complicated, and two hypothetical progenitors were included. Hypothetical progenitor I evolved into hypothetical progenitor II by adding 15 spacers. E14033 lost one spacer of hypothetical progenitor I, but E12049, PSUO78, and APEC O2-211 lost multiple spacers of hypothetical progenitor II. Therefore, E14033 was closer to the recent progenitor of RST21-1 strains than others in terms of spacer evolution. At present, the available data are not sufficient to answer it. It is reasonable to ask why most strains tend to delete spacers. The spacers of representative RST3-1, RST4-4, and RST21-1 strains targeted plasmids and bacteriophages. In contrast to the RST3-1 spacers, mainly targeting plasmids (93.3%), the RST4-4 (100%) and RST21-1 (90.0%) spacers mainly targeted bacteriophage genes (Table S7). In a previous study, the targets of spacers were chromosomal regions (12%), plasmids (31%), and phages (57%), and the bias of the targets of RST3-1 strains toward plasmids and RST4-4 and RST21-1 toward phages observed in the present study was unexpected [34].



Figure 3. Evolutionary model based on the spacer sequences of CRISPRs in O78 APEC strains according to RST: (**A**) RST3-1, (**B**) RST4-4, and (**C**) RST21-1. The CRISPR spacers of genetically related strains of each RST were compared, and their progenitor–progeny relationships were hypothesized in terms of spacer contents. In the case of RST3-1 strains, E19057 and E123 may be different progenies of APEC O78 and may have lost different numbers of spacers in CRT2. In the case of RST4-4, E19025, E18005, and NCTC11129 may have lost different numbers of spacers in CRT1 and CRT2 of a hypothetical progenitor. A RST3-1_G2 strain, WS3294A, shares the same hypothetical progenitor of RST4-4, reflecting their evolutionary relatedness. Only RST4-4 strains E19025, E18005, and NCTC11129 acquired one spacer (CRT2_11). In the case of RST21-1, E14033 may be a progeny of hypothetical progenitor I, and APEC O2-211, PSU-O78, and E12049 are progenies of hypothetical progenitor II. All the RST21-1 strains may have experienced different deletions in CRT1 and CRT2.

2.6. Evolutionary Linkage between Early-Appearing RSTs

Genome types RST2-1_G1 (O26), RST2-1_G2, and RST2-1_G3 exhibited high genomic coverage/identity related to RST3-2 (91/99.87%, O111)/3-5 (96/99/95%, O26), RST3-7 (93/99.97%), and RST3-3 (96/99.94%)/3-8 (99/99.89%), respectively (Figure 4). The RST3-1 O78 APEC strains were classified under RST3-1_G1 and showed high genomic coverage/identities and CRISPR spacer contents related to non-O78 and O8 human strains of the same genome type (Table S5). They showed higher genomic coverage/identities in relation to RST4-12 (93/99.74%) and RST4-27 (94/99.71%) (Figure 4). No genetic linkage was observed between the RST2-1 and RST3-1 strains, and they may have evolved independently of each other. RST3-1_G2 displayed relatively high genomic coverage/identity (92/99.97%) and shared some CRISPR spacers with RST4-4 O78 APEC strains (Figure 4; Table S5).



Figure 4. Evolutionary relationships of RSTs appearing early. *E. coli* lineages ranging from RST2-1 to RST 4-27 were predicted using the measure of genomic coverage (%)/identity (%).

2.7. Evolution of O Serogroups in Different RSTs

On the basis of genome sequence coverage/identity and RSTs, we hypothesized the evolution of O serogroups. The O antigens of simpler structures and components may appear earlier than more complex O antigens and much simpler O serogroups may appear in earlier RSTs [35]. The major O serogroup of RST2-1 is O26, which is composed of linearly linked trisaccharides without side branches $[\rightarrow 3)$ -L-Rha- $(\alpha 1 \rightarrow 4)$ -L-FucNAc- $(\alpha 1 \rightarrow 3)$ -DGlcNAc- $(\beta 1 \rightarrow]$ [3]. The genome sequence coverages/identities between O26 strains are 96–99% and 99.96–99.99%. However, other strains with different O antigen structures, such as linearly linked tetrasaccharides (O71, O118, and O186) and pentasaccharides (O69 and O103), show high genome sequence coverages/identities (97–98%/99.97–99.98%) related to O26 strains (Table S4). The O antigen structures and O-AGC sequences apparently differ between them, and the different O-AGCs may have been horizontally transferred to the genomes of the prevalent O26 strains [3,36,37]. In contrast to clearly different O157:H7 EHEC strains (RST25-1), the evolutionary links between O26, O103, and O111 ETEC strains were obscure, but the high genomic coverage and identity observed between RST2-1 O26 ETEC strains and other RST2-1 O103 and RST3-2 O111 strains can explain the parallel evolution of various O serogroup ETEC strains (Figure 4) [38].

The prevalence of the O8 serogroup among RST3-1_G1 strains may reflect the early appearance of the O8 antigen with a simple structure and components, $[\rightarrow 3)$ -D-Man- $(\beta 1 \rightarrow 2)$ -D-Man- $(\alpha 1 \rightarrow 2)$ -D-Man- $(\alpha 1 \rightarrow]$, during *E. coli* evolution. Most RST3-1_G1 strains originated from healthy or clinical human specimens and may have evolved in humans. O78 serogroup RST3-1_G1 strains were only present in chickens and turkeys, and they showed more complex mPPTs than human non-O78 RST3-1_G1 strains (0 to 2 prophages) (Table S5). This may support the possibility that the O78 serogroup appeared in chickens and evolved via the acquisition of more prophages in the intestines of poultry. The O serogroups of RST3-1_G2 were mixed with various serogroups, including O8 and O78. The only RST3-1_G2 human-origin O78 strain, WS3294A, was ETEC, which contained a more complex and unique mPPT (4-6-7-11) than other RST3-1 O78 strains (Table S2). The close genetic linkage observed between non-O78 strains of RST3-1_G2 and RST4-4 O78 strains may suggest the acquisition of O78-AGC by some strains within RST3-1_G2 from RST3-1_G1 or RST4-4 O78 strains. No O78 RST4-4 strains, except for Korean APEC strains, have been identified to date (Table S9). The RST3-1 O78 APEC strain AH01 was classified as RST3-1_G3 and is the only O78 avian strain in the genome type. In the RST3-1_G4 genome type, three O78 strains (ATCC 43896, L3_E36, and 00-3279) were identified, and some of them were pathogenic and showed high genomic coverage/identity (98–99%/99.99%) and

similar mPPTs to each other (Table S2). However, their mPPTs were different from those of other RST3-1 avian O78 strains in terms of the presence of prophage 7. The presence of a relatively simple O78 antigen, consisting of four linearly linked carbohydrates of two different components (D-mannose and D-GlcNAc) in RST3-1 and RST4-4 may support the early appearance of the O78 antigen during *E. coli* evolution (Figure 2). The distribution of the O78 serogroup in other RSTs in branch I-1 can be explained by the maintenance of the O78 serogroup during the evolution of *rpoB*.

The O78 serogroup appeared abruptly among RST21-1 strains in branch I-2. The RST21-1 strains showing high genomic coverage/identity and common CRISPR spacer contents had variable O serogroups, and most of the O serogroups contained more evolved structures, except O78 (Table S8). The majority of RST21-1 strains (45.8%, 11/24) originated from avian hosts and belonged to the non-O78 serogroup, except for the PSUO78 strain. Some still possessed simpler mPPT types (2-5, 2-7, and 2-4-5). Therefore, RST21-1 O78 APEC strains may have evolved from non-O78 RST21-1 strains through the acquisition of the prevalent O78-AGC in avian hosts.

Thus, we found several new cases of an O serogroup shift, such as the case of O157:H7 *E. coli* originating from O55 *E. coli* in RST25-1, identified based on the RSTing scheme and simple genomic coverages/identities [39]. Although the RST3-1, RST4-30, and RST6-1 O78 ETEC strains have been reported in humans, the relatively low zoonotic risk of major O78 APEC strains can be supported by the low frequencies of O78 *E. coli* strains in clinical human cases at present (Table S2).

2.8. Profiling Virulence Genes of O78 APEC Strains

The presence of 24 representative virulence genes grouped into six categories was confirmed (Tables 1 and S10). All O78 strains were found to possess the *fimH*, *crlA*, *iss*, sitA, iroN, hlyF, ompT, and malX genes. Additionally, the following genes were detected at various frequencies: fyuA (90.6%), iutA (87.5%), irp2 (87.5%), mat (75.0%), hra (75.0%), tsh (75.0%), traT (68.8%), and ireA (65.6%). As previously reported, the redundant possession of siderophores and the absence of toxin genes such as stx1, stx2, lt, and st were verified [5]. The virulence gene content varied among strains, ranging from 9 to 20, and there was an evident accumulation of multiple virulence genes (Table 1). Additionally, the accumulation orders and patterns of virulence genes are hinted at in our results (Table S10). In the case of adhesion genes, *fimH* and *crlA* may have been acquired first, after which *hra* and mat were acquired by the majority of O78 strains (75.0%, 24/32). Subsequently, papC or *iha* were optionally acquired. Siderophore genes may have acquired *sitA/iroN* followed by fyuA, iutA/irp2, ireA, and chuA. Regarding toxin genes, hlyF may have been acquired first, followed sequentially by tsh, pic/vat, and astA. The distribution of virulence genes on the chromosome and/or plasmids was determined using the genome sequences of the E19057, E123, E18005, E19025, E12049, and E14033 strains, and the results are summarized in Figure S3. All of the adhesion genes, some siderophore genes (fyuA, irp2, ireA, and chuA), some toxin genes (*pic*, *vat*, and *astA*), one invasion gene (*tia*), and *malX* were present only on the chromosome. *iss, sitA*, and *ompT* were redundantly present on the chromosome and plasmid. traT, iroN, iutA, hlyF, and tsh were present only on the plasmids, and the origin sequences of multiple virulence gene-bearing plasmids were classified under the same type, IncFIB.

2.9. Antibiograms of O78 APEC Strains

The resistance of the strains to 24 antibiotics of eight classes and their resistance genes are summarized in Table S11. Among multidrug-resistant (MDR) strains, the frequency of RST types reveals that RST4-4 is the most predominant at 61.1%. This is followed by RST3-1 at 16.7% and RST21-1 at 11.1%. Within each RST group, the prevalence of MDR strains was as follows: RST3-1 at 27.3%, RST4-4 at 78.6%, and RST21-1 at 50%. The prevalence of multidrug resistance (MDR) in the RST4-4 strains was observed to be significantly elevated in comparison to the RST3-1 strains (p < 0.05), as determined through the application of the

Mann–Whitney U test (Table 4). In particular, MDR RST4-4 and MDR RST21-1 strains also possess many virulence genes that need to receive special attention.

The antibiotics tested in this study associated with the highest resistance were amoxicillin (100%), enrofloxacin (65.6%), oxytetracycline (56.3%), cefazolin (50.0%), streptomycin (28.1%), sulfamethoxazole/trimethoprim (21.9%), and florfenicol (6.3%). There was no strain showing resistance against colistin (Table S11). The recorded resistance rates are similar to other reports in Korea [4,40]. During 1980-2005, resistance to streptomycin and tetracycline was highest among Korean APEC strains, reaching 84.2%, followed by resistance to enrofloxacin (71.3%) and ampicillin (67.3%) [41]. The maintenance of antibiotic resistance is costly when the associated resistance mechanism is specific, and the decreased resistance to tetracycline and streptomycin in this study may be somehow related to the decreased usage of these antibiotics in chicks on poultry farms. However, high resistance frequencies of ampicillin and enrofloxacin still revealed no effect of the efforts to reduce antibiotic abuse in the field. Resistance to florfenicol is low at present, but it has become the preferred antibiotic for the treatment of APEC in the field. Therefore, changes in the resistance and prevalence of the *floR* gene need to be monitored. The resistance to cephalosporins and the acquisition of related resistance genes observed during 2014-2019 may be related to the recent increase in ceftiofur usage for the *in ovo* inoculation of embryonated chicken eggs for bacterial clearance in Korea, as in other countries [42,43]. However, it is noteworthy that the reduction in the use of certain antibiotics, such as tetracyclines and penicillins, did not necessarily lead to a decrease in antibiotic-resistant bacterial populations. In some cases, there was even a marginal increase observed. These findings suggest a complex interplay of factors influencing the persistence and evolution of antibiotic resistance, extending beyond the simple dynamics of antibiotic presence or absence [24,40–43].

Resistance to penicillins was associated with *bla_{TEM}* (31.3%), *bla_{SHV}* (3.1%), and PBP3 mutations (D350N and S357N), and extended resistance to cephalosporins was associated with *bla_{CTX-1}* (28.1%), *bla_{CTX-9}* (9.4%), and *ampC* (6.3%). Quinolone resistance is related to single or multiple mutations in gyrA, gyrB, and/or parC, and the S83L mutation of gyrA provides intermediate resistance to enrofloxacin and is shared by 90.6% of strains. For complete resistance to quinolones, additional mutations, such as the D87N of gyrA or the S492N of gyrB, were needed. Resistance to aminoglycosides was associated with *strA/strB* (21.9%), *aadA1/A2* (18.8%), and *aac*(3)II (18.8%). Resistance to tetracyclines was associated with *tetB* (50.0%) and *tetA* (18.8%). Resistance to sulfonamides was associated with sul2 (37.5%) and sul1 (15.6%), and resistance to trimethoprim was associated with *dfrA1* (12.9%). Resistance to sulfamethoxazole/trimethoprim discs (SXT) was determined by *dfrA1* in combination with *sul1* or *sul2*. Resistance to phenicols was associated with floR (6.3%). Only class 1 and 2 integrons were present in seven and one strains among the tested O78 APEC strains (Tables 5 and S11). The locations of resistance genes are summarized in Figure S3. The presence of class 1 integrons in E12049 and E14033 may be related to the presence of *aadA1/sul1* genes in similar plasmids in terms of size and origin sequence type, IncI1-1 α (Tables 3 and S11) [41,44]. Thus, the genotypic and phenotypic antibiograms matched each other well in the present study. Based on a comparison of the contents and numbers of resistance genes of O78 APEC strains, the different genotypes are depicted and theoretically interconnected in Figure 5. Most O78 APEC strains (21/32) shared a minimal gyrA gene mutation (S83L) and possessed 19 different combinations of tetA, tetB, ampC, bla_{TEM}, bla_{SHV}, bla_{CTX-M} groups 1 and 9, sul1, sul2, dfrA1, aadA, aac(3)II, *strA/B*, and *floR*. Therefore, some genotypes can be interconnected via the acquisition or deletion of resistance genes. Although we addressed specific genes related to antibiotic resistance, other unknown genes and cryptic mechanisms may be involved. In the case of MDR strains, it is very common that the resistance mechanisms are nonspecific and, therefore, are associated with mechanisms related to transport, cellular homeostasis or cellular aggregation, and a decrease in the area/volume ratio. Additionally, we need to consider the combinatorial capacity of the gene regulatory network to optimize resistance phenotypes [24,45–49].

	RST3-1	RST4-4	RST5-1	RST12-6	RST21-1				
Frequency out of	3/18	11/18	1/18	1/18	2/18				
MDR strains (%) ^a	(16.7%)	(61.1%) ^b	(5.6%)	(5.6%)	(11.1%)				
MDR Frequency out of	3/11	11/14	1/1	1/1	2/4				
each RST (%)	(27.3%)	(78.6%)	(100%)	(100%)	(50%)				

Table 4. Frequency of RSTs among MDR O78 APEC strains (*n* = 18).

^a Significant correlation between RST and MDR (p < 0.05). ^b Significant difference between RST3-1 and RST4-4 (p < 0.05).

Table 5. Frequency of antibiotic resistance-related genetic markers of O78 APEC strains (*n* = 32 including the E123 strain).

Antibiotics Resistance- Related Genes	gyrA	tetB	sul2	bla _{TEM}	bla _{CTX-M} -G1	strA/strB	tetA	aadA1 /A2	aac(3)II	sul1	dfrA1	bla _{CTX-M} -G9	ampC	floR	bla _{SHV}
Frequency (%)	29	16	12	10	9	7	6	6	6	5	4	3	2	2	1
	(90.6)	(50.0)	(37.5)	(31.3)	(28.1)	(21.9)	(18.8)	(18.8)	(18.8)	(15.6)	(12.5)	(9.4)	(6.3)	(6.3)	(3.1)



Figure 5. Theoretical interconnections between O78 APEC strains (n = 32) with different contents of antibiotic resistance genes (genotypes). This schematic illustrates relationships between different genotypes.

2.10. Profiles of Microcins and Colicins of O78 APEC Strains

To date, various microcins and colicins have been reported among *E. coli* strains [50–52]. First, we conducted an in silico analysis to screen for their presence in six Korean O78 APEC strains (Table 6). Microcin V, microcin J25, microcin H47, colicin Ia, colicin Ib, and colicin M were present on the chromosomes or plasmids. Except for E123 and E14033, all others possessed double or triple bacteriocins. We screened the presence of the above bacteriocins using PCR for the remaining O78 strains, and the results are summarized in Table S12. The frequency of microcin V (78.1%) was the highest, followed by colicin M (65.6%), microcin J25 (53.1%), colicin Ib (53.1%), microcin H47 (15.6%), and colicin Ia (3.1%).

<u>Etucio</u>	рст	Chromosomo		No. of			
Strain	K31	Chromosome	1	2	3	4	Microcin/Colicin ^a
E123	3-1	None	Microcin V (Colicin V)	None			1
E19057	3-1	None	Colicin Ia	Microcin V (Colicin V)			2
E18005	4-4	microcin H47	Microcin J25 Colicin M Microcin V (Colicin V) Colicin B				4
E19025	4-4	None	(pseudo) Microcin J25 Colicin M Microcin V (Colicin V) Colicin B (pseudo)	None			3
E12049	21-1	None	Colicin M	Colicin Ib			2
E14033	21-1	None	None	Collcin Ib	None	None	1

Table 6. Contents of microcins and colicins in the representative RSTs.

^a Counted numbers exclude colicin B (classified as pseudogenes).

We tested the bacteriocin activity of each O78 APEC strain against DH5a and six representative O78 APEC strains using a modified spot assay. Except for E16023, all the other O78 strains showed inhibitory activity with different intensities (from + to ++++). RST3-1 strains, E123 and E19057, were highly susceptible to bacteriocin produced by 11 (34.4%) and 5 (15.6%) of the O78 strains in comparison with RST21-1 E14033 (6.3%) and RST4-4 E19025 (3.1%). Among the twelve O78 strains killing at least one O78 strain, RST4-4 strains were more frequent (18.8%, 6/32; 42.9%, 6/14 RST4-4) than the others, including RST3-1 (9.4%, 3/32; 27.3%, 3/11 RST3-1), RST21-1 (6.3%, 2/32; 50.0%, 2/4 RST21-1), and RST12-6 (3.1%, 1/32) strains. A significant difference was observed in the number of the six bacteriocins assessed between RST3-1 and RST4-4 (p < 0.05), as determined using the Kruskal–Wallis test. (Table S12). In contrast to other strains killing one or two O78 strains, the RST4-4 strains E15016 and E15025 killed three strains (RST3-1 E123, E19057, and RST21-1 E14033) (Table S13). The development of relatively broad colicidal activity among RST4-4 strains may be a cause of the prevalence of RST4-4 compared with other RSTs.

3. Materials and Methods

3.1. Bacterial Strains and Identification

Thirty-one O78 APEC strains from 2012 to 2020 and a previously reported O78 APEC strain (E123) were isolated from poultry cases consigned to our laboratory (Table 1). Lactosenonfermenting colonies on the MacConkey agar (BD Difco, Becton Dr, NJ, USA) were identified using MALDI (matrix-assisted laser desorption ionization)–TOF (time-of-flight) mass spectrometry (Bruker, Billerica, MA, USA) and the bacterial species were further confirmed using 16S rRNA sequencing [53]. Bacteria were cultured in Luria–Bertani (LB) broth (Duchefa, Netherlands) at 37 °C and preserved in 20% glycerol LB broth stocks at -60 °C until use.

3.2. Serotyping and Phylogrouping of E. coli

The O serogroup of APEC was determined using an agglutination test, employing commercial O78-specific antisera in accordance with the manufacturer's protocol (Denka Seiken, Tokyo, Japan). Additionally, PCR was performed utilizing serogroup-specific primers, as described in a previous study [54]. In silico serotyping of O and H antigens was carried out utilizing SerotypeFinder (v.2.0.1, https://cge.food.dtu.dk/services/SerotypeFinder/) accessed on 1 May 2023 in conjunction with EcoSP (http://ecosp.dmicrobe.cn/) software [55,56] To identify the phylogenetic group, a multiplex PCR assay developed by Clermont was applied following the described protocol [57].

3.3. rpoB Sequence Typing (RSTing) and Molecular Prophage Typing (mPPTing)

We refined a previously reported *rpoB* sequence type (RST) by implementing some modifications. The PCR assay and sequencing were conducted as in a previous study [6].

The complete *rpoB* nucleotide sequences of *E. coli* (3029 strains) and other bacteria were collected from GenBank (National Center for Biotechnology Information, NCBI) (up to 5 April 2023) (Table 2). Upon aligning the whole *rpoB* sequences, including those of O78 APEC with MEGA X (v.11.0.13) [58], haplotypes were determined based on SNPs using DnaSP (v.6.12.03) [59]. Initial RST numbers were assigned by counting SNPs in the *rpoB* sequences, and subsequent sub-numbers were serialized according to the distinct haplotypes. The *rpoB* network was constructed using the median-joining network method with the aid of popART (v.1.7) [29,30]. Bayesian phylogenetic trees were constructed using MrBayes (v.2.2.4) [60], implemented in Geneious Prime (v.2023.2.1) (https://www.geneious.com, Auckland, New Zealand). Labels for the figures were edited using Adobe Illustrator 2023 (Adobe Inc., San Jose, CA, USA).

In silico molecular prophage typing was performed as previously reported, with a modification to add information about the copy numbers of each mPPT [7].

3.4. Molecular Pathotyping

The presence of 34 virulence genes in O78 APEC strains was investigated using 34 primers based on a combination of a 5-set multiplex PCR and four single PCR assays, as reported previously (Table S14) [61–66].

3.5. Antibiotic Susceptibility Testing and Molecular Profiling of Antibiotic Resistance Genes

Sensitivity to various antibiotics listed in supplementary Table S11 was assessed utilizing the disk diffusion method, according to Kirby and Bauer [67], in conjunction with the guidelines provided by the Clinical and Laboratory Standards Institute (CLSI) [68]. For the determination of colistin susceptibility, the broth microdilution method (BMD) was employed as described in previous studies [69,70] and in accordance with the EUCAST guidelines (www.eucast.org, Basel, Switzerland).

Subsequently, 37 antimicrobial resistance genes were amplified from all O78 APEC isolates using PCR and utilizing the primers listed in Table S14, as previously reported [71–88]. Mutations in the quinolone resistance-determining regions (QRDRs) of *gyrA*, *gyrB*, *parC*, and *parE* were identified via sequencing after PCR, as reported previously [76,77].

3.6. Comparative Genomics

Whole-genome sequencing (WGS) was performed using de novo assembly and PacBio sequencing technology (PacBio) after constructing a SMRTbell long-read library. To ensure accurate genome sequences, error correction was applied utilizing a TruSeq DNA library (insert size: 350 bp) in conjunction with Illumina high-quality data (100 bp paired-end). DFAST was utilized for gene annotation [89]. MLST 2.0. and PubMLST were applied to

match *E. coli* MLST alleles [90,91]. CRISPRs and *cas* genes in the genome were detected using CRISPRCasFinder (v.1.1.2), and the matching of spacer sequences was accomplished with the CRISPRTarget tool along with GenBank data using BLASTN [92–94]. The visualization of genomic analysis was performed using the BLAST Ring Image Generator (BRIG) (Figure S2) [95].

3.7. Identification of Bacteriocin Genes in O78 APEC

The primers utilized for molecular detection of bacteriocins in this study are described in Table S14 [50]. The *E. coli* cells were cultured overnight in LB broth, and the bacterial DNA was obtained by lysing the cells at 95 °C for 10 min. Aliquots of 20 μ L of a reaction mixture containing 2 μ L DNA, 2 μ L primer set, and 16 μ L DW were added to the Maxime PCR PreMix strips (i-StarTaq) kit (iNtRon, Seongnam, Republic of Korea). The amplification conditions were as follows: 1 cycle at 95 °C for 5 min, 30 cycles of 95 °C for 30 s, 55 °C for 30 s, and 72 °C for 1 min, and a final cycle of 5 min at 72 °C, followed by a hold at 12 °C.

3.8. Antimicrobial Activity Test of Bacteriocin

Brief screening for sensitivity to bacteriocin from O78 APEC was conducted using a modified spot assay method. Bacterial broth cultures were diluted in PBS to achieve a turbidity of 0.1–0.2 MacFarland, ensuring similar numbers of seven indicator *E. coli* strains (DH5 α , E19057, E123, E18005, E19025, E12049, and E14033). Three milliliters of the diluent was evenly spread onto square LB agar plates (SPL, Pocheon, Republic of Korea) and allowed to dry until no liquid remained. Overnight bacterial cultures of 32 O78 APEC strains and DH5 α in LB broth were concentrated 100-fold, and a 10 μ L volume of the fluid was spotted onto the indicator lawn plates. After overnight incubation at 37 °C, the sensitivity to O78 bacteriocin was determined based on the thickness of clear zones around the spotted sites: -, no inhibition zone; +, 1–2 mm; ++, 3–4 mm; +++, 5–6 mm; ++++, \geq 7 mm.

3.9. Statistical Analysis

Fisher's exact tests were used to determine the significant association between the prevalence of RST and MDR strains. The Mann–Whitney U test was utilized to compare the MDR frequencies among the RST groups. Additionally, the presence of distinct virulence genes and bacteriocins across different RST groups was analyzed using the Kruskal–Wallis test, followed by the Dunn–Bonferroni correction. All statistical analyses were conducted using SPSS (v.26.0) (IBM Corp., Armonk, NY, USA), and the results were interpreted with 95% confidence intervals. To ensure reliability, all experiments in this study were repeated at least twice.

4. Conclusions

Taken together, the RSTing and network analysis in the present study provided new insights into the evolutionary relationships between *E. coli* and other related bacteria. The major RSTs, RST3-1 and RST4-4, of O78 serogroup *E. coli* may have settled down and adapted to poultry to become APEC by accumulating virulence and antibiotic resistance genes. The zoonotic potential of O78 APEC strains is not currently high, but the prevalence of MDR RST4-4 and the appearance of new RST21-1 O78 APEC strains bearing multiple virulence genes encourage the continuous monitoring of their prevalence and virulence in poultry and their zoonotic risk.

Supplementary Materials: The following supporting information can be downloaded at: https:// www.mdpi.com/article/10.3390/antibiotics12121714/s1, Figure S1: Bayesian phylogenetic analysis of *E. coli rpoB* sequence types; Figure S2: Comparison of genome sequences of O78 APEC strains; Figure S3: The location of virulence genes and antibiotic resistance genes in the O78 APEC genome; Table S1: *rpoB* sequence types (RSTs) of *E. coli* strains; Table S2: O78 *E. coli* strains among the analyzed *E. coli* strains; Table S3: RSTs of *Shigella* species; Table S4: In silico genetic characterization of RST2-1 *E. coli* strains; Table S5: In silico genetic characterization of RST3-1 *E. coli* strains; Table S6: Comparison

19 of 23

of CRISPR spacers of the representative O78 *E. coli* strains; Table S7: Targets of CRISPR spacers of the representative O78 APEC strains; Table S8: In silico genetic characterization of RST21-1 *E. coli* strains; Table S9: In silico genetic characterization of RST4-4 *E. coli* strains; Table S10: Profiles and frequency of virulence genes of O78 APEC strains; Table S11: Genotypic and phenotypic correlations of antibiotic resistance among O78 APEC strains; Table S12: Profiles of bacteriocins and colicins among O78 APEC strains; Table S13: Colicin activity of O78 APEC strains; Table S14: Primer sets used in this study.

Author Contributions: Conceptualization, E.-J.H. and H.-J.K.; methodology, E.-J.H., S.-M.H. and H.-J.K.; software, E.-J.H. and S.-M.A.; validation, E.-J.H. and H.-J.K.; formal analysis, E.-J.H. and H.-J.K.; investigation, E.-J.H., S.-M.H., S.-J.K. and H.-W.K.; resources, E.-J.H., S.-M.H., S.-J.K. and H.-W.K.; data curation, E.-J.H. and S.-M.H.; writing—original draft preparation, E.-J.H.; writing—review and editing, E.-J.H. and H.-J.K.; visualization, E.-J.H. and S.-M.A.; supervision, K.-S.C. and H.-J.K.; project administration, E.-J.H., K.-S.C. and H.-J.K.; funding acquisition, K.-S.C. and H.-J.K. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No. 2023R1A2C1007801) and also by the Research Institute of Veterinary Science of Seoul National University's BK21 program.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The WGS data for the O78 APEC chromosome and plasmids are available in the NCBI database under BioProject PRJNA978016. Accession numbers of the individual genomes are described in Table 3.

Conflicts of Interest: Author Hyuk-Joon Kwon was employed by the GeNiner Inc. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- 1. Nataro, J.P.; Kaper, J.B. Diarrheagenic Escherichia coli. Clin. Microbiol. Rev. 1998, 11, 142–201. [CrossRef] [PubMed]
- Russo, T.A.; Johnson, J.R. Proposal for a new inclusive designation for extraintestinal pathogenic isolates of *Escherichia coli*: ExPEC. J. Infect. Dis. 2000, 181, 1753–1754. [CrossRef] [PubMed]
- 3. Liu, B.; Furevi, A.; Perepelov, A.V.; Guo, X.; Cao, H.; Wang, Q.; Reeves, P.R.; Knirel, Y.A.; Wang, L.; Widmalm, G. Structure and genetics of *Escherichia coli* O antigens. *FEMS Microbiol. Rev.* **2020**, *44*, 655–683. [CrossRef]
- Kim, Y.B.; Yoon, M.Y.; Ha, J.S.; Seo, K.W.; Noh, E.B.; Son, S.H.; Lee, Y.J. Molecular characterization of avian pathogenic *Escherichia* coli from broiler chickens with colibacillosis. *Poult. Sci.* 2020, 99, 1088–1095. [CrossRef]
- 5. Jeong, Y.W.; Kim, T.E.; Kim, J.H.; Kwon, H.J. Pathotyping avian pathogenic *Escherichia coli* strains in Korea. *J. Vet. Sci.* 2012, 13, 145–152. [CrossRef] [PubMed]
- Kwon, H.-J.; Seong, W.-J.; Kim, T.-E.; Won, Y.-J.; Kim, J.-H. *rpoB* gene sequencing for phylogenetic analysis of avian pathogenic *Escherichia coli. Korean J. Vet. Res.* 2015, 55, 31–39. [CrossRef]
- Kwon, H.J.; Seong, W.J.; Kim, J.H. Molecular prophage typing of avian pathogenic *Escherichia coli*. Vet. Microbiol. 2013, 162, 785–792. [CrossRef]
- 8. Garrett, S.C. Pruning and Tending Immune Memories: Spacer Dynamics in the CRISPR Array. *Front. Microbiol.* **2021**, *12*, 664299. [CrossRef]
- 9. Shin, H.; Kim, Y.; Unno, T.; Hur, H.G. Prevalence and Characterization of CRISPR Locus 2.1 Spacers in *Escherichia coli* Isolates Obtained from Feces of Animals and Humans. *Microbiol. Spectr.* 2023, *11*, e0493422. [CrossRef]
- 10. Zhou, Z.; Charlesworth, J.; Achtman, M. HierCC: A multi-level clustering scheme for population assignments based on core genome MLST. *Bioinformatics* **2021**, *37*, 3645–3646. [CrossRef]
- 11. Ding, W.; Baumdicker, F.; Neher, R.A. panX: Pan-genome analysis and exploration. Nucleic Acids Res. 2018, 46, e5. [CrossRef]
- 12. Ko, D.S.; Kim, D.; Kim, E.K.; Kim, J.H.; Kwon, H.J. Evolution of a major bovine mastitic genotype (*rpoB* sequence type 10-2) of *Staphylococcus aureus* in cows. *J. Microbiol.* **2019**, *57*, 587–596. [CrossRef]
- Ko, D.S.; Kim, N.H.; Kim, E.K.; Ha, E.J.; Ro, Y.H.; Kim, D.; Choi, K.S.; Kwon, H.J. Comparative genomics of bovine mastitis-origin Staphylococcus aureus strains classified into prevalent human genotypes. Res. Vet. Sci. 2021, 139, 67–77. [CrossRef]
- 14. Guinane, C.M.; Ben Zakour, N.L.; Tormo-Mas, M.A.; Weinert, L.A.; Lowder, B.V.; Cartwright, R.A.; Smyth, D.S.; Smyth, C.J.; Lindsay, J.A.; Gould, K.A.; et al. Evolutionary genomics of *Staphylococcus aureus* reveals insights into the origin and molecular basis of ruminant host adaptation. *Genome Biol. Evol.* **2010**, *2*, 454–466. [CrossRef]

- Herron-Olson, L.; Fitzgerald, J.R.; Musser, J.M.; Kapur, V. Molecular correlates of host specialization in *Staphylococcus aureus*. *PLoS* ONE 2007, 2, e1120. [CrossRef] [PubMed]
- Lowder, B.V.; Guinane, C.M.; Ben Zakour, N.L.; Weinert, L.A.; Conway-Morris, A.; Cartwright, R.A.; Simpson, A.J.; Rambaut, A.; Nubel, U.; Fitzgerald, J.R. Recent human-to-poultry host jump, adaptation, and pandemic spread of *Staphylococcus aureus*. *Proc. Natl. Acad. Sci. USA* 2009, 106, 19545–19550. [CrossRef] [PubMed]
- Viana, D.; Comos, M.; McAdam, P.R.; Ward, M.J.; Selva, L.; Guinane, C.M.; Gonzalez-Munoz, B.M.; Tristan, A.; Foster, S.J.; Fitzgerald, J.R.; et al. A single natural nucleotide mutation alters bacterial pathogen host tropism. *Nat. Genet.* 2015, 47, 361–366. [CrossRef] [PubMed]
- 18. Dahal, R.H.; Choi, Y.J.; Kim, S.; Kim, J. Differentiation of *Escherichia fergusonii* and *Escherichia coli* Isolated from Patients with Inflammatory Bowel Disease/Ischemic Colitis and Their Antimicrobial Susceptibility Patterns. *Antibiotics* 2023, 12, 154. [CrossRef]
- Gomes, T.A.T.; Ooka, T.; Hernandes, R.T.; Yamamoto, D.; Hayashi, T. Escherichia albertii Pathogenesis. EcoSal Plus 2020, 9. [CrossRef] [PubMed]
- Cobo-Simon, M.; Hart, R.; Ochman, H. Escherichia coli: What Is and Which Are? Mol. Biol. Evol. 2023, 40, msac273. [CrossRef] [PubMed]
- Gordienko, E.N.; Kazanov, M.D.; Gelfand, M.S. Evolution of pan-genomes of *Escherichia coli*, *Shigella* spp., and *Salmonella enterica*. J. Bacteriol. 2013, 195, 2786–2792. [CrossRef] [PubMed]
- Kim, N.H.; Ha, E.J.; Ko, D.S.; Lee, C.Y.; Kim, J.H.; Kwon, H.J. Molecular evolution of *Salmonella enterica* subsp. *enterica* serovar Gallinarum biovar Gallinarum in the field. *Vet. Microbiol.* 2019, 235, 63–70. [CrossRef] [PubMed]
- 23. Krakauer, D.C.; Plotkin, J.B. Redundancy, antiredundancy, and the robustness of genomes. *Proc. Natl. Acad. Sci. USA* 2002, *99*, 1405–1409. [CrossRef] [PubMed]
- 24. Smith, H.W. Persistence of tetracycline resistance in pig E. coli. *Nature* 1975, 258, 628–630. [CrossRef] [PubMed]
- Crossman, L.C.; Chaudhuri, R.R.; Beatson, S.A.; Wells, T.J.; Desvaux, M.; Cunningham, A.F.; Petty, N.K.; Mahon, V.; Brinkley, C.; Hobman, J.L.; et al. A commensal gone bad: Complete genome sequence of the prototypical enterotoxigenic *Escherichia coli* strain H10407. *J. Bacteriol.* 2010, 192, 5822–5831. [CrossRef]
- 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. The genome sequence of avian pathogenic *Escherichia coli* strain O1:K1:H7 shares strong similarities with human extraintestinal pathogenic *E. coli* genomes. *J. Bacteriol.* 2007, 189, 3228–3236. [CrossRef]
- 27. Severinov, K.; Soushko, M.; Goldfarb, A.; Nikiforov, V. RifR mutations in the beginning of the *Escherichia coli rpoB* gene. *Mol. Gen. Genet.* **1994**, 244, 120–126. [CrossRef]
- 28. Doolittle, R.F.; Feng, D.F.; Tsang, S.; Cho, G.; Little, E. Determining divergence times of the major kingdoms of living organisms with a protein clock. *Science* **1996**, 271, 470–477. [CrossRef]
- Bandelt, H.J.; Forster, P.; Rohl, A. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* 1999, 16, 37–48. [CrossRef]
- Leigh, J.W.; Bryant, D.; Nakagawa, S. popart: Full-feature software for haplotype network construction. *Methods Ecol. Evol.* 2015, 6, 1110–1116. [CrossRef]
- Rasko, D.A.; Rosovitz, M.J.; Myers, G.S.; Mongodin, E.F.; Fricke, W.F.; Gajer, P.; Crabtree, J.; Sebaihia, M.; Thomson, N.R.; Chaudhuri, R.; et al. The pangenome structure of *Escherichia coli*: Comparative genomic analysis of *E. coli* commensal and pathogenic isolates. *J. Bacteriol.* 2008, 190, 6881–6893. [CrossRef] [PubMed]
- Thomas, C.M.; Thomson, N.R.; Cerdeno-Tarraga, A.M.; Brown, C.J.; Top, E.M.; Frost, L.S. Annotation of plasmid genes. *Plasmid* 2017, 91, 61–67. [CrossRef] [PubMed]
- Redondo-Salvo, S.; Bartomeus-Penalver, R.; Vielva, L.; Tagg, K.A.; Webb, H.E.; Fernandez-Lopez, R.; de la Cruz, F. COPLA, a taxonomic classifier of plasmids. *BMC Bioinform.* 2021, 22, 390. [CrossRef]
- 34. Touchon, M.; Charpentier, S.; Clermont, O.; Rocha, E.P.; Denamur, E.; Branger, C. CRISPR distribution within the *Escherichia coli* species is not suggestive of immunity-associated diversifying selection. *J. Bacteriol.* **2011**, 193, 2460–2467. [CrossRef]
- 35. Gupta, R.S. The natural evolutionary relationships among prokaryotes. *Crit. Rev. Microbiol.* **2000**, *26*, 111–131. [CrossRef] [PubMed]
- 36. Lerouge, I.; Vanderleyden, J. O-antigen structural variation: Mechanisms and possible roles in animal/plant–microbe interactions. *FEMS Microbiol. Rev.* 2002, *26*, 17–47. [CrossRef]
- DebRoy, C.; Fratamico, P.M.; Yan, X.; Baranzoni, G.; Liu, Y.; Needleman, D.S.; Tebbs, R.; O'Connell, C.D.; Allred, A.; Swimley, M.; et al. Comparison of O-Antigen Gene Clusters of All O-Serogroups of *Escherichia coli* and Proposal for Adopting a New Nomenclature for O-Typing. *PLoS ONE* 2016, *11*, e0147434. [CrossRef]
- Ogura, Y.; Ooka, T.; Iguchi, A.; Toh, H.; Asadulghani, M.; Oshima, K.; Kodama, T.; Abe, H.; Nakayama, K.; Kurokawa, K.; et al. Comparative genomics reveal the mechanism of the parallel evolution of O157 and non-O157 enterohemorrhagic *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 2009, 106, 17939–17944. [CrossRef]
- 39. Wang, L.; Huskic, S.; Cisterne, A.; Rothemund, D.; Reeves, P.R. The O-antigen gene cluster of *Escherichia coli* O55:H7 and identification of a new UDP-GlcNAc C4 epimerase gene. *J. Bacteriol.* **2002**, *184*, 2620–2625. [CrossRef]
- Jeong, J.; Lee, J.Y.; Kang, M.S.; Lee, H.J.; Kang, S.I.; Lee, O.M.; Kwon, Y.K.; Kim, J.H. Comparative Characteristics and Zoonotic Potential of Avian Pathogenic *Escherichia coli* (APEC) Isolates from Chicken and Duck in South Korea. *Microorganisms* 2021, 9, 946. [CrossRef]

- Kim, T.-E.; Jeong, Y.-W.; Cho, S.-H.; Kim, S.-J.; Kwon, H.-J. Chronological study of antibiotic resistances and their relevant genes in Korean avian pathogenic *Escherichia coli* isolates. *J. Clin. Microbiol.* 2007, 45, 3309–3315. [CrossRef] [PubMed]
- Agunos, A.; Leger, D.F.; Carson, C.A.; Gow, S.P.; Bosman, A.; Irwin, R.J.; Reid-Smith, R.J. Antimicrobial use surveillance in broiler chicken flocks in Canada, 2013-2015. *PLoS ONE* 2017, 12, e0179384. [CrossRef] [PubMed]
- Verrette, L.; Fairbrother, J.M.; Boulianne, M. Effect of Cessation of Ceftiofur and Substitution with Lincomycin-Spectinomycin on Extended-Spectrum-beta-Lactamase/AmpC Genes and Multidrug Resistance in *Escherichia coli* from a Canadian Broiler Production Pyramid. *Appl. Environ. Microbiol.* 2019, *85*, e00037-19. [CrossRef]
- 44. de Los Santos, E.; Lavina, M.; Poey, M.E. Strict relationship between class 1 integrons and resistance to sulfamethoxazole in *Escherichia coli. Microb. Pathog.* **2021**, *161*, 105206. [CrossRef]
- 45. Shis, D.L.; Bennett, M.R.; Igoshin, O.A. Dynamics of Bacterial Gene Regulatory Networks. *Annu. Rev. Biophys.* 2018, 47, 447–467. [CrossRef]
- Molstad, S.; Erntell, M.; Hanberger, H.; Melander, E.; Norman, C.; Skoog, G.; Lundborg, C.S.; Soderstrom, A.; Torell, E.; Cars, O. Sustained reduction of antibiotic use and low bacterial resistance: 10-year follow-up of the Swedish Strama programme. *Lancet Infect. Dis.* 2008, *8*, 125–132. [CrossRef] [PubMed]
- Nikaido, H. Molecular Basis of Bacterial Outer Membrane Permeability Revisited. *Microbiol. Mol. Biol. Rev.* 2003, 67, 593–656. [CrossRef]
- Blair, J.M.; Webber, M.A.; Baylay, A.J.; Ogbolu, D.O.; Piddock, L.J. Molecular mechanisms of antibiotic resistance. *Nat. Rev. Microbiol.* 2015, 13, 42–51. [CrossRef]
- Hall, C.W.; Mah, T.F. Molecular mechanisms of biofilm-based antibiotic resistance and tolerance in pathogenic bacteria. *FEMS Microbiol. Rev.* 2017, 41, 276–301. [CrossRef]
- Gordon, D.M.; O'Brien, C.L. Bacteriocin diversity and the frequency of multiple bacteriocin production in *Escherichia coli*. *Microbiology* 2006, 152, 3239–3244. [CrossRef]
- 51. Baquero, F.; Lanza, V.F.; Baquero, M.R.; Del Campo, R.; Bravo-Vazquez, D.A. Microcins in *Enterobacteriaceae*: Peptide Antimicrobials in the Eco-Active Intestinal Chemosphere. *Front. Microbiol.* **2019**, *10*, 2261. [CrossRef]
- 52. Cameron, A.; Zaheer, R.; Adator, E.H.; Barbieri, R.; Reuter, T.; McAllister, T.A. Bacteriocin Occurrence and Activity in *Escherichia coli* Isolated from Bovines and Wastewater. *Toxins* **2019**, *11*, 475. [CrossRef] [PubMed]
- 53. Weisburg, W.G.; Barns, S.M.; Pelletier, D.A.; Lane, D.J. 16S ribosomal DNA amplification for phylogenetic study. *J. Bacteriol.* **1991**, 173, 697–703. [CrossRef] [PubMed]
- 54. DebRoy, C.; Fratamico, P.M.; Roberts, E. Molecular serogrouping of *Escherichia coli. Anim. Health Res. Rev.* 2018, 19, 1–16. [CrossRef] [PubMed]
- Joensen, K.G.; Tetzschner, A.M.; Iguchi, A.; Aarestrup, F.M.; Scheutz, F. Rapid and Easy In Silico Serotyping of *Escherichia coli* Isolates by Use of Whole-Genome Sequencing Data. J. Clin. Microbiol. 2015, 53, 2410–2426. [CrossRef] [PubMed]
- Dong, A.; Liu, C.; Hua, X.; Yu, Y.; Guo, Y.; Wang, D.; Liu, X.; Chen, H.; Wang, H.; Zhu, L. Bioinformatic analysis of structures and encoding genes of *Escherichia coli* surface polysaccharides sheds light on the heterologous biosynthesis of glycans. *BMC Genom.* 2023, 24, 168. [CrossRef]
- 57. Clermont, O.; Christenson, J.K.; Denamur, E.; Gordon, D.M. The Clermont *Escherichia coli* phylo-typing method revisited: Improvement of specificity and detection of new phylo-groups. *Environ. Microbiol. Rep.* **2013**, *5*, 58–65. [CrossRef] [PubMed]
- Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* 2018, 35, 1547–1549. [CrossRef]
- Rozas, J.; Ferrer-Mata, A.; Sanchez-DelBarrio, J.C.; Guirao-Rico, S.; Librado, P.; Ramos-Onsins, S.E.; Sanchez-Gracia, A. DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. *Mol. Biol. Evol.* 2017, 34, 3299–3302. [CrossRef]
- 60. Huelsenbeck, J.P.; Ronquist, F. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 2001, 17, 754–755. [CrossRef]
- Ewers, C.; Li, G.; Wilking, H.; Kiessling, S.; Alt, K.; Antao, E.M.; Laturnus, C.; Diehl, I.; Glodde, S.; Homeier, T.; et al. Avian pathogenic, uropathogenic, and newborn meningitis-causing *Escherichia coli*: How closely related are they? *Int. J. Med. Microbiol.* 2007, 297, 163–176. [CrossRef]
- Janben, T.; Schwarz, C.; Preikschat, P.; Voss, M.; Philipp, H.C.; Wieler, L.H. Virulence-associated genes in avian pathogenic *Escherichia coli* (APEC) isolated from internal organs of poultry having died from colibacillosis. *Int. J. Med. Microbiol.* 2001, 291, 371–378. [CrossRef] [PubMed]
- 63. Schubert, S.; Rakin, A.; Karch, H.; Carniel, E.; Heesemann, J. Prevalence of the High-Pathogenicity Island Yersinia Species among *Escherichia coli* Strains That Are Pathogenic to Humans. *Infect. Immun.* **1998**, *66*, 480–485. [CrossRef] [PubMed]
- 64. Wang, G.; Clark, C.G.; Rodgers, F.G. Detection in *Escherichia coli* of the genes encoding the major virulence factors, the genes defining the O157:H7 serotype, and components of the type 2 Shiga toxin family by multiplex PCR. *J. Clin. Microbiol.* **2002**, 40, 3613–3619. [CrossRef] [PubMed]
- Johnson, T.J.; Wannemuehler, Y.; Doetkott, C.; Johnson, S.J.; Rosenberger, S.C.; Nolan, L.K. Identification of minimal predictors of avian pathogenic *Escherichia coli* virulence for use as a rapid diagnostic tool. *J. Clin. Microbiol.* 2008, 46, 3987–3996. [CrossRef] [PubMed]
- 66. Toma, C.; Lu, Y.; Higa, N.; Nakasone, N.; Chinen, I.; Baschkier, A.; Rivas, M.; Iwanaga, M. Multiplex PCR assay for identification of human diarrheagenic *Escherichia coli*. *J. Clin. Microbiol.* **2003**, *41*, 2669–2671. [CrossRef] [PubMed]
- 67. Hudzicki, J. Kirby-Bauer disk diffusion susceptibility test protocol. Am. Soc. Microbiol. 2009, 15, 55–63.

- CLSI. Performance Standards for Antimivrobial Susceptibility Testing. In CLSI Supplement M100, 30th ed.; Clinical and Laboratory Standards Institute: Wayne, PA, USA, 2020; ISBN 978-1-68440-066-9 (Print); ISBN 978-1-68440-067-6 (Electronic).
- 69. Wiegand, I.; Hilpert, K.; Hancock, R.E. Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances. *Nat. Protoc.* **2008**, *3*, 163–175. [CrossRef]
- Jayol, A.; Nordmann, P.; Andre, C.; Poirel, L.; Dubois, V. Evaluation of three broth microdilution systems to determine colistin susceptibility of Gram-negative bacilli. J. Antimicrob. Chemother. 2018, 73, 1272–1278. [CrossRef]
- 71. Jing Yu, W.; Pan, T.; En Hui, C.; Li Qin, W.; Wan Hua, L.; Juan Juan, R.; Ning, W.; Yuan Hao, Q.; Hung Jen, L. Characterization of antimicrobial resistance and related resistance genes in *Escherichia coli* strains isolated from chickens in China during 2007–2012. *Afr. J. Microbiol. Res.* 2013, 7, 5238–5247. [CrossRef]
- 72. Dallenne, C.; Da Costa, A.; Decre, D.; Favier, C.; Arlet, G. Development of a set of multiplex PCR assays for the detection of genes encoding important beta-lactamases in Enterobacteriaceae. *J. Antimicrob. Chemother.* **2010**, *65*, 490–495. [CrossRef]
- Chen, S.; Zhao, S.; White, D.G.; Schroeder, C.M.; Lu, R.; Yang, H.; McDermott, P.F.; Ayers, S.; Meng, J. Characterization of multiple-antimicrobial-resistant *Salmonella* serovars isolated from retail meats. *Appl. Environ. Microbiol.* 2004, 70, 1–7. [CrossRef]
- 74. Xiang, T.; Chen, C.; Wen, J.; Liu, Y.; Zhang, Q.; Cheng, N.; Wu, X.; Zhang, W. Resistance of *Klebsiella pneumoniae* Strains Carrying *bla* (NDM-1) Gene and the Genetic Environment of *bla* (NDM-1). *Front. Microbiol.* **2020**, *11*, 700. [CrossRef] [PubMed]
- 75. Ramadan, A.A.; Abdelaziz, N.A.; Amin, M.A.; Aziz, R.K. Novel blaCTX-M variants and genotype-phenotype correlations among clinical isolates of extended spectrum beta lactamase-producing *Escherichia coli. Sci. Rep.* **2019**, *9*, 4224. [CrossRef] [PubMed]
- 76. Komp Lindgren, P.; Karlsson, A.; Hughes, D. Mutation rate and evolution of fluoroquinolone resistance in *Escherichia coli* isolates from patients with urinary tract infections. *Antimicrob. Agents Chemother.* **2003**, *47*, 3222–3232. [CrossRef] [PubMed]
- 77. Giraud, E.; Brisabois, A.; Martel, J.-L.; Chaslus-Dancla, E. Comparative Studies of Mutations in Animal Isolates and Experimental In Vitro- and In Vivo-Selected Mutants of *Salmonella* spp. Suggest a Counterselection of Highly Fluoroquinolone-Resistant Strains in the Field. *Antimicrob. Agents Chemother.* 1999, 43, 2131–2137. [CrossRef]
- 78. Shahrani, M.; Dehkordi, F.S.; Momtaz, H. Characterization of *Escherichia coli* virulence genes, pathotypes and antibiotic resistance properties in diarrheic calves in Iran. *Biol. Res.* 2014, 47, 28. [CrossRef] [PubMed]
- 79. Kozak, G.K.; Boerlin, P.; Janecko, N.; Reid-Smith, R.J.; Jardine, C. Antimicrobial resistance in *Escherichia coli* isolates from swine and wild small mammals in the proximity of swine farms and in natural environments in Ontario, Canada. *Appl. Environ. Microbiol.* **2009**, *75*, 559–566. [CrossRef] [PubMed]
- Aarestrup, F.M.; Lertworapreecha, M.; Evans, M.C.; Bangtrakulnonth, A.; Chalermchaikit, T.; Hendriksen, R.S.; Wegener, H.C. Antimicrobial susceptibility and occurrence of resistance genes among *Salmonella enterica* serovar Weltevreden from different countries. *J. Antimicrob. Chemother.* 2003, *52*, 715–718. [CrossRef]
- 81. Ibrahim, R.A.; Cryer, T.L.; Lafi, S.Q.; Basha, E.A.; Good, L.; Tarazi, Y.H. Identification of *Escherichia coli* from broiler chickens in Jordan, their antimicrobial resistance, gene characterization and the associated risk factors. *BMC Vet. Res.* 2019, 15, 159. [CrossRef]
- 82. Perreten, V.; Boerlin, P. A new sulfonamide resistance gene (sul3) in *Escherichia coli* is widespread in the pig population of Switzerland. *Antimicrob. Agents Chemother.* 2003, 47, 1169–1172. [CrossRef] [PubMed]
- 83. Sunde, M. Prevalence and characterization of class 1 and class 2 integrons in *Escherichia coli* isolated from meat and meat products of Norwegian origin. *J. Antimicrob. Chemother.* **2005**, *56*, 1019–1024. [CrossRef]
- 84. Liu, J.; Zhang, Z.; Feng, Y.; Hu, H.; Yu, Y.; Qiu, L.; Liu, H.; Guo, Z.; Huang, J.; Du, C.; et al. Molecular Detection of the *mcr* Genes by Multiplex PCR. *Infect. Drug Resist.* 2020, *13*, 3463–3468. [CrossRef] [PubMed]
- Maynard, C.; Fairbrother, J.M.; Bekal, S.; Sanschagrin, F.; Levesque, R.C.; Brousseau, R.; Masson, L.; Lariviere, S.; Harel, J. Antimicrobial resistance genes in enterotoxigenic *Escherichia coli* O149:K91 isolates obtained over a 23-year period from pigs. *Antimicrob. Agents Chemother.* 2003, 47, 3214–3221. [CrossRef] [PubMed]
- 86. Dolejska, M.; Cizek, A.; Literak, I. High prevalence of antimicrobial-resistant genes and integrons in *Escherichia coli* isolates from black-headed gulls in the Czech Republic. *J. Appl. Microbiol.* **2007**, *103*, 11–19. [CrossRef] [PubMed]
- Allen, V.G.; Farrell, D.J.; Rebbapragada, A.; Tan, J.; Tijet, N.; Perusini, S.J.; Towns, L.; Lo, S.; Low, D.E.; Melano, R.G. Molecular Analysis of Antimicrobial Resistance Mechanisms in *Neisseria gonorrhoeae* Isolates from Ontario, Canada. *Antimicrob. Agents Chemother.* 2011, 55, 703–712. [CrossRef] [PubMed]
- Kargar, M.; Mohammadalipour, Z.; Doosti, A.; Lorzadeh, S.; Japoni-Nejad, A. High Prevalence of Class 1 to 3 Integrons among Multidrug-Resistant Diarrheagenic *Escherichia coli* in Southwest of Iran. *Osong Public. Health Res. Perspect.* 2014, *5*, 193–198. [CrossRef] [PubMed]
- 89. Tanizawa, Y.; Fujisawa, T.; Nakamura, Y. DFAST: A flexible prokaryotic genome annotation pipeline for faster genome publication. *Bioinformatics* **2018**, *34*, 1037–1039. [CrossRef]
- Larsen, M.V.; Cosentino, S.; Rasmussen, S.; Friis, C.; Hasman, H.; Marvig, R.L.; Jelsbak, L.; Sicheritz-Ponten, T.; Ussery, D.W.; Aarestrup, F.M.; et al. Multilocus sequence typing of total-genome-sequenced bacteria. *J. Clin. Microbiol.* 2012, 50, 1355–1361. [CrossRef]
- 91. Jolley, K.A.; Bray, J.E.; Maiden, M.C.J. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. *Wellcome Open Res.* 2018, *3*, 124. [CrossRef]
- Couvin, D.; Bernheim, A.; Toffano-Nioche, C.; Touchon, M.; Michalik, J.; Neron, B.; Rocha, E.P.C.; Vergnaud, G.; Gautheret, D.; Pourcel, C. CRISPRCasFinder, an update of CRISRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins. *Nucleic Acids Res.* 2018, 46, W246–W251. [CrossRef] [PubMed]

- Biswas, A.; Gagnon, J.N.; Brouns, S.J.; Fineran, P.C.; Brown, C.M. CRISPRTarget: Bioinformatic prediction and analysis of crRNA targets. RNA Biol. 2013, 10, 817–827. [CrossRef] [PubMed]
- 94. Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulos, J.; Bealer, K.; Madden, T.L. BLAST+: Architecture and applications. *BMC Bioinform.* 2009, 10, 421. [CrossRef] [PubMed]
- 95. Alikhan, N.F.; Petty, N.K.; Ben Zakour, N.L.; Beatson, S.A. BLAST Ring Image Generator (BRIG): Simple prokaryote genome comparisons. *BMC Genom.* 2011, 12, 402. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.