JOURNAL OF BACTERIOLOGY, Aug. 2006, p. 5975–5983 Vol. 188, No. 16
0021-9193/06/$08.00+0 doi:10.1128/JB.00204-06

Complete DNA Sequence of a ColBM Plasmid from Avian Pathogenic *Escherichia coli* Suggests that It Evolved from Closely Related ColV Virulence Plasmids†

Timothy J. Johnson, Sara J. Johnson, and Lisa K. Nolan*

Department of Veterinary Microbiology and Preventive Medicine, College of Veterinary Medicine, 1802 Elwood Drive, VMRI 2, Iowa State University, Ames, Iowa 50011

Received 7 February 2006/Accepted 12 May 2006

Avian pathogenic *Escherichia coli* (APEC), an extraintestinal pathogenic *E. coli* causing colibacillosis in birds, is responsible for significant economic losses for the poultry industry. Recently, we reported that the APEC pathotype was characterized by possession of a set of genes contained within a 94-kb cluster linked to a ColV plasmid, pAPEC-O2-ColV. These included *sitABCD*, genes of the aerobactin operon, *hlyF*, *iss*, genes of the salmochelin operon, and the 5′ end of *cvaB* of the ColV operon. However, the results of gene prevalence studies performed among APEC isolates revealed that these traits were not always linked to ColV plasmids. Here, we present the complete sequence of a 174-kb plasmid, pAPEC-O1-ColBM, which contains a putative virulence cluster similar to that of pAPEC-O2-ColV. These two F-type plasmids share remarkable similarity, except that they encode the production of different colicins; pAPEC-O2-ColV contains an intact ColV operon, and pAPEC-O1-ColBM encodes the colicins B and M. Interestingly, remnants of the ColV operon exist in pAPEC-O1-ColBM, hinting that ColBM-type plasmids may have evolved from ColV plasmids. Among APEC isolates, the prevalence of ColBM sequences helps account for the previously observed differences in prevalence between genes of the "conserved" portion of the putative virulence cluster of pAPEC-O2-ColV and those genes within its "variable" portion. These results, in conjunction with Southern blotting and probing of representative ColBM-positive strains, indicate that this "conserved" cluster of putative virulence genes is primarily linked to F-type virulence plasmids among the APEC isolates studied.

Avian pathogenic *Escherichia coli* (APEC), an extraintestinal pathogenic *E. coli* causing colibacillosis in birds, is responsible for significant economic losses for the poultry industry (7), yet the virulence mechanisms underlying APEC virulence are imperfectly understood. Recent studies have sought to define the APEC pathotype and have shown that several genes occur frequently among APEC, regardless of the avian host species or lesion of origin (12, 36, 37). Many of these genes that are characteristic of the APEC pathotype are linked to large plasmids, and many of these plasmids encode production of the bacteriocin ColV (9, 23, 36, 37), leading to their being termed ColV plasmids.

ColV plasmids have long been recognized for their association with *E. coli* virulence (43, 51, 54). However, ColV production, the namesake trait of ColV plasmids, does not appear to contribute to the disease-causing abilities of *E. coli* (35), suggesting that other ColV plasmid-linked traits might be responsible for ColV plasmids' association with virulence. Indeed, we have recently completed the first sequence of a 180-kb ColV plasmid, revealing the presence of a 94-kb cluster of putative virulence genes (23). Based on studies of the prevalence of the genes of this putative virulence cluster among avian *E. coli* isolates, it appears that the cluster contains a "conserved" region, with its genes occurring in nearly 75% or greater of APEC examined, and a "variable" region, with its genes occurring in significantly fewer of the APEC isolates (23). The "conserved" portion of the cluster contains four iron transport and acquisition systems, including the aerobactin and salmochelin siderophore systems and the *sit* and *eit* ABC transporter systems (23). Additionally, the conserved portion of the cluster contains the increased serum survival gene (*iss*), previously implicated in increased virulence and serum resistance (2); *hlyF*, a gene encoding a putative avian *E. coli* hemolysin (29); *ompT*, an outer membrane protease (36); the RepFIB replicon (14); and portions of the ColV operon, including *cvaA* and the 5′ portion of *cvaB*, both ColV export genes (15). Interestingly, the split in gene prevalence between the conserved and variable portions of this virulence cluster occurs within the *cvaB* gene of the ColV operon (23). That is, the 5′ end of *cvaB* and many of its upstream genes occur significantly more often among APEC isolates than do the 3′ end of *cvaB* and many of its downstream genes, suggesting that this virulence cluster may occur in sites in the APEC genome other than on ColV plasmids. In particular, *iroN*, *sitA*, *iutA*, *iss*, *cvaA*, and the 5′ end of *cvaB* occur more frequently among APEC isolates than *cvaC*, the ColV operon's structural gene (23).

Based upon available extraintestinal pathogenic *E. coli* sequencing data and the current literature (8, 30, 33, 48, 52, 53), likely alternative locations of these genes include pathogenicity islands (PAIs) within the bacterial chromosome or non-ColV plasmids. In uropathogenic *E. coli* (UPEC) causing urinary tract infections in humans, several PAIs have been sequenced that contain genes of pAPEC-O2-ColV's putative virulence

* Corresponding author. Mailing address: Department of Veterinary Microbiology and Preventive Medicine, College of Veterinary Medicine, 1802 Elwood Drive, VMRI 2, Iowa State University, Ames, IA 50011. Phone: (515) 294-3470. Fax: (515) 294-3839. E-mail: lknolan@iastate.edu.

† Supplemental material for this article may be found at http://jb.asm.org/.

cluster along with microcins sharing considerable homology with components of the ColV operon (8, 33, 53). Among APEC isolates, studies have identified non-ColV plasmids that possess other genes of pAPEC-O2-ColV's putative virulence cluster, such as genes of the aerobactin operon and *tsh* (16).

The localization of this virulence cluster to "aerobactin" plasmids, which do not encode ColV, might explain the observed link between aerobactin production, complement resistance, and virulence reported by Ike et al. (19). It might also account for the virulence attributed to an aerobactin-producing plasmid described by Ginns et al. (16). Also intriguing in this regard is the observation by Dozois et al. (10) that *tsh*, too, could be found in sites other than on ColV plasmids.

In the present study, a non-ColV plasmid also containing this putative virulence cluster is described. This plasmid is a 174-kb ColBM plasmid known as pAPEC-O1-ColBM. ColBM plasmids have received little attention in the literature (27, 32, 40). These plasmids were originally identified in UPEC strains, where they were found to encode the colicins B and M (32). At 29,453 Da, colicin M (ColM) is the smallest known colicin (17). It inhibits the regeneration of bactoprenyl-P during peptidoglycan synthesis (17). The ColM operon contains the structural gene for ColM activity (*cma*) and the ColM immunity gene (*cmi*), which encodes a protein that binds colicin M to inactivate it (17). Colicin B (ColB) is a 58,000-Da protein which acts to inhibit the membrane potential via channel formation in the cytoplasmic membrane (17). Similar to ColM, ColB activity is mediated by a structural gene (*cba*) and a corresponding immunity gene (*cbi*). These two colicins have been well studied, and the operons encoding them are known to be found together on large, transmissible plasmids (32). Little is known about the plasmids on which the ColB and ColM operons reside (27, 40).

In order to better understand the nature of APEC virulence and evolution, a ColBM plasmid harboring an APEC virulence cluster was sequenced. Additionally, a large collection of APEC isolates was examined for selected genes occurring in this plasmid to determine the prevalence of ColBM sequences among avian *E. coli* isolates and their association with APEC virulence. Furthermore, plasmid preparations of ColBM-positive strains were probed for *cbi* and *iss* in order to verify the plasmid location of such sequences in these strains.

MATERIALS AND METHODS

Bacterial strains and plasmids. pAPEC-O1-ColBM was originally derived from APEC strain O1 (O1:K1), which was isolated from the lung of a turkey with colibacillosis. This APEC strain possesses at least four plasmids, pAPEC-O1-ColBM (174 kb), pAPEC-O1-R (240 kb), pAPEC-O1-Cryptic1 (120 kb), and pAPEC-O1-Cryptic2 (60 kb). These plasmids are transmissible to other bacterial strains by conjugation (22).

Isolates used for the gene prevalence studies were obtained from a variety of sources within the United States, including Georgia, Nebraska, North Dakota, and Minnesota. These isolates originated from sites of infection from birds diagnosed with colibacillosis (APEC).

The positive control strain used for multiplex PCR was APEC O1. *E. coli* DH5α was used as a negative control for all of the genes studied (38). All bacterial strains and subclones were stored at −70°C in brain heart infusion broth (Difco Laboratories, Detroit, MI) with 10% glycerol until use (39).

DNA isolation and preparation for PCR. Total DNA to be used as template for PCR was obtained from APEC O1 and each of the 451 APEC isolates using a boiling lysis procedure previously described (21).

Shotgun library construction and sequencing. pAPEC-O1-ColBM sequences were collected from a random shotgun library of APEC O1 genomic DNA. DNA was sheared, concentrated, and desalted using standard protocols (38). DNA was end repaired (30 min, 15°C; 100-μl reaction mixture: 2 μg sheared DNA, 15 U T4 DNA polymerase, 10 U *E. coli* DNA polymerase [MBI Fermentas, Vilnius, Lithuania], 500 μM each deoxynucleoside triphosphate, 10 μl Yellow Tango buffer [MBI Fermentas]), desalted, and tailed with an extra A residue (30 min, 50°C; 100-μl reaction mixture: 2 μg sheared DNA, 50 μM each dCTP, dGTP, and dTTP, 2 mM dATP, 20 U *Taq* polymerase [MBI Fermentas], 10 μl Yellow Tango buffer). A-tailed DNA was then size fractionated by electrophoresis, and the 1.5- to 2.5-kb fraction was isolated and purified using standard methods (38) prior to cloning into pGEM-T (Promega, Madison, WI).

Shotgun sequencing was performed by MWG Biotech, Inc. (Hedersberg, Germany). Briefly, plasmid clones were grown for 20 h in 1.8 ml LB broth supplemented with 200 μg ml^{-1} ampicillin in deep-well boxes. Plasmid DNA was prepared on a RoboPrep 2500 DNA-Prep-Robot (MWG-Biotech, Ebersberg, Germany) using the NucleoSpin Robot-96 plasmid kit (Macherey & Nagel, Dueren, Germany) and sequenced from both ends with standard primers using the Big Dye terminator chemistry (Applied Biosystems, Foster City, CA). The data were collected with ABI 3700 and ABI 3730xl capillary sequencers.

Gap closure of pAPEC-O1-ColBM was accomplished using the pooled primer technique described by Tettelin et al. and used elsewhere (23, 45). The final gaps were closed using standard PCR.

Assembly and annotation. Sequencing reads were assembled using SeqMan software from DNAStar (Madison, WI). Open reading frames (ORFs) in the plasmid sequence were identified using GeneQuest from DNAStar (Madison, WI), followed by manual inspection. The initiation codon giving the longest coding region was used, including GUG and UUG. ORFs were considered if they were larger than 66 amino acids in size. Translated ORFs were then compared to known protein sequences using BLAST (NCBI; 12/05). Those with greater than 30% amino acid sequence identity over greater than 60% of the sequence were considered matches. Hypothetical proteins with greater than 30% amino acid sequence identity to one or more previously published proteins were classified as conserved hypothetical proteins, and ORFs with less than 30% amino acid sequence identity to any published sequences were classified as hypothetical proteins. The DNA sequence was also analyzed to identify noncoding features. G+C contents of individual ORFs were analyzed using GeneQuest (DNAStar). Insertion sequences and repetitive elements were identified using IS FINDER (http://www-is.biotoul.fr/). Codon usage for each ORF was analyzed using the codon adaptation index versus the *E. coli* K-12 MG1655 genome (3, 52).

Gene prevalence studies. From the completed sequence of pAPEC-O1-ColBM, ColBM-related genes were selected for gene prevalence studies. PCR primers were created for each gene using PrimerSelect from DNAStar, and these primers were used together in a multiplex PCR panel to screen for selected genes of pAPEC-O1-ColBM among the 451 APEC isolates. The genes studied included *cma*, the colicin M activity gene, and *cbi*, the colicin B immunity gene.

All primers, annealing temperatures, and expected amplicon sizes are listed in Table 1. Primers were obtained from Integrated DNA Technologies (Coralville, IA). Genes were amplified in a multiplex panel using a modified version of the multiplex PCR technique previously described (23, 36, 37). PCR was performed with Amplitaq Polymerase Gold (Applied Biosystems, Foster City, CA) according to the manufacturer's instructions. Conditions used for PCR were as

TABLE 1. Primers used in gene prevalence studies

Primer	Gene	Sequence (5′ to 3′)	$T_{annealing}$ (°C)	Amplicon size (bp)	Reference
COLB F	*cbi*	ACAAGACAGCACCA GTTATGGGTATT	63	430	This study
COLB R		GTTGTTGGTTTTGT TGGCGTAGTTAT			
COLM F	*cma*	CAGCGCCATTACCC CATAAATAGTGA	63	498	This study
COLM R		GGTTCGTTCGCCGG TGTAAGCGTTAG			
CVAC F	*cvaC*	CACACACAAACGGG AGCTGTT	63	679	23
CVAC R		CTTCCCGCAGCATA GTTCCAT			

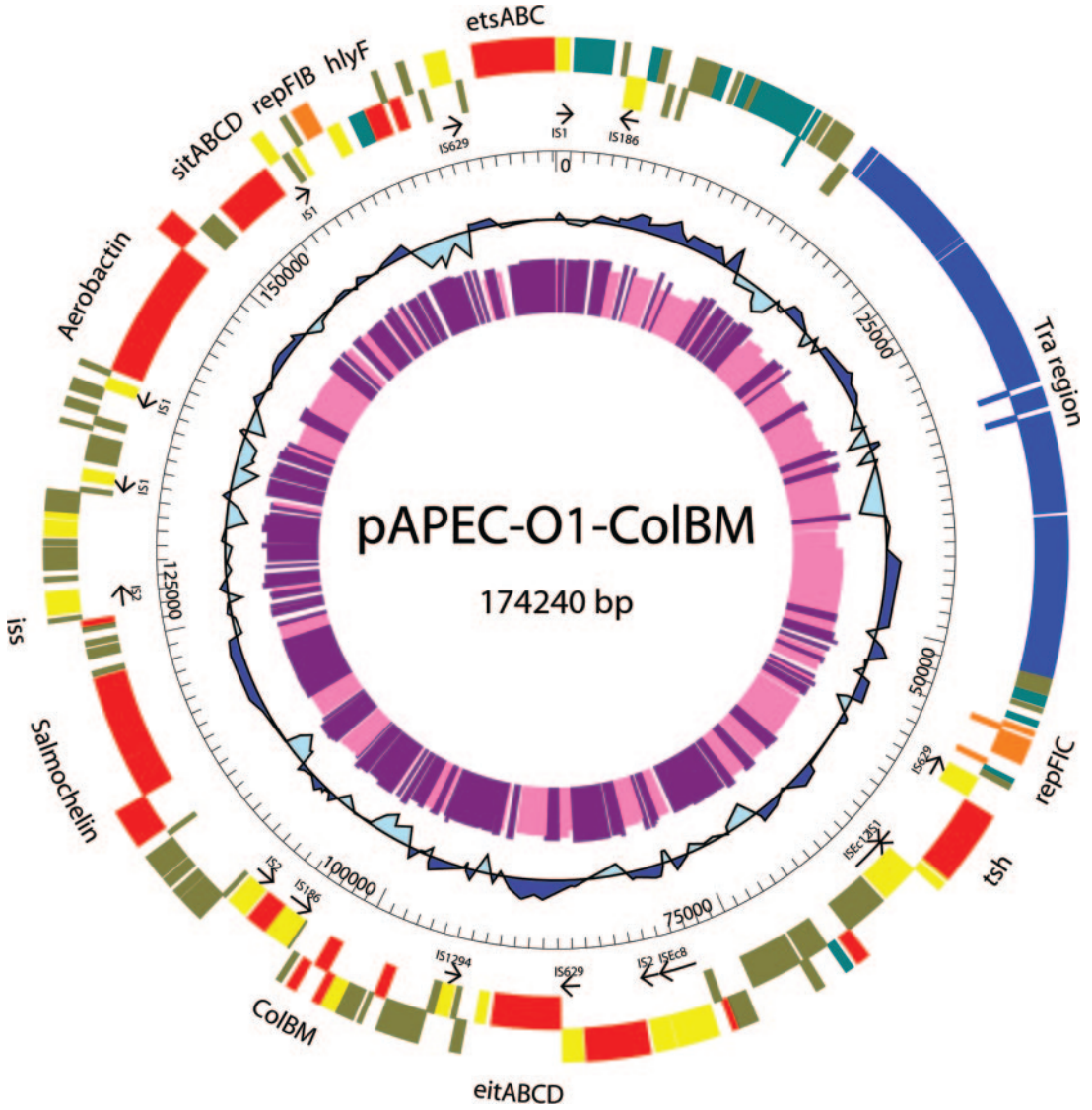

FIG. 1. Map of pAPEC-O1-ColBM. The outer two circles show ORFs in forward and reverse orientation, categorized by color as follows: red, virulence associated; yellow, mobile elements; dark blue, plasmid transfer; orange, plasmid replication; olive, unknown; teal, plasmid maintenance. The third circle shows IS elements with their directions of transcription indicated by arrows. The fourth circle shows the scale in base pairs. The fifth circle shows G+C content plotted against average G+C content for the plasmid of 49.6%, with light blue indicating lower G+C values and dark blue indicating higher G+C values. The sixth circle shows the CAI (1-CAI is plotted), with pink rays indicating CAI values of <0.25 and purple rays indicating CAI values of >0.25. Maps were created using GenVision from DNAStar.

follows: 5 min at 94°C; 30 cycles of 30 seconds at 94°C, 30 seconds at 60°C, and 3 min at 72°C; and a final extension of 10 min at 72°C. Amplicons were visualized on 2.0% Tris-acetate-EDTA (TAE) agarose gels alongside a 1-kb ladder (Promega), and if an amplicon of expected size was observed, then the isolate was considered positive for that particular gene.

Southern hybridization studies. Plasmid DNA was harvested according to the method of Wang and Rossman (50). The DNA obtained was subsequently separated in 0.7% agarose gels using a horizontal gel electrophoresis unit with a constant voltage power source run at 4 V/cm for 6 h. Appropriate digoxigenin-labeled molecular size standards were included on gels (Roche Applied Science, Indianapolis, IN). Amplicons to be labeled as probes for *iss* and *cbi* were generated through DNA amplification techniques mentioned above. Amplicons were separated by horizontal gel electrophoresis, excised from the agarose, and purified with GENECLEAN (Bio 101, Carlsbad, CA). Purified amplicons were digoxigenin labeled, and Southern hybridizations and detections were carried out according to the manufacturer's recommendations (Roche Applied Science).

Comparative genomics. Portions of pAPEC-O1-ColBM were compared to similar plasmid regions using the Artemis comparison tool (4). Protein alignments were performed using the ClustalW algorithm within the DNAStar package (LaserGene). Linear DNA alignments were visualized using Redasoft Visual Cloning version 3.0 (Redasoft Corporation, Toronto, Ontario, Canada).

Nucleotide sequence accession number. The complete sequence and annotation of pAPEC-O1-ColBM are deposited in GenBank under the accession number DQ381420.

RESULTS

Overview of pAPEC-O1-ColBM. Approximately 2,000 clones from the genomic library of APEC O1 were used to assemble the complete sequence of pAPEC-O1-ColBM at approximately 12-

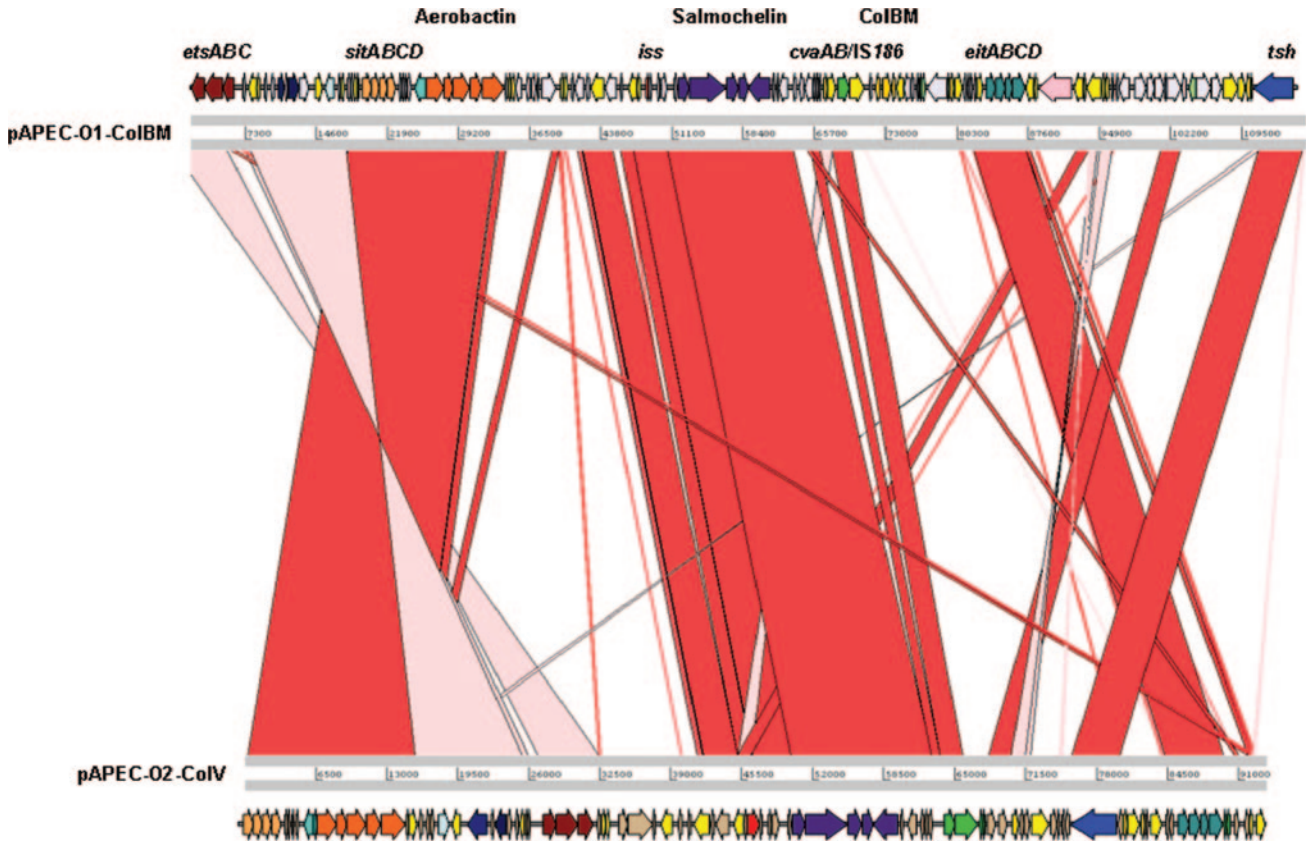

FIG. 2. Nucleotide comparison of the putative virulence regions of pAPEC-O1-ColBM and pAPEC-O2-ColV using the Artemis comparison tool. Same-strand DNA homology is shaded red, while inverse homology is shaded pink. Genes are color coded as follows: maroon, *etsABC*; light orange, *sitABCD*; dark orange, aerobactin operon; purple, salmochelin operon; green, ColV operon; aqua green, *eitABCD*; pink, putative autotransporter/adhesin; blue, *tsh*; red, *iss*; yellow, mobile elements; tan and gray, genes of unknown function.

fold coverage. An end-sequenced fosmid library of APEC O1 was used to verify the final assembly. This plasmid was 174,242 bp in size and found to contain virulence-associated genes, plasmid transfer genes, plasmid maintenance genes, mobile genetic elements, and genes encoding hypothetical proteins of unknown function (Fig. 1). Analysis of the sequence revealed 198 ORFs (see the supplemental material). Of these coding regions, 33 (16.7%) shared homology to IS elements or transposases, 34 (17.2%) with proteins involved in plasmid transfer, and 33 (16.7%) with APEC-associated putative virulence genes. Seventy-four (37.4%) of the identified ORFs on pAPEC-O1-ColBM translated into proteins of unknown function. The overall G+C content of the plasmid was 49.6%. The G+C contents of the individual ORFs were plotted against a map of pAPEC-O1-ColBM, allowing for deviations from the average G+C content of the plasmid to be visualized (Fig. 1). Several regions appeared to deviate sharply from the plasmid G+C content average, including *hlyF*, *eitABCD*, the ColBM operon, and several hypothetical proteins. Additionally, codon usage was analyzed using the codon adaptation index (CAI) with a cutoff value of 0.25 (Fig. 1). These data often mirrored those of the G+C analysis, where ORFs with an extreme G+C content also exhibited differential codon usage. From this analysis, several regions of potentially "foreign" DNA could be discerned, including the ColBM region, the salmochelin operon, *iss*, the RepFIB region, the *etsABC* region, and several regions containing hypothetical proteins (Fig. 1). Interestingly, the aerobactin operon, *sitABCD* gene cluster, and *tsh* gene had CAI values greater than 0.25, suggesting that they were derived from the *E. coli* backbone.

Virulence-associated genes of pAPEC-O1-ColBM. pAPEC-O1-ColBM was found to contain several genes that have previously been associated with APEC virulence (23, 36, 37). These genes were found within a 116-kb region of the plasmid (Fig. 2) and included the following, in order: *etsABC*, genes of a putative ABC transport system (23); *sitABCD*, genes of another ABC transport system involved in iron and manganese transport (5, 23); *iucABCD* and *iutA*, genes of the aerobactin siderophore system (6); *iss*, the increased serum survival gene involved in complement resistance (18, 31, 34, 42); *iroBCDEN*, genes of the salmochelin siderophore system (11); *cvaA* and *cvaB* (truncated), genes of the ColV operon (15); *eitABCD*, genes of a putative iron transport system (23); and *tsh*, the temperature-sensitive hemagglutinin gene (11). All of these genes have been previously localized to a single ColV plasmid, pAPEC-O2-ColV (22, 23). Nucleotide comparison of the 116-kb putative virulence region of pAPEC-O1-ColBM with the 94-kb putative virulence region of pAPEC-O2-ColV revealed several large blocks of highly homologous DNA (Fig. 2). In particular, the portion of pAPEC-O2-ColV previously

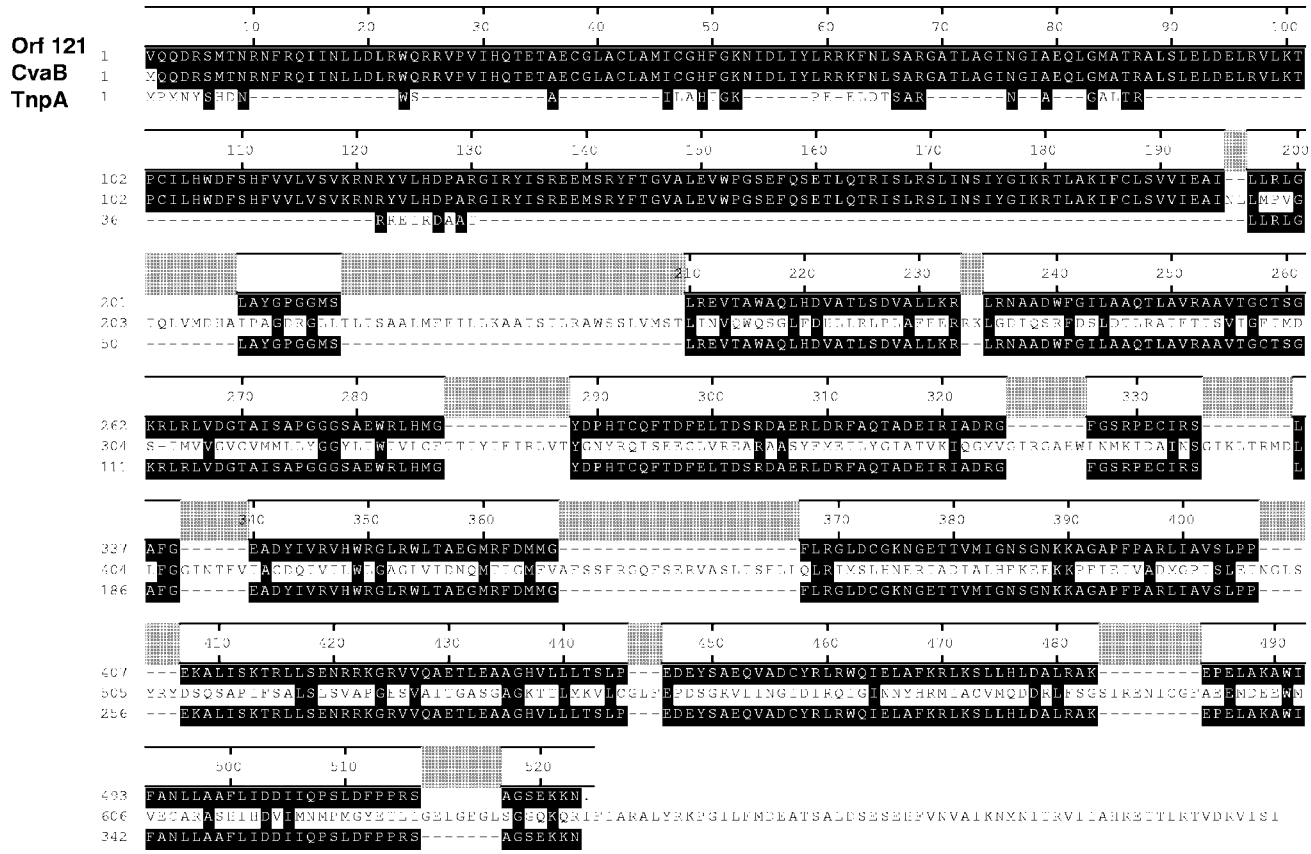

FIG. 3. Amino acid alignment of the CvaB-IS*186* fusion (Orf 121; top) with CvaB (middle) and IS*186* transposase (bottom). Amino acids shaded gray correspond to matches with Orf 121.

identified as "conserved" among APEC isolates is nearly identical to the region of pAPEC-O1-ColBM containing the same genes. The major deviation from this sequence was an inversion and relocation of a colinear block of DNA containing the *etsABC* gene cluster, *hlyF*, *ompT*, and the RepFIB replicon from the 3′ end of the aerobactin operon in pAPEC-O2-ColV to the 5′ end of the *sitABCD* gene cluster in pAPEC-O1-ColBM. The region of pAPEC-O2-ColV that we had previously described as "variable" among APEC isolates shared much less homology with corresponding regions of pAPEC-O1-ColBM. Further analysis of the point at which there is a sharp change in the homology between pAPEC-O2-ColV's and pAPEC-O1-ColBM's virulence clusters revealed that pAPEC-O1-ColBM contains an intact *cvaA* gene of the ColV operon, but it contains a truncated *cvaB* gene and lacks *cvaC* and *cvi* altogether. The 5′ end of the *cvaB* gene in pAPEC-O1-ColBM was found to be fused with an IS*186* element (Fig. 3). This appeared to be the "breakpoint" between high and low DNA homology between pAPEC-O1-ColBM and pAPEC-O2-ColV. Downstream of this fusion was the *colBM* gene cluster, encoding colicins B and M, followed by *eitABCD*, a previously described putative iron transport system (23), followed by a large coding region sharing low identity to a second putative autotransporter/adhesin and *tsh*. This region was also characterized by a large number of hypothetical coding regions and repetitive elements.

Replication and transfer regions of pAPEC-O1-ColBM. pAPEC-O1-ColBM contains a transfer region spanning 32,138 bp, which is similar to those of pAPEC-O2-ColV (23), the F plasmid (GenBank accession number AP00918), and *Escherichia coli* plasmid 1658/97 (GenBank accession number AF550679). All predicted proteins within the transfer region except for *artA*, *trbF*, and *traS* shared 94% or greater BLAST homology with published sequences of these plasmids. The sequence of *traS* varied considerably from previously published sequences, sharing only 26% protein homology with its closest matches. Just upstream of the origin of transfer was *ssb*, the single-stranded binding protein, and *psiA* and *psiB*, plasmid SOS inhibition genes (25). Also found within this region were *sok* and *mok*, involved in postsegregational killing to ensure plasmid stability (46). Downstream of the *traX* gene was *finO*, a fertility inhibition gene (47), *srnB*, a modulator of postsegregational killing (13), and a putative toxin-antitoxin plasmid stability system.

Downstream of the transfer region was one of the two plasmid replicons of pAPEC-O1-ColBM. This region shared greatest homology with the RepFIIA and RepFIC replicons (GenBank BLASTX), albeit with some interesting deviations. For example, the first 600 bp of this region contained the *copB* repressor and flanking DNA sequences from the RepFIC region, but not the *repA1* gene of this replicon that it has been found to repress (26). Instead, this region possessed the *repA1* gene from the

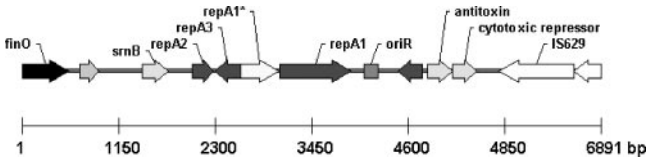

FIG. 4. Genetic map of the RepFIIA region of pAPEC-O1-ColBM. Arrows represent ORFs in the direction of transcription. *repA2*, *repA3*, and *repA1* are components of the RepFIIA replicon, while *repA1∗* is a component of the RepFIC replicon.

RepFIIA replicon (Fig. 4). However, *copB* and *repA1* were separated by two hypothetical proteins, one of which is novel and the other of which shares 100% identity over only 26% of *repA1* from the RepFIC replicon. Downstream of *repA1* in pAPEC-O1-ColBM were the origin of replication (*oriR*) and *repA4* of the RepFIIA replicon. Overall, this replicon shared the most similarities with the RepFIIA replicons of the *E. coli* plasmid 1658/97 (GenBank accession number AF550679) and the *E. coli* ColV plasmid pRK100 (1), although remnants of what has been defined as the RepFIC replicon were also present within this region.

The second replication region found within pAPEC-O1-ColBM was the RepFIB replicon. Within this region were the *repA* replication gene and a site-specific integrase (*int*) involved in recombination (14). This region shared greatest homology with the RepFIB replicons from pAPEC-O2-ColV (23) and *E. coli* plasmid 1658/97 (GenBank accession number AF550679).

Prevalence of plasmid-related genes in APEC. A collection of 451 APEC isolates was examined for *cbi*, the colicin B immunity gene, and *cma*, the colicin M activity gene. Among APEC isolates, *cbi* and *cma* occurred 28% of the time (Table 2). These results were then compared to previous gene prevalence results obtained from this collection of isolates (23) (Table 2). The *cvaC* gene of the ColV operon was previously found to occur at a rate of 64.6% among the APEC isolates tested. Genes of the "conserved" region of pAPEC-O2-ColV, such as the 5′ end of *cvaB*, *iss*, and the aerobactin and *sit* operons, occurred at a significantly higher rate, suggesting that they might occur elsewhere in the APEC genome. Further examination of these isolates for their possession of *cvaC* of the ColV operon and *cbi* and *cma* of the ColB and ColM operons revealed that *cvaC* occurred in conjunction with either *cbi*, *cma*, or both about 20% of the time. *cvaC* itself occurred in the absence of *cbi* and *cma* about 42% of the time. Some isolates lacked *cvaC*, with 16% of these containing genes of the ColB and ColM operons alone. No isolate contained *cma* in the absence of *cvaC* and *cbi*, but several possessed *cbi* while lacking *cma*. Seventeen percent of the APEC examined lacked *cbi*, *cma*, and *cvaC* altogether. When plasmids from strains containing *cbi* or *cma* but not *cvaC* were analyzed by Southern blotting using *cbi* and *iss* sequence probes, both genes localized to large plasmids in nearly every case (Fig. 5).

TABLE 2. Occurrence of genes of the ColV (*cvaC*), ColB (*cbi*), and ColM (*cma*) operons among APEC isolates

Gene combination	% APEC (n = 451)	Operon(s) present
cvaC-cbi-cma	9.8	ColV, ColB, ColM
cvaC-cbi	8.6	ColV, ColB
cvaC-cma	1.5	ColV, ColM
cvaC	47.5	ColV
cbi-cma	13.1	ColB, ColM
cbi	2.9	ColB
cma	0.0	ColM
Absence of *cvaC*, *cbi*, *cma*	16.6	None

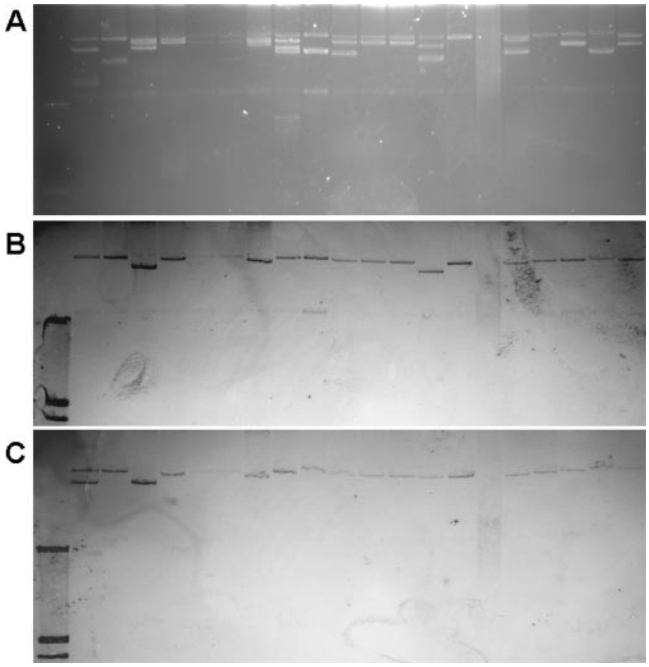

FIG. 5. (A) Agarose gel electrophoresis of selected isolates with a *cbi*$^+$, mutant *cvaC* gene profile. The first lane is a digoxigenin-labeled Southern blot marker with a top band of 21.5 kb. The next 20 lanes are selected *cbi*$^+$ mutant *cvaC* APEC isolates. (B) Southern blot of an agarose gel probed with the *iss* amplicon. (C) Southern blot of an agarose gel probed with the *cbi* amplicon.

DISCUSSION

The plasmid sequenced in this study provides an example of the occurrence of the putative virulence cluster of pAPEC-O2-ColV (23) on a non-ColV plasmid. Comparison of the putative virulence clusters of pAPEC-O2-ColV and pAPEC-O1-ColBM revealed striking similarities (Fig. 2). Greatest DNA homology was observed between the "conserved" portions of these clusters. The primary difference seen within this region was an inversion and relocation of the RepFIB replicon and its surrounding area. However, this is not surprising, considering the mosaic nature of these regions. In both pAPEC-O2-ColV and pAPEC-O1-ColBM, the *sitABCD* and aerobactin systems are flanked by inverted IS*1* elements, providing a means for their transposition. Additionally, several other intact and truncated mobile elements and phage-related elements exist within this region. Thus, if the virulence cluster of pAPEC-O1-ColBM is derived from an ancestral ColV plasmid, recombinational events may have been mediated by IS elements or phages.

Overall, 67% of the predicted proteins of pAPEC-O1-ColBM, not including IS elements, are also found on pAPEC-O2-ColV. Although pAPEC-O1-ColBM and pAPEC-O2-ColV

TABLE 3. Occurrence of Sit and aerobactin among publicly available sequences

System and strain	Organism	Location(s)	Reference
Sit system			
APEC O2	*E. coli*	pAPEC-O2-ColV	AY545598
APEC O1	*E. coli*	pAPEC-O1-ColBM	DQ381420
APEC O1	*E. coli*	Chromosome	DQ335213
APEC χ1722	*E. coli*	pAPEC-1	AY598030
UPEC CFT073	*E. coli*	Chromosome	NC_004431
2a 301	*Shigella flexneri*	Chromosome	NC_004337
2a 2457T	*S. flexneri*	Chromosome	AE016984
Sd197	*Shigella dysenteriae*	Chromosome	CP000034
SL1344	*Salmonella enterica* serovar Typhimurium	Centisome 63-associated PAI	AF128999
Aerobactin operon			
APEC O2	*E. coli*	pAPEC-O2-ColV	AY545598
APEC O1	*E. coli*	pAPEC-O2-ColV	AY545598
N/A	*E. coli*	pColV-K30	X76100
UPEC CFT073	*E. coli*	*pheV*-associated PAI	NC_004431
Nissle 1917	*E. coli*	*pheV*-associated island	AJ586888
2a 301	*S. flexneri*	*pheU*-associated island	NC_004337
M90T	*S. flexneri*	*selC*-associated PAI	AF141323

are similar in size, possess RepFIB and RepFIIA replicons, have nearly identical transfer regions, and possess putative virulence clusters that are very similar in terms of size and orientation, they also have some key differences. One important deviation occurs within the virulence clusters, at the breakpoint where DNA homology between these plasmids diverges. We previously hypothesized that the *cvaB* gene of the ColV operon might have served as a breakpoint during mobilization of the conserved portion of its virulence cluster (23). Analysis of the virulence cluster of pAPEC-ColBM revealed a truncated *cvaB* gene near the ColB and ColM operons, further supporting this idea. Thus, we speculate that pAPEC-O1-ColBM and similar plasmids might have evolved from an ancestral ColV plasmid.

We analyzed 451 APEC isolates for the presence of genes of the ColB and ColM operons in an effort to further explain the previously observed differences in the occurrence of genes of the virulence cluster of pAPEC-O2-ColV (23). Significant differences in prevalence among APEC isolates were previously observed for the "conserved" portion of pAPEC-O2-ColV's putative virulence cluster, compared to the "variable" portion of the cluster. When ColBM- and ColV-specific sequences were compared among APEC isolates, we found that our previous results were more easily interpreted. Sixty-four percent of the APEC isolates examined contained only ColBM- or ColV-specific sequences but not both, likely corresponding to strains possessing ColV or ColBM plasmids with multiple APEC virulence factors. Southern blotting and probing for *cbi* and *iss* confirmed that large ColBM plasmids are the likely location of virulence-associated genes occurring in ColV-negative strains. Taken together, the data in this study suggest that the differences previously observed between the frequencies of *cvaC* and genes of the conserved region of pAPEC-O2-ColV are mostly due to the presence of these genes on ColBM plasmids among APEC isolates.

While the majority of the APEC isolates examined appeared to possess either a ColV or ColBM plasmid, exceptions to the rule were also found. Ten percent of the APEC isolates examined possessed both ColV- and ColBM-associated sequences. We offer two possible explanations for this observation. First, these regions might sometimes occur together on the same plasmid. Perhaps the ColV operon is not always truncated within the *cvaB* gene when this virulence cluster occurs elsewhere, and instead the ColV operon remains intact. Alternatively, perhaps these plasmids sometimes occur together in the same strain. While these plasmids are both IncFIB and thus incompatible by definition, it has been shown that plasmids such as these are capable of coexistence in the same strain (44). This coexistence becomes possible when plasmids possess more than one replicon and are capable of using alternative replication mechanisms in the presence of plasmids of the same Inc group (44). It certainly seems plausible that coexistence of such plasmids may occur in APEC, as pAPEC-O2-ColV and pAPEC-O1-ColBM each contain two plasmid replicons. Also, about 11% of the APEC isolates examined possessed *cbi* and not *cma* of the ColBM region, suggesting that there may be an additional subset of plasmids that exist among APEC isolates and carry virulence-associated genes.

While the combining of gene prevalence results with genomics can be enlightening, some obvious caveats exist. For example, we previously found that *sitA*, *iutA*, and *iroN* occurred more often among APEC isolates than ColV- or ColBM-associated sequences, collectively (23) (Table 2). Therefore, it seems likely that they can also occur in other locations in the genome. Analysis of the GenBank database demonstrates that these genes can be found on the bacterial chromosome, as well as on these plasmids (Table 3). Study of the APEC O1 genome has provided further insight into this discussion, as the *sit* operon in APEC O1 occurs on pAPEC-O1-ColBM and also in its chromosome (GenBank accession number DQ335213) in a location similar to where it is found in UPEC CFT073 (51). Thus, the occurrence of chromosomally encoded APEC virulence factors cannot be ruled out.

We previously identified a cluster of virulence-associated genes on a ColV plasmid which occurred at similar rates among APEC isolates and were strongly associated (23). This

study identifies a non-ColV plasmid location for this cluster, which is pAPEC-O1-ColBM. Analysis of this ColBM plasmid indicates that it may have evolved from ColV plasmids, and based on the results of this study, it appears that the primary locations of virulence-associated genes, which define the APEC pathotype, are on ColV and/or ColBM plasmids. ColV plasmids appear to be the primary type of RepFIB virulence plasmid occurring among APEC isolates, and this is supported by the wealth of recent literature describing ColV-associated sequences occurring among APEC isolates (12, 20, 23, 24, 28, 29, 36, 37, 39, 41, 48, 49). However, ColBM plasmids are an exception to this rule, and the sequence of pAPEC-O1-ColBM provides interesting insights into the possible evolution of F-type plasmids and also into the nature of plasmid-mediated APEC virulence.

ACKNOWLEDGMENTS

This project was funded, in part, by the Roy J. Carver Charitable Trust Fund and the Iowa State University Biotechnology Council, Provost's Office, and College of Veterinary Medicine Dean's Office.

REFERENCES

1. **Ambrozic, J., A. Ostroversnik, M. Starcic, I. Kuhar, M. Grabnar, and D. Zgur-Bertok.** 1998. *Escherichia coli* ColV plasmid pRK100: genetic organization, stability and conjugal transfer. Microbiology **144:**343–352.
2. **Binns, M. M., D. L. Davies, and K. G. Hardy.** 1979. Cloned fragments of the plasmid ColV,I-K94 specifying virulence and serum resistance. Nature **279:** 778–781.
3. **Blattner, F. R., G. Plunkett III, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau, and Y. Shao.** 1997. The complete genome sequence of *Escherichia coli* K-12. Science **277:**1453–1474.
4. **Carver, T. J., K. M. Rutherford, M. Berriman, M. A. Rajandream, B. G. Barrel, and J. Parkhill.** 2005. ACT: the Artemis comparison tool. Bioinformatics **21:**3422–3423.
5. **Crosa, J. H., A. R. Mey, and S. M. Payne.** 2004. Iron transport in bacteria. ASM Press, Washington, D.C.
6. **de Lorenzo, V., and J. B. Neilands.** 1986. Characterization of *iucA* and *iucC* genes of the aerobactin system of plasmid ColV-K30 in *Escherichia coli*. J. Bacteriol. **167:**350–355.
7. **Dho-Moulin, M., and J. M. Fairbrother.** 1999. Avian pathogenic *Escherichia coli* (APEC). Vet. Res. **30:**299–316.
8. **Dobrindt, U., G. Blum-Oehler, G. Nagy, G. Schneider, A. Johann, G. Gottschalk, and J. Hacker.** 2002. Genetic structure and distribution of four pathogenicity islands (PAI I_{536} to PAI IV_{536}) of uropathogenic *Escherichia coli* strain 536. Infect. Immun. **70:**6365–6372.
9. **Doetkott, D. M., L. K. Nolan, C. W. Giddings, and D. L. Berryhill.** 1996. Large plasmids of avian *Escherichia coli* isolates. Avian Dis. **40:**927–930.
10. **Dozois, C. M., M. Dho-Moulin, A. Bree, J. M. Fairbrother, C. Desautels, and R. Curtiss III.** 2000. Relationship between the *tsh* autotransporter and pathogenicity of avian *Escherichia coli* and localization and analysis of the *tsh* genetic region. Infect. Immun. **68:**4145–4154.
11. **Dozois, C. M., F. Daigle, and R. Curtiss III.** 2003. Identification of pathogen-specific and conserved genes expressed in vivo by an avian pathogenic *Escherichia coli* strain. Proc. Natl. Acad. Sci. USA **100:**247–252.
12. **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.
13. **Gerdes, K., L. K. Poulsen, T. Thisted, A. K. Nielsen, J. Martinussen, and P. H. Andreasen.** 1990. The *hok* killer gene family in gram-negative bacteria. New Biol. **2:**946–956.
14. **Gibbs, M. D., A. J. Spiers, and P. L. Bergquist.** 1993. RepFIB: a basic replicon of large plasmids. Plasmid **29:**165–179.
15. **Gilson, L., H. K. Mahanty, and R. Kolter.** 1987. Four plasmid genes are required for colicin V synthesis, export, and immunity. J. Bacteriol. **169:** 2466–2470.
16. **Ginns, C. A., M. L. Benham, L. M. Adams, K. G. Whithear, K. A. Bettelheim, B. S. Crabb, and G. F. Browning.** 2000. Colonization of the respiratory tract by a virulent strain of avian *Escherichia coli* requires carriage of a conjugative plasmid. Infect. Immun. **68:**1535–1541.
17. **Harkness, R. E., and T. Olschlager.** 1991. The biology of colicin M. FEMS Microbiol. Rev. **8:**27–41.
18. **Horne, S. M., S. J. Pfaff-McDonough, C. W. Giddings, and L. K. Nolan.** 2000. Cloning and sequencing of the *iss* gene from a virulent avian *Escherichia coli*. Avian Dis. **44:**179–184.
19. **Ike, K., K. Kawahara, H. Danbara, and K. Kume.** 1992. Serum resistance and aerobactin iron uptake in avian *Escherichia coli* mediated by conjugative 100-megadalton plasmid. J. Vet. Med. Sci. **54:**1091–1098.
20. **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.
21. **Johnson, J. R., and J. J. Brown.** 1996. A novel multiply primed polymerase chain reaction assay for identification of variant *papG* genes encoding the Gal(α1–4)Gal-binding PapG adhesins of *Escherichia coli*. J. Infect. Dis. **173:** 920–926.
22. **Johnson, T. J., C. W. Giddings, S. M. Horne, P. S. Gibbs, R. E. Wooley, J. Skyberg, P. Olah, R. Kercher, J. S. Sherwood, S. L. Foley, and L. K. Nolan.** 2002. Location of increased serum survival gene and selected virulence traits on a conjugative R plasmid in an avian *Escherichia coli* isolate. Avian Dis. **46:**342–352.
23. **Johnson, T. J., K. E. Rodriguez-Siek, S. J. Johnson, and L. K. Nolan.** 2006. DNA sequence of a ColV plasmid and prevalence of selected plasmid-encoded virulence genes among avian *Escherichia coli*. J. Bacteriol. **188:**745–758.
24. **Li, G., C. Laturnus, C. Ewers, and L. H. Wieler.** 2005. Identification of genes required for avian *Escherichia coli* septicemia by signature-tagged mutagenesis. Infect. Immun. **73:**2818–2827.
25. **Loh, S., R. Skurray, J. Celerier, M. Bagdasarian, A. Bailone, and R. Devoret.** 1990. Nucleotide sequence of the *psiA* (plasmid SOS inhibition) gene located on the leading region of plasmids F and R6-5. Nucleic Acids Res. **18:**4597.
26. **Maas, R., and C. Wang.** 1997. Role of the RepA1 protein in RepFIC plasmid replication. J. Bacteriol. **179:**2163–2168.
27. **Massie, B., D. Dea, and A. Sasarman.** 1984. A detailed physical and genetic map of two R.ColBM IncFIII plasmids. Plasmid **11:**105–108.
28. **McPeake, S. J., J. A. Smyth, and H. J. Ball.** 2005. Characterisation of avian pathogenic *Escherichia coli* (APEC) associated with colisepticemia compared to faecal isolates from healthy birds. Vet. Microbiol. **110:**245–253.
29. **Morales, C., M. D. Lee, C. Hofacre, and J. J. Maurer.** 2004. Detection of a novel virulence gene and a *Salmonella* virulence homologue among *Escherichia coli* isolated from broiler chickens. Foodborne Pathog. Dis. **1:**160–165.
30. **Moss, J. E., T. J. Cardozo, A. Zychlinsky, and E. A. Groisman.** 1999. The *selC*-associated SHI-2 pathogenicity island of *Shigella flexneri*. Mol. Microbiol. **33:**74–83.
31. **Nolan, L. K., S. M. Horne, C. W. Giddings, S. L. Foley, T. J. Johnson, A. M. Lynne, and J. Skyberg.** 2003. Resistance to serum complement, *iss*, and virulence of avian *Escherichia coli*. Vet. Res. Commun. **27:**101–110.
32. **Oelschlager, T., E. Schramm, and V. Braun.** 1984. Cloning and expression of the activity and immunity genes of colicins B and M on ColBM plasmids. Mol. Gen. Genet **196:**482–487.
33. **Oelschlaeger, T. A., U. Dobrindt, and J. Hacker.** 2002. Pathogenicity islands of uropathogenic *E. coli* and the evolution of virulence. Int. J. Antimicrob. Agents **19:**517–521.
34. **Pfaff-McDonough, S. J., S. M. Horne, C. W. Giddings, J. O. Ebert, C. Doetkott, M. H. Smith, and L. K. Nolan.** 2000. Complement resistance-related traits among *Escherichia coli* isolates from apparently healthy birds and birds with colibacillosis. Avian Dis. **44:**23–33.
35. **Quackenbush, R. L., and S. Falkow.** 1979. Relationship between colicin V activity and virulence in *Escherichia coli*. Infect. Immun. **24:**562–564.
36. **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.
37. **Rodriguez-Siek, K. E., C. W. Giddings, C. Doetkott, T. J. Johnson, M. K. Fakhr, and L. K. Nolan.** 2005. Comparison of *Escherichia coli* isolates implicated in human urinary tract infection and avian colibacillosis. Microbiology **151:**2097–2110.
38. **Sambrook, J., and D. W. Russell.** 2001. Molecular cloning: a laboratory manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
39. **Sanderson, K. E., and D. R. Zeigler.** 1991. Storing, shipping, and maintaining records on bacterial strains. Methods Enzymol. **204:**248–264.
40. **Sasarman, A., B. Massie, M. Zollinger, H. Gagne-Tellier, F. Shareck, S. Garzon, and R. Morisset.** 1980. Naturally-occurring R.ColBM plasmids belong to the IncFIII incompatibility group. J. Gen. Microbiol. **119:**475–483.
41. **Schouler, C., F. Koffmann, C. Amory, S. Leroy-Setrin, and M. Moulin-Schouleur.** 2004. Genomic subtraction for the identification of putative new virulence factors of an avian pathogenic *Escherichia coli* strain of O2 serogroup. Microbiology **150:**2973–2984.
42. **Skyberg, J., S. M. Horne, C. W. Giddings, R. E. Wooley, P. S. Gibbs, and L. K. Nolan.** 2003. Characterizing avian *Escherichia coli* isolates with multiplex polymerase chain reaction. Avian Dis. **47:**1441–1447.
43. **Smith, H. W., and M. B. Huggins.** 1980. The association of the O18, K1 and H7 antigens and the ColV plasmid of a strain of *Escherichia coli* with its virulence and immunogenicity. J. Gen. Microbiol. **121:**387–400.
44. **Starcic-Erjevac, M., W. Gaastra, J. van Putten, and D. Zgur-Bertok.** 2003.

Identification and the origin of replications and partial characterization of plasmid pRK100. Plasmid **50:**102–112.

45. **Tettelin, H., D. Radune, S. Kasif, H. Khouri, and S. L. Salzberg.** 1999. Optimized multiplex PCR: efficiently closing a whole-genome shotgun sequencing project. Genomics **62:**500–507.
46. **Thisted, T., and K. Gerdes.** 1992. Mechanism of post-segregational killing by the *hok*/*sok* system of plasmid R1. J. Mol. Biol. **223:**41–54.
47. **Timmis, K. N., I. Andres, and M. Achtman.** 1978. Fertility repression of F-like conjugative plasmids: physical mapping of the R6-5 *finO* and *finP* cistrons and identification of the FinO protein. Proc. Natl. Acad. Sci. USA **75:**5836–5840.
48. **Tivendale, K. A., J. L. Allen, C. A. Ginns, B. S. Crabb, and G. F. Browning.** 2004. Association of *iss* and *iucA*, but not *tsh*, with plasmid-mediated virulence of avian pathogenic *Escherichia coli*. Infect. Immun. **72:**6554–6560.
49. **Vandekerchove, D., F. Vandemaele, C. Adriaensen, M. Zaleska, J. P. Hernalsteens, L. D. Baets, P. Butaye, F. V. 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.
50. **Wang, Z., and T. G. Rossman.** 1994. Large-scale supercoiled plasmid preparation by acidic phenol extraction. BioTechniques **16:**460–463.
51. **Waters, V. L., and J. H. Crosa.** 1991. Colicin V virulence plasmids. Microbiol. Rev. **55:**437–450.
52. **Wei, J., M. B. Goldberg, V. Burland, M. M. Venkatesan, W. Deng, G. Fournier, G. F. Mayhew, G. Plunkett III, D. J. Rose, A. Darling, B. Mau, N. T. Perna, S. M. Payne, L. J. Runyen-Janecky, S. Zhou, D. C. Schwartz, and F. R. Blattner.** 2003. Complete genome sequence and comparative genomics of *Shigella flexneri* serotype 2a strain 2457T. Infect. Immun. **71:**2775–2786.
53. **Welch, R. A., V. Burland, G. Plunkett III, P. Redford, P. Roesch, D. Rasko, E. L. Buckles, S. R. Liou, A. Boutin, J. Hackett, D. Stroud, G. F. Mayhew, D. J. Rose, S. Zhou, D. C. Schwartz, N. T. Perna, H. L. Mobley, M. S. Donnenberg, and F. R. Blattner.** 2002. Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*. Proc. Natl. Acad. Sci. USA **99:**17020–17024.
54. **Wooley, R. E., L. K. Nolan, J. Brown, P. S. Gibbs, C. W. Giddings, and K. S. Turner.** 1993. Association of K-1 capsule, smooth lipopolysaccharides, traT gene, and colicin V production with complement resistance and virulence of avian *Escherichia coli*. Avian Dis. **37:**1092–1096.