8-2006

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  
Iowa State University

Sara J. Johnson  
Iowa State University

Lisa K. Nolan  
Iowa State University, lknolan@iastate.edu

Follow this and additional works at: http://lib.dr.iastate.edu/vmpm_pubs

Part of the Genomics Commons, Veterinary Microbiology and Immunobiology Commons, and the Veterinary Preventive Medicine, Epidemiology, and Public Health Commons

The complete bibliographic information for this item can be found at http://lib.dr.iastate.edu/vmpm_pubs/27. For information on how to cite this item, please visit http://lib.dr.iastate.edu/howtocite.html.
Complete DNA Sequence of a ColBM Plasmid from Avian Pathogenic Escherichia coli Suggests that It Evolved from Closely Related ColV Virulence Plasmids

Abstract
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.

Disciplines
Genomics | Veterinary Microbiology and Immunobiology | Veterinary Preventive Medicine, Epidemiology, and Public Health

Comments
This article is from Journal of Bacteriology 188, no. 16 (August 2006): 5975–5983, 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 (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, 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, pAPEC-O2-ColV, 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...
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 ColM 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 (cbb) 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-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 tailing an extra A residue (30 min, 50°C, 100-μl reaction mixture: 2 μg sheared DNA, 50 μM each dCTP, dGTP, and dTTP, 2 mU 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.** This study 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 ampicillin in deep-well plates. Plasmid DNA was then prepared on a RoboPrep 2500 DNA-Prep Robot (MWG-Biotech, Ebersberg, Germany) using the NucleoSpin Robot-96 plasmid kit (Machery & Nagel, Düren, 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.

**Gene prevalence studies.** The complete sequence 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.** 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 (Coraville, 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

<table>
<thead>
<tr>
<th>Primer</th>
<th>Gene</th>
<th>Sequence (5’ to 3’)</th>
<th>Temperature (°C)</th>
<th>Amplicon size (bp)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>COLB F</td>
<td>cbi</td>
<td>ACAAGACAGACGAC</td>
<td>63</td>
<td>430</td>
<td>This study</td>
</tr>
<tr>
<td>COLB R</td>
<td></td>
<td>GTATTGGGTATT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COLM F</td>
<td>cma</td>
<td>CACGCCCATATACCC</td>
<td>63</td>
<td>498</td>
<td>This study</td>
</tr>
<tr>
<td>COLM R</td>
<td></td>
<td>CATTAATAGTGGA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CVAC F</td>
<td>cvaC</td>
<td>CACACACAAACGCGG</td>
<td>63</td>
<td>679</td>
<td>23</td>
</tr>
<tr>
<td>CVAC R</td>
<td></td>
<td>AGCTGTTC</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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-
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 average G+C content of the plasmid (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); troBCDEN, genes of the salmochelin siderophore system (11); cvaA and cvaB (truncated), genes of the ColV operon (15); etsABCD, 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. 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, etsABCD; pink, putative autotransporter/adhesin; blue, tsh; red, iss; yellow, mobile elements; tan and gray, genes of unknown function.
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 \textit{etsABC} gene cluster, \textit{hlyF}, \textit{ompT}, and the RepFIB replicon from the 3’/H11032 end of the aerobactin operon in pAPEC-O2-ColV to the 5’/H11032 end of the \textit{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 \textit{cvaA} gene of the ColV operon, but it contains a truncated \textit{cvaB} gene and lacks \textit{cvaC} and \textit{cvi} altogether. The 5’ end of the \textit{cvaB} gene in pAPEC-O1-ColBM was found to be fused with an IS186 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 \textit{colBM} gene cluster, encoding colicins B and M, followed by \textit{eitABCD}, a previously described putative iron transport system (23), followed by a large coding region sharing low identity to a second putative auto-transporter/adhesin and \textit{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 \textit{Escherichia coli} plasmid 1658/97 (GenBank accession number AF550679). All predicted proteins within the transfer region except for \textit{artA}, \textit{trbF}, and \textit{traS} shared 94% or greater BLAST homology with published sequences of these plasmids. The sequence of \textit{traS} varied considerably from previously published sequences, sharing only 26% protein homology with its closest matches. Just upstream of the origin of transfer was \textit{ssb}, the single-stranded binding protein, and \textit{psiA} and \textit{psiB}, plasmid \textit{SOS} inhibition genes (25). Also found within this region were \textit{sok} and \textit{mok}, involved in postsegregational killing to ensure plasmid stability (46). Downstream of the \textit{traX} gene was \textit{finO}, a fertility inhibition gene (47), \textit{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 \textit{copB} repressor and flanking DNA sequences from the RepFIC region, but not the \textit{repA1} gene of this replicon that it has been found to repress (26). Instead, this region possessed the \textit{repA1} gene from the
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 CoIB 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 *cma* sequence probes, both genes localized to large plasmids in nearly every case (Fig. 5).

**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 IS1 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 2. Occurrence of genes of the ColV (cvaC), CoIB (cbi), and CoLM (cma) operons among APEC isolates**

<table>
<thead>
<tr>
<th>Gene combination</th>
<th>% APEC (n = 451)</th>
<th>Operon(s) present</th>
</tr>
</thead>
<tbody>
<tr>
<td>cvaC-chi-cma</td>
<td>9.8</td>
<td>ColV, CoIB, CoLM</td>
</tr>
<tr>
<td>cvaC-chi</td>
<td>8.6</td>
<td>ColV, CoIB</td>
</tr>
<tr>
<td>cvaC-cma</td>
<td>1.5</td>
<td>CoIB, CoLM</td>
</tr>
<tr>
<td>cvaC</td>
<td>47.5</td>
<td>CoIV</td>
</tr>
<tr>
<td>chi-cma</td>
<td>13.1</td>
<td>CoIB, CoLM</td>
</tr>
<tr>
<td>chi</td>
<td>2.9</td>
<td>CoIB</td>
</tr>
<tr>
<td>cma</td>
<td>0.0</td>
<td>CoLM</td>
</tr>
<tr>
<td>Absence of cvaC, chi, cma</td>
<td>16.0</td>
<td>None</td>
</tr>
</tbody>
</table>
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 its 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 genome views 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 DQ2235213) 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).
study identifies a non-ColV plasmid location for this cluster, which is pAPEC-O1-ColBMM. Analysis of this ColBMM 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 ColBMM 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, ColBMM plasmids are an exception to this rule, and the sequence of pAPEC-O1-ColBMM 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


