

Comparison of *Escherichia coli* isolates implicated in human urinary tract infection and avian colibacillosis

Kylie E. Rodriguez-Siek,^{1†} Catherine W. Giddings,² Curt Doetkott,³ Timothy J. Johnson,¹ Mohamed K. Fakhr² and Lisa K. Nolan¹

Correspondence

Lisa K. Nolan

lknolan@iastate.edu

¹Department of Veterinary Microbiology and Preventive Medicine, College of Veterinary Medicine, Iowa State University, Ames, IA 50011, USA

^{2,3}Department of Veterinary and Microbiological Sciences² and Information Technology Services³, North Dakota State University, Fargo, ND 58105, USA

Since avian pathogenic *Escherichia coli* (APEC) and human uropathogenic *E. coli* (UPEC) may encounter similar challenges when establishing infection in extraintestinal locations, they may share a similar content of virulence genes and capacity to cause disease. In the present study, 524 APEC and 200 UPEC isolates were compared by their content of virulence genes, phylogenetic group, and other traits. The two groups showed substantial overlap in terms of their serogroups, phylogenetic groups and virulence genotypes, including their possession of certain genes associated with large transmissible plasmids of APEC. Based on these results, the propensity of both groups to cause extraintestinal infections, and a well-documented ability of avian *E. coli* to spread to human beings, the potential for APEC to act as human UPEC or as a reservoir of virulence genes for UPEC should be considered. However, significant differences in the prevalence of the traits occurred across the two groups, suggesting that if APEC are involved in human urinary tract infections, they are not involved in all of them.

Received 20 July 2004

Revised 3 February 2005

Accepted 14 April 2005

INTRODUCTION

Urinary tract infections (UTIs) are among the most common bacterial infections (Foxman, 2002), causing significant morbidity and mortality and costing the healthcare system of the United States well over a billion dollars annually (Foxman, 2002; Russo & Johnson, 2003). Most community-acquired UTIs are due to uropathogenic *E. coli* (UPEC) infections (Johnson, 1991; Donnenberg & Welch, 1996). UPEC are uniquely endowed with various virulence traits, enabling them to survive and grow in urine and other extraintestinal environments. The abilities of UPEC to grow extraintestinally may enable them to cause a variety of diseases, not just urinary tract ones. This broad potential to cause disease has led Russo & Johnson (2000) to propose that UPEC be incorporated in a new category known as extraintestinal pathogenic *E. coli* (ExPEC). Similarly, avian pathogenic *Escherichia coli* (APEC) are known for their broad potential to cause extraintestinal diseases in birds

(Barnes *et al.*, 2003). Since APEC and human ExPEC may encounter similar challenges when establishing infection in these locations, they may share a similar content of virulence genes and capacity to cause disease. If these pathogens are similar in genetic constitution, the potential of APEC to serve as a source of human ExPEC or as a reservoir of virulence genes for human ExPEC would need to be considered. Of special interest in this regard are certain genes, found on large plasmids in APEC, that are known to occur on UPEC plasmids (Sorsa *et al.*, 2003) and pathogenicity islands (PAIs) (Dobrindt *et al.*, 2001, 2002; Oelschlaeger *et al.*, 2002b). As a first step in assessing the potential of APEC to cause human extraintestinal diseases, 200 UPEC and 524 APEC were compared for their content of virulence genes, including many implicated in ExPEC virulence as well as those associated with APEC plasmids. Other traits were also examined. These results were then statistically analysed in an effort to detect similarities and differences between the APEC and UPEC isolates studied here.

METHODS

Bacterial strains. A total of 724 isolates were used in this study, including 200 isolates from cases of human UTIs, kindly provided by Dr Paul Carson (Meritcare Hospital, Fargo, ND), and 524 *E. coli* isolates implicated in avian colibacillosis that were obtained from

Abbreviations: APEC, avian pathogenic *Escherichia coli*; ddH₂O, double distilled water; ExPEC, extraintestinal pathogenic *Escherichia coli*; LDA, linear discriminant analysis; PAI, pathogenicity island; PCs, principal components; PCA, principal components analysis; UPEC, uropathogenic *E. coli*; UTI, urinary tract infection.

[†]Done in partial fulfillment of the requirements for the MS degree in Microbiology at North Dakota State University.

various locations within the USA. A subset of these APEC has been described previously (Rodriguez-Siek *et al.*, 2005). Isolates were serogrouped through the *Escherichia coli* Reference Center (Pennsylvania State University) and screened for traits associated with the virulence of ExPEC and APEC, using the techniques described below. Organisms were stored at -80°C in brain heart infusion broth (Difco) with 10% (v/v) glycerol until use (Sanderson & Zeigler, 1991).

Haemolytic reaction. Test and control organisms were assessed for their haemolytic reaction on 5% sheep blood agar plates using standard methods (Forbes *et al.*, 1998).

Fermentation of lactose. Test and control organisms were assessed for lactose utilization by plating on MacConkey agar (Difco) using standard methods (Forbes *et al.*, 1998).

Phylogenetic typing. Isolates were assigned to phylogenetic groups according to the method of Clermont *et al.* (2000). Using this method, isolates are assigned to one of four groups (A, B1, B2 or D)

based on their possession of two genes (*chuA* and *yjaA*) and a DNA fragment (TSPE4.C2) (Table 1), as determined by PCR. Boiled lysates of overnight cultures were used as a source of template DNA for this study (Johnson & Brown, 1996). Amplification was performed in a 25 μl reaction mixture containing 18.3 μl double-distilled H_2O (dd H_2O), 2.5 μl 10 \times PCR Buffer (Invitrogen), 1.0 μl 50 mM MgCl_2 , 0.5 μl of a 2.5 mM dNTP mixture (USB), 0.075 μl of 0.1 mM upper and lower primers (Integrated DNA Technologies) (Table 2), 0.25 μl (5 U μl^{-1}) *Taq* DNA polymerase (Invitrogen) and 2.0 μl template DNA. The reaction mixture was subjected to the following parameters in a Mastercycler Gradient thermocycler (Brinkmann Eppendorf): 4 min at 94°C , 30 cycles of 5 s at 94°C and 10 s at 59°C , and a final extension step of 5 min at 72°C , followed by a hold at 4°C .

Samples were subjected to horizontal gel electrophoresis in 1.5% (w/v) agarose, and the size of the amplicons was determined by comparison to the Hi-Lo DNA marker (Minnesota Molecular Inc.). Strains known to possess or lack the genes of interest were examined with each

Table 1. ExPEC/APEC genes used in virulence genotyping and phylogenetic typing

Descriptions of genes encoding components of certain adhesins (i.e. genes encoding parts of the P pilus, *papA*; *papC*; *papEF*; *papG*, including *papG* alleles I, II, and III; the S pilus, *sfa* and the gene encoding the S fimbrial tip, *sfaS*; the Type 1 fimbrial adhesin, *fimH*; the F1C fimbrial tip, *focG*; and other genes encoding portions of miscellaneous adhesins, *iha*; *afa*; *gafD*; and *bmaE*); toxins (*cnf-1* and *cdtB*); protectins (*kpsMT* K1, *kpsMT* II, *kpsMT* III and *rfa*); siderophores (*fyuA*); and other miscellaneous structures [*ibeA*, *ompT*, and *malX* (PAI)] can be found in Johnson *et al.* (2001). Also, the description of *papG* allele I' can be found in Johnson & Stell (2000).

Gene	Description	Source or reference
pTJ100-related genes		
<i>cvaC</i> *	Structural gene for the colicin V operon	Gilson <i>et al.</i> (1987)
<i>iroN</i> †	Catecholate siderophore receptor gene	Russo <i>et al.</i> (1999)
<i>iss</i> *	Increased serum survival gene	Binns <i>et al.</i> (1979); Chuba <i>et al.</i> (1989); Pfaff-McDonough <i>et al.</i> (2000)
<i>iucC</i> †	Involved in aerobactin synthesis	de Lorenzo & Neilands (1986); de Lorenzo <i>et al.</i> (1986)
<i>iutA</i> †	Ferric aerobactin receptor gene; iron transport	de Lorenzo <i>et al.</i> (1986)
<i>sitA</i> †	Putative iron transport gene	Runyen-Janecky <i>et al.</i> (2003)
<i>traT</i> *	Outer-membrane protein gene; serum resistance	Achtman <i>et al.</i> (1977); Moll <i>et al.</i> (1980)
<i>tsh</i> ‡	Temperature-sensitive haemagglutinin gene	Provence & Curtiss (1994)
Iron-related		
<i>feoB</i>	Gene which mediates ferric iron uptake	Runyen-Janecky <i>et al.</i> (2003)
<i>ireA</i>	Encodes an iron-responsive element	Russo <i>et al.</i> (2001)
<i>irp-2</i>	Iron-repressible gene associated with yersiniabactin synthesis	Schubert <i>et al.</i> (1998)
Toxins		
<i>hlyD</i>	Transport gene of the haemolysin operon	Hacker <i>et al.</i> (1988)
Miscellaneous		
<i>fliC</i> (H7)	Produces flagellin protein associated with the H7 antigen group	Fields <i>et al.</i> (1997)
Phylogenetic typing		
<i>chuA</i>	Haem transport gene	Clermont <i>et al.</i> (2000)
<i>yjaA</i>	Gene of unknown function from the <i>E. coli</i> K-12 genome	Clermont <i>et al.</i> (2000)
TSPE4.C2	Anonymous DNA fragment	Bonacorsi <i>et al.</i> (2000); Clermont <i>et al.</i> (2000)

*These genes are listed as pTJ100-related, but they could also be listed as protectins.

†These genes are listed as pTJ100-related, but they could also be listed with the iron-related genes.

‡This gene is listed as pTJ100-related, but it also could be listed in the miscellaneous group.

Table 2. Primers used for the amplification studies

Gene	Amplicon size (bp)	Primer sequence (5'–3')	Reference
pTJ100-related genes			
<i>cvaC</i> *	679	F cacacacaaacgggagctgtt R cttcccgagcatagttccat	Johnson & Stell (2000)
<i>iroN</i> †	667	F aagtcaaagcaggggtgcccg R gacgccgacattaagacgcag	J. R. Johnson protocols
<i>iss</i> *	323	F cagcaaccggaaccacttgatg R agcattgccagagcggcagaa	J. R. Johnson protocols
<i>iucC</i> †	541	F cgccgtggctggggaag R cagccggttcaccaagtatcactg	Skyberg <i>et al.</i> (2003)
<i>iutA</i> †	302	F ggctggacatcatgggaactgg R cgtcgggaacgggtagaaatcg	Johnson & Stell (2000)
<i>sitA</i> †	608	F aggggggcacaaactgattctcg R taccgggcccgttttctgtgc	Runyen-Janecky <i>et al.</i> (2003)
<i>traT</i> *	290	F ggtgtggtgcgatgagcacag R cacggttcagccatccctgag	Johnson & Stell (2000)
<i>tsh</i> ‡	420	F gggaaatgacctgaatgctgg R ccgctcatcagtcagtaccac	Maurer <i>et al.</i> (1998)
Adhesins			
<i>afa</i>	594	F ggcagagggccggcaacaggc R cccgtaacgcgcagcatctc	Johnson & Stell (2000)
<i>bmaE</i>	507	F atggcgctaacttgccatgctg R agggggacatatagcccccttc	Johnson & Stell (2000)
<i>fimH</i>	508	F tcgagaacggataagccgtgg R gcagtcacctgccctccgga	Johnson & Stell (2000)
<i>gafD</i>	952	F tgttggaccgtctcagggtctc R tcccggaaactgctgttact	Johnson & Stell (2000)
<i>iha</i>	829	F ctggcggaggctctgagatca R tccttaagctcccgcggctga	J. R. Johnson protocols
<i>papA</i>	717	F atggcagtggtgtttgttg R cgtcccaccatacgtgctcttc	Johnson & Stell (2000)
<i>papC</i>	205	F gtggcagtatgagtaatgaccgtta R atatcctttctgcaggatgaata	Johnson & Stell (2000)
<i>papEF</i>	326	F gcaacagcaacgctggtgcatcat R agagagagccactcttatacggaca	Johnson & Stell (2000)
<i>papG</i> §			
<i>papG</i> allele I§	461	F tcgtgctcagggtccggaattt R tggcatccccaacattatcg	Johnson & Stell (2000)
<i>papG</i> allele I'§	479	F ctactatagttcatgctcaggtc R cctgcatctccaccattatcga	Johnson & Stell (2000)
<i>papG</i> allele II§	190	F gggatgagcgggcctttgat R cgggcccccaagtaactcg	Johnson & Stell (2000)
<i>papG</i> allele III§	258	F ggctgcaatggatttacctgg R ccaccaaatgaccatgccagac	Johnson & Stell (2000)
<i>sfa</i>	410	F ctccggagaactgggtcatcttac R cggaggagtaattacaacctggca	Johnson & Stell (2000)
<i>sfaS</i>	244	F gtggatacgacgattactgtg R ccgccagattccctgtattc	Johnson & Stell (2000)
<i>focG</i>	364	F cagcacaggcagtgatagca R gaatgtcgcctgccattgct	Johnson & Stell (2000)
Iron-related			
<i>feoB</i>	470	F aattggcgtgcatgaagataactg R agctggcgacctgatagaacaatg	Runyen-Janecky <i>et al.</i> (2003)

Table 2. cont.

Gene	Amplicon size (bp)	Primer sequence (5'–3')	Reference
<i>fyuA</i>	787	F tgattaaccccgacgggaa R cgagtaggcacgatgttgta	Johnson & Stell (2000)
<i>ireA</i>	254	F gatgactcagccacgggtaa R ccaggactcacctcacgaat	J. R. Johnson protocols
<i>irp-2</i>	287	F aaggattcgtgttacggac R tcgtcgggcagcgtttcttct	Janben <i>et al.</i> (2001)
Protectins			
<i>kpsMT</i> (K1)	153	F tagcaaacgttctattggtgc use with <i>kpsMT</i> II-Reverse	Johnson & Stell (2000)
<i>kpsMT</i> II	272	F gcgcatttgctgatactgttg R catccagacgataagcatgagca	Johnson & Stell (2000)
<i>kpsMT</i> III	392	F tcctcttgctactattccccct R aggcgtatccatccctcctaac	Johnson & Stell (2000)
<i>rfc</i>	788	F atccatcaggaggggactgga R aaccataccaaccaatgcgag	Johnson & Stell (2000)
Toxins			
<i>cdtB</i>	430	F gaaaataaatggaacacacatgtccg F' gaaagtaaatggaatataaatgtccg R aaatctcctgcaatcatccagtta R' aaatcaccaagaatcatccagtta	Johnson & Stell (2000)
<i>cnf-1</i>	1105	F atcttatactggatgggatcatcttg R gcagaacgacgttctcataagtatc	Johnson & Stell (2000)
<i>hlyD</i>	904	F ctccggtacgtgaaaaggac R gccctgattactgaagcctg	J. R. Johnson protocols
Miscellaneous			
<i>fliC</i> (H7)	547	F acgatgcaggcaactgacg R ggggtggtcgttgcaaac	J. R. Johnson protocols
<i>ibeA</i>	171	F aggcaggtgtgcccgcgtac R tgggtgtccggcaaacatgc	Johnson & Stell (2000)
<i>ompT</i>	559	F atctagccgaagaaggaggc R cccgggtcatagtgttcac	J. R. Johnson protocols
<i>malX</i> (PA1)	925	F ggacatcctgttacagcgcgca R tcgccaccaatcacagccgaac	Johnson & Stell (2000)
Phylogenetic typing			
<i>chuA</i>	279	F gacgaaccaacggctcaggat R tgccgccagttaccaagaca	Clermont <i>et al.</i> (2000)
<i>yjaA</i>	211	F tgaagtgtcaggagacgtg R atggagaatgcgttctcaac	Clermont <i>et al.</i> (2000)
TSPE4.C2	152	F gagtaatgtcggggcattca R cgcgccaacaagtattacg	Clermont <i>et al.</i> (2000)

*These genes are listed as pTJ100-related, but they could also be listed with the protectins.

†These genes are listed as pTJ100-related, but they could also be listed with the iron-related genes.

‡This gene is listed as pTJ100-related, but it also could be listed in the miscellaneous group.

\$*spgG* genotypes were assessed using an allele-specific assay developed by Johnson & Brown (1996). Protocols and control strains for this procedure were obtained directly from Dr James R. Johnson, VA Medical Center, Minneapolis, MN.

amplification procedure. An isolate was considered to contain a gene of interest if it produced an amplicon of the expected size (Table 2). Isolates were assigned to phylogenetic group A if they exhibited the profile: *chuA* (–), *yjaA* (+/–) and TSPE4.C2 (–); B1 by the profile of *chuA* (–), *yjaA* (+/–) and TSPE4.C2 (+); B2 by the genotype of *chuA* (+), *yjaA* (+) and TSPE4.C2 (+/–); or D if they were *chuA* (+), *yjaA* (–) and TSPE4.C2 (+/–).

Virulence genotyping. Test and control organisms were examined for the presence of several genes (Table 1) known for their association with ExPEC or APEC virulence, using multiple PCR assays. The APEC genes of interest have been associated with APEC plasmids, such as pTJ100 (Johnson *et al.*, 2002, 2004). pTJ100 is a large plasmid known to contain many of the genes associated with APEC virulence, including the iron-acquisition operons, aerobactin

(Johnson *et al.*, 2002), *iro* (GenBank accession no. AY545598) and *sit* (GenBank accession no. AY553855); the ColV operon (Johnson *et al.*, 2002); certain genes associated with serum resistance, *iss* (Johnson *et al.*, 2002) and *traT* (Johnson *et al.*, 2004); and *tsh* (Johnson *et al.*, 2002). Interestingly, pTJ100 also shows similarities in genetic content to a UPEC plasmid (Sorsa *et al.*, 2003) and PAI (Dobrindt *et al.*, 2001, 2002; Oelschlaeger *et al.*, 2002a,b).

All primers used in amplification of the virulence genes were obtained from Sigma-Genosys and Integrated DNA Technologies (Table 2). Template DNA for all amplifications was generated as described elsewhere (Johnson & Brown, 1996). Targeted genes were amplified in multiplex procedures or singly; all amplification procedures have been documented elsewhere. Six different multiplex procedures, targeting different combinations of genes, were used. Five of these were described by Johnson & Stell (2000) and targeted: 1) a PAI of UPEC CFT073, *papA*, *fimH*, *kpsMT* III, *papEF*, *ireA* and *ibeA*; 2) *cnf-1*, *fyuA*, *iroN*, *bmaE*, *sfa*, *iutA* and *papG* allele III; 3) *hlyD*, *rfe*, *ompT*, *papG* allele I', *papG* allele I, *kpsMT* II and *papC*; 4) *gafD*, *cvaC*, *fliC* (H7), *cdtB*, *focG*, *traT* and *papG* allele II; and 5) *papG* allele 1, *papG* alleles 2 and 3, *iha*, *afa*, *iss*, *sfaS* and *kpsMT* (K1). Amplification for each of these groups of genes was performed in 25 µl reaction mixtures that included 2 µl template DNA, 12.775 µl ddH₂O, 2.5 µl 10 × PCR buffer (Invitrogen), 4 µl 25 mM MgCl₂, 0.25 µl AmpliTaq Gold *Taq* (5 U µl⁻¹) (Roche Molecular Biosystems), 0.625 µl of each 10 mM dNTP (USB) and 0.075 µl 0.1 mM upper and lower primers. The amount of ddH₂O varied according to the number of primers used in each group/panel. These reaction mixtures were subjected to the following conditions in a Mastercycler Gradient thermocycler: 12 min at 95 °C to activate the AmpliTaq Gold *Taq*, 25 cycles of 30 s at 94 °C, 30 s at 63 °C, and 3 min at 68 °C, with a final cycle of 10 min at 72 °C, followed by a hold at 4 °C.

sita, *feoB* and *irp-2* were detected with another multiplex procedure (Rodriguez-Siek *et al.*, 2005). In this case, the 25 µl reaction mixture included 2 µl template DNA, 18.9 µl ddH₂O, 1.5 µl 10 × PCR Buffer (Invitrogen), 0.75 µl 50 mM MgCl₂, 0.25 µl DNA *Taq* Polymerase (Invitrogen) 1.0 µl of a dNTP mixture with a concentration of 2.5 mM of each dNTP (USB) and 0.1 µl 0.1 mM upper and lower primers (Table 2). This mixture was subjected to 5 min at 94 °C, 30 cycles of 30 s at 94 °C, 30 s at 59 °C, and 30 s at 72 °C, with a final 5 min cycle at 72 °C, followed by a 4 °C hold.

tsh and *iucC* were identified using individual amplification procedures (Rodriguez-Siek *et al.*, 2005). For amplification of *tsh*, a 25 µl reaction mixture was used that included 15.88 µl ddH₂O, 2.5 µl 10 × PCR Buffer (Invitrogen), 2.0 µl 50 mM MgCl₂, 0.25 µl DNA *Taq* Polymerase (Invitrogen), 2.0 µl of a dNTP mixture with a concentration of 2.5 mM of each dNTP (USB) and 0.25 µl 0.1 mM upper and lower primers (Table 2). This mixture was subjected to 5 min at 95 °C, 1 min at 95 °C, 9 cycles of 30 s at 55 °C, 30 s at 72 °C, and 30 s at 94 °C, followed by 25 cycles of 30 s at 55 °C and 1 min at 72 °C, followed by 7 min at 72 °C with a final hold at 4 °C. For amplification of *iucC*, a 25 µl reaction mixture was used, containing 15.15 µl ddH₂O, 2.5 µl 10 × PCR Buffer (Invitrogen), 3.0 µl 50 mM MgCl₂, 0.15 µl DNA *Taq* Polymerase (Invitrogen) 2.0 µl of a dNTP mixture with a concentration of 2.5 mM of each dNTP (USB) and 0.25 µl 0.1 mM upper and lower primers (Table 2). This mixture was subjected to the conditions described by Skyberg *et al.*, 2003.

All samples were subjected to horizontal gel electrophoresis in 1.5 % agarose, and the size of the amplicons was determined by comparison to the Hi-Lo DNA marker (Minnesota Molecular Inc.). Positive and negative controls were examined with each amplification procedure, and all amplification procedures were repeated three times to reduce the possibility of false negatives. An isolate was considered to contain the gene of interest if it produced an amplicon of the expected size (Table 2).

Biostatistics. The null hypothesis that the relative proportions of each of the four phylogenetic groups was equal across the APEC and UPEC isolates was tested with the chi-square test of homogeneity (Snedecor & Cochran, 1980). The null hypothesis that the proportion of UPEC containing each virulence gene was equal to the proportion of APEC isolates with that gene was tested by a Z-test on the difference between the proportions (Snedecor & Cochran, 1980). In a further attempt to discern patterns among all isolates based on their content of virulence genes (*papG* allele I' was excluded, as it was absent in all isolates), multivariate statistics were used. A linear discriminant analysis (LDA) was used to determine if isolate type (APEC or UPEC) could be predicted, based on the virulence genes present (Huberty, 1994). Although use of data from binary variables in an LDA, as done here, violates the assumption of multivariate normality, LDA was used, since parametric LDA can be very robust in spite of such violations (McLachlan, 1992). Additionally, a cluster analysis of the isolates was run using the average linkage method based upon Jaccard's dissimilarity coefficient calculated from the presence of virulence genes (S Institute, Inc., 2004). In order to better discern patterns among the isolates, results of the cluster and discriminant analyses, along with the isolates' virulence genotypes, phylogenetic groups and states of origin, were used to construct a single figure based on principles of Eisen *et al.* (1998).

Additionally, all isolates were subjected to principal components analysis (PCA), based on their content of virulence genes. Essentially, a PCA allows a reduction of the dimensionality of the dataset through mathematical transformation of the data into a new set of variables [the principal components (PCs)]. As a result of this analysis, the variability of the data can be expressed in fewer dimensions than it would have otherwise (Morrison, 1976). In the present study, PCA was used to reduce 37 dimensions (one per each gene or allele studied) to two PCs. Then, the two PCs, which accounted for much of the variation in the original data, were plotted against one another using different colours to denote whether an isolate was an APEC or UPEC. Results of the cluster and discriminant analyses were also superimposed on the PCA, allowing us to better visualize areas of overlap between APEC and UPEC.

RESULTS

Serogroups among APEC and UPEC

Many of the UPEC and APEC isolates used in this study could not be classified using standard antisera: 12.5 % of UPEC and 29.4 % of APEC were non-typable. Additionally, some isolates in both groups were classified as rough, showed multiple serogroups, or autoagglutinated, preventing their assignment to a serogroup. Overall, only about 70 % of both groups was typed to a single serogroup. Among those that could be classified into a single serogroup, a high degree of diversity was seen (Table 3). Both UPEC and APEC contained isolates identified as O1, O2, O4, O6, O8, O11, O15, O18, O19, O21, O23, O25, O36, O75, O77, O82, O83, O86, O109, O112 and O117. About 86 % of the typable UPEC fell into a serogroup that was also found among the APEC of this study. That is, a large number of the UPEC overlapped in serogroup with the APEC, whereas only about 14 % of the UPEC were assigned to UPEC-unique serogroups. By contrast, among the typable APEC, only about 39 % were classified into a serogroup shared with UPEC. Most typable APEC were classified into an APEC-unique serogroup (about 62 %). APEC isolates fell into 60

Table 3. Serogroups among the APEC and UPEC isolates

Some isolates were not serogrouped because they failed to react with standard antisera, were rough, autoagglutinated, or were classified into more than one serogroup. In addition, serogrouping was not performed on three of the APEC isolates. Values show the number of isolates of each serogroup.

Serogroup	APEC (<i>n</i> =524)	UPEC (<i>n</i> =200)
Shared serogroups		
O1	9	13
O2	63	16
O4	1	10
O6	1	32
O8	12	6
O11	7	1
O15	1	4
O18	11	12
O19	7	1
O21	1	3
O23	1	1
O25	9	9
O36	7	1
O75	1	11
O77	3	1
O82	1	1
O83	2	1
O86	1	1
O109	4	1
O112	1	1
O117	1	1
Unshared serogroups		
O3	0	2
O5	4	0
O7	3	0
O9	6	0
O12	0	1
O16	0	2
O17	2	0
O20	5	0
O28	2	0
O29	2	0
O32	0	1
O35	4	0
O39	0	1
O41	1	0
O43	0	1
O51	3	0
O53	3	0
O55	1	0
O61	1	0
O64	0	1
O73	1	0
O74	1	0
O78	125	0
O85	1	0
O91	2	0
O100	3	0

Table 3. cont.

Serogroup	APEC (<i>n</i> =524)	UPEC (<i>n</i> =200)
O102	0	1
O103	3	0
O105	0	3
O106	2	0
O108	2	0
O111	14	0
O114	2	0
O115	2	0
O118	1	0
O119	6	0
O120	3	0
O123	1	0
O127	1	0
O128	1	0
O132	4	0
O134	1	0
O141	0	1
O143	1	0
O145	0	1
O149	0	1
O150	0	1
O153	3	0
O154	0	1
O157	0	1
O158	2	0
O159	1	0
O167	2	0
O171	4	0
O175	0	1

serogroups, with 39 of these unique to APEC. Among the UPEC, 37 different serogroups were found, with 17 of these unique to UPEC. Serogroups containing at least 5% or more of the typable isolates for UPEC were O6>O2>O1>O18>O75>O4, and for APEC, they were O78>O2. Among these, only the O78 group was not found in both the APEC and UPEC isolates. Interestingly, the O2 serogroup was among the most commonly occurring serogroups in both APEC and UPEC.

Lactose utilization and haemolytic reactions by APEC and UPEC

In general, most isolates, regardless of host source, were able to ferment lactose. However, 8.4% of APEC and 7.5% of UPEC did not give a positive reaction on MacConkey agar for lactose fermentation. One APEC strain haemolysed blood agar, whereas 16% of UPEC strains caused clearing of the agar around areas of bacterial growth (i.e. they were β haemolytic), and one UPEC isolate caused an α -haemolytic reaction.

Phylogenetic typing of APEC and UPEC

Some APEC and UPEC isolates were assigned to all four phylogenetic groups (Table 4). The majority of APEC fell

into group A, whereas the majority of UPEC were found in group B2. A substantial number of both APEC (29.6 %) and UPEC (18.5 %) were assigned to group D. Analysis of the results demonstrated that there is strong evidence to reject the hypothesis of homogeneity of relative proportions for each of the four phylogenetic groups across isolate source (APEC or UPEC) ($\chi^2 = 151.02$, $df = 3$, $P < 0.0001$). That is, APEC and UPEC differed significantly in their assignments to phylogenetic groups.

Virulence genotyping of APEC and UPEC

Table 5 summarizes the content of virulence genes for the two groups of isolates. Virtually all isolates contained *feoB* and *fimH* (98 % or greater). All the pTJ100-associated genes occurred in the majority of APEC. In fact, these plasmid-related genes were the most commonly occurring genes among APEC, with *iroN*, *iss*, *iucC*, *iutA*, *sitA* and *traT* occurring in 75 % or more of APEC isolates, and *cvaC* and *tsh* occurring in 66.8 % and 63.2 % of APEC, respectively. Plasmid-associated genes occurred in a substantial number of UPEC as well. *sitA* occurred in about 86 % of UPEC, and *traT* and *iss* occurred in the majority of UPEC studied. A substantial number of UPEC also contained the remaining plasmid genes (*iroN*, *iucC*, *iutA* and *tsh*), except for *cvaC*, which was found in only about 7.5 % of the UPEC isolates.

Also, other genes not linked to pTJ100 were frequently found in both APEC and UPEC isolates. Genes of the yersinabactin operon (*fyuA* and *irp-2*) and *ompT*, which encodes a protease (Cavard & Lazdunski, 1990), commonly occurred in both UPEC and APEC. *kpsMT* II, involved in the synthesis of group II capsules, and a marker of a PAI from UPEC strain CFT073 occurred commonly in UPEC (~75 %). These did not occur as commonly in APEC isolates, but were found in more than 15 % of the isolates tested. *papA*, *C*, *E* and *F* occurred in about half of the UPEC tested, and they contained various *papG* alleles. About a third of the UPEC had *papG* allele I and about a quarter had

papG alleles II and/or III. Interestingly, few APEC were identified as containing *papA* (7.4 %), but almost 40 % had *papC*, *E*, *F* and *G* allele II. Other genes of interest that occurred in 15–49 % of the APEC isolates included *ireA* and *kpsMT* (K1); 15–49 % of the UPEC isolates contained *iha*, *sfa*, *sfaS*, *ireA*, *kpsMT* (K1), *cnf-1*, *hlyD*, *fliC* (H7) and *ibeA*. Although the differences in distributions of *bmaE*, *gafD*, *afa*, *fimH*, *feoB* and *sitA* between the two groups were not significantly different, the two groups did vary significantly, often to a great degree, in the distribution of the other genes tested (Table 5). Few isolates of either group produced amplicons with the *bmaE* and *gafD* primers, and none produced an amplicon corresponding to *papG* allele I'. Also of interest, six of the APEC isolates had virulence genotypes that were identical to one or more of the UPEC isolates (Table 6).

Many of the isolates' identities as APEC or UPEC were predictable using discriminant analysis, which was based on virulence genotypes. However, several isolates' identities were incorrectly predicted by discriminant analysis. Twenty of the 524 APEC were misclassified as UPEC, while 29 of 200 UPEC were misclassified as APEC (Fig. 1). These errors in prediction of group membership did not appear to be related to phylogenetic type. Of the 20 APEC isolates that were incorrectly predicted to be UPEC, eight fell into phylogenetic group A (3.6 %), two into B1 (2.2 %), five into B2 (2.2 %) and five into D (2.7 %). Of the 29 UPEC isolates that were predicted to be APEC, six were assigned to group A (2.7 %), one to B1 (1.1 %), 13 to B2 (5.7 %) and nine to D (4.9 %). However, misclassification was more likely when counts of virulence genes were low.

In addition, all *E. coli* were subjected to cluster analysis based on their virulence genotypes. Many clusters and subclusters were identified (Fig. 1). For ease of description, clusters were sequentially numbered from 1 to 6, beginning with cluster 1, which is found at the top of Fig. 1. Cluster 1, a heavily mixed cluster of APEC and UPEC (74 APEC isolates and 44 UPEC isolates), contained many of the isolates that were misclassified by discriminant analysis. Isolates of this cluster contained fewer virulence genes than did isolates of other clusters, and members of this cluster could be found in all four phylogenetic groups. Cluster 2 contained the bulk of the APEC isolates (363 of 524 APEC), some UPEC (7), and several misclassified isolates. Isolates found in this cluster typically contained a full complement of pTJ100-related genes. This cluster could be roughly divided into two subclusters of isolates based on possession of the *pap* operon (*papC*, *E*, *F* and *G* allele II, but not A) and *ireA*. Isolates in cluster 2 were assigned to a variety of phylogenetic groups, including those not typically associated with pathogenicity. In general, cluster 3 was composed of APEC and UPEC (45 UPEC and 63 APEC isolates) containing several pTJ100-related genes, but lacking genes of the *pap* operon. Isolates of this cluster also tended to contain capsular genes, *malX* (PAI) and *ibeA*, and to be classified into phylogenetic groups, B2 or D, which are typically linked to pathogenicity.

Table 4. Assignment of isolates to phylogenetic groups

Analysis of the results demonstrated that there is strong evidence to reject the hypothesis of homogeneity of relative proportions for each of the four phylogenetic groups across isolate source (APEC or UPEC) ($\chi^2 = 151.02$, $df = 3$, $P < 0.0001$). That is, APEC and UPEC differed significantly in their assignments to phylogenetic groups. Values in parentheses show the proportion of isolates in each group as a percentage of the total.

	Number of isolates in group:				Total
	A	B1	B2	D	
APEC	199 (38.0)	81 (15.5)	97 (18.5)	147 (28.1)	524
UPEC	21 (10.5)	12 (6.0)	130 (65.0)	37 (18.5)	200
Total	220	93	227	184	724

Table 5. Prevalence of virulence genes in APEC and UPEC

Values show the number of isolates; values in parentheses show the proportion of isolates containing each gene as a proportion of the total number of isolates.

Gene	APEC (<i>n</i> = 524)		UPEC (<i>n</i> = 200)		Z-number	<i>P</i> value
	Number of isolates	(Percentage of APEC)	Number of isolates	(Percentage of UPEC)		
pTJ100-related						
<i>cvaC</i> *	350	(66·8)	15	(7·5)	14·764	< 0·0001
<i>iroN</i> †	454	(86·6)	80	(40·0)	11·644	< 0·0001
<i>iss</i> *	427	(81·5)	121	(60·5)	5·565	< 0·0001
<i>iucC</i> †	407	(77·7)	78	(39·0)	9·438	< 0·0001
<i>iutA</i> †	420	(80·2)	67	(33·5)	11·332	< 0·0001
<i>sitA</i> †	453	(86·5)	171	(85·5)	0·329	0·7419
<i>traT</i> *	407	(77·7)	141	(70·5)	1·969	0·0489
<i>tsh</i> ‡	331	(63·2)	79	(39·5)	5·697	< 0·0001
Adhesins						
<i>afa</i>	42	(8·0)	20	(10·0)	−0·834	0·4042
<i>bmaE</i>	2	(0·4)	4	(2·0)	−1·795	0·0727
<i>fimH</i>	514	(98·1)	198	(99·0)	−0·913	0·3613
<i>focG</i>	0	(0·0)	25	(12·5)	−6·213	< 0·0001
<i>gafD</i>	2	(0·4)	0	(0·0)	1·052	0·2927
<i>iha</i>	15	(2·9)	53	(26·5)	−8·035	< 0·0001
<i>papA</i>	39	(7·4)	99	(49·5)	−11·213	< 0·0001
<i>papC</i>	203	(38·7)	103	(51·5)	−3·085	0·0020
<i>papEF</i>	197	(37·6)	102	(51·3)	−3·247	0·0012
<i>papG</i> allele I	2	(0·4)	64	(32·0)	−10·327	< 0·0001
<i>papG</i> allele I'	0	(0·0)	0	(0·0)	—	—
<i>papG</i> allele II	205	(39·1)	48	(24·0)	3·915	< 0·0001
<i>papG</i> allele III	3	(0·6)	51	(25·5)	−8·907	< 0·0001
<i>sfa</i>	22	(4·2)	63	(31·5)	−8·578	< 0·0001
<i>sfaS</i>	19	(3·6)	52	(26·0)	−7·578	< 0·0001
Iron-related						
<i>feoB</i>	521	(99·4)	198	(99·0)	0·582	0·5604
<i>fyuA</i>	307	(58·6)	160	(80·0)	−5·585	< 0·0001
<i>ireA</i>	242	(46·2)	48	(24·0)	5·592	< 0·0001
<i>irp-2</i>	301	(57·4)	167	(83·5)	−6·873	< 0·0001
Protectins						
<i>kpsMT</i> (K1)	83	(15·8)	61	(30·7)	−4·181	< 0·0001
<i>kpsMT</i> II	130	(24·8)	155	(77·5)	−12·682	< 0·0001
<i>kpsMT</i> III	8	(1·5)	13	(6·5)	−3·049	0·0023
<i>rfc</i>	3	(0·6)	16	(8·0)	−4·412	< 0·0001
Toxins						
<i>cdtB</i>	6	(1·1)	16	(8·0)	−3·948	< 0·0001
<i>cnf-1</i>	6	(1·1)	55	(27·5)	−9·052	< 0·0001
<i>hlyD</i>	4	(0·8)	62	(31·0)	−9·953	< 0·0001
Miscellaneous						
<i>fliC</i> (H7)	25	(4·8)	40	(20·0)	−5·562	< 0·0001
<i>ibeA</i>	77	(14·7)	60	(30·0)	−4·420	< 0·0001
<i>ompT</i>	367	(70·0)	167	(83·5)	−3·835	0·0001
<i>malX</i> (PAI)	87	(16·6)	149	(74·5)	−13·987	< 0·0001

*These genes are listed as pTJ100-related, but they could also be listed with the protectins.

†These genes are listed as pTJ100-related, but they could also be listed with the iron-related genes.

‡These genes are listed as pTJ100-related, but they could also be listed in the miscellaneous group.

Table 6. APEC and UPEC isolates with identical virulence genotypes

+, Contains the gene; -, lacks the gene; M, reacted to an experimental antiserum; w, produced a weak reaction to the antiserum.

Type	Serogroup	Phylo- genetic group	<i>cvaC</i>	<i>iron</i>	<i>iss</i>	<i>tucC</i>	<i>tutA</i>	<i>sita</i>	<i>traT</i>	<i>tsh</i>	<i>afu</i>	<i>bmaE</i>	<i>fimH</i>	<i>focG</i>	<i>gadD</i>	<i>ihfA</i>	<i>papA</i>	<i>papC</i>	<i>papEF</i>	<i>papG</i> allele I	<i>papG</i> allele I'	<i>papG</i> allele II	<i>papG</i> allele III	<i>sfa</i>	<i>sfas</i>	<i>feoB</i>	<i>fyuA</i>	<i>ireA</i>	<i>tvp-2</i>	<i>kpsMT</i> (K1)	<i>kpsMT</i> II	<i>kpsMT</i> III	<i>rjc</i>	<i>cdtB</i>	<i>cnf-1</i>	<i>hlyD</i>	<i>fliC</i> (H7)	<i>ibcA</i>	<i>ompT</i>	<i>malX</i> (PAI)	
UPEC	1	B2	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
UPEC	M	B2	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
APEC	2	D	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
UPEC	6	B2	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
APEC	2	B2	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
UPEC	-	D	-	-	-	+	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
APEC	X9	A	-	-	-	+	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
UPEC	109w	A	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
APEC	86	B1	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
UPEC	8	D	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
APEC	-	B1	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
UPEC	25	D	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
APEC	-	B2	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

Cluster 4 was another mixed cluster (37 UPEC and 20 APEC) containing some misclassified isolates. It was dominated by UPEC containing the *pap* genes (*papA*, C, E, F and G allele II) and *malX* (PAI). Additionally, several of these isolates had genes for capsular antigen. Most isolates in this cluster were assigned to phylogenetic groups B2 and D. Cluster 5 was composed of a single APEC isolate. Cluster 6 was heavily dominated by UPEC (67 UPEC and 3 APEC). Typical members of this cluster contained some plasmid genes, *malX* (PAI), and genes encoding adhesins (*pap* and/or others), capsule, and a toxin. Most were classified into phylogenetic group B2. In general, isolates of cluster 6 contained more virulence genes than other isolates.

The data obtained on virulence genotypes of APEC and UPEC isolates were also subjected to PCA using SAS in an effort to provide a visual assessment of the degree of separation among the clusters of isolates. The two PCs were then plotted against one another (Fig. 2), and the results of the cluster and discriminant analyses were superimposed on this plot. Again, overlap among APEC and UPEC isolates was apparent in all clusters, but cluster 6 isolates, which were predominately UPEC, were found to be quite distinct, as indicated by their distance from the other isolates on this plot. Notice that the greatest numbers of misclassified isolates occurred in the lower left-hand portion of the PCA plot and were primarily from the uppermost mixed cluster of isolates that are typically characterized by low numbers of virulence genes.

Origin of APEC and UPEC isolates

Although UPEC isolates originated from a single state, they exhibited great diversity in their complement of virulence genes and assignments to cluster and phylogenetic groups (Fig. 1). By contrast, the majority of APEC isolates were found in a single cluster, despite the fact that they originated from at least nine states (Fig. 1).

DISCUSSION

Although it is widely accepted that the source of *E. coli* causing most UTIs is the colonic flora of affected individuals (Salysers & Whitt, 1994), there is no consensus as to the source of the urovirulent clones inhabiting the colon. Although some UTIs might involve sexual transmission (Foxman *et al.*, 1997; Johnson & Delavari, 2002) or exposure to the excreta of dogs with UTIs (Johnson *et al.*, 2000), the major source of these urovirulent clones is unknown. A plausible route of transmission for any bacterium colonizing the gut is a fecal-oral one. If the fecal-oral route is a possible means by which urovirulent *E. coli* come to inhabit the human intestine, it would seem important to determine what food might serve as their vehicle. Poultry is a candidate vehicle based on the findings of others (Levy *et al.*, 1976; Linton *et al.*, 1977; Ojeniyi, 1989; van den Bogaard *et al.*, 2002; Johnson *et al.*, 2003) that show transmission of avian *E. coli* from poultry to humans or similarities between avian *E. coli* and UPEC.

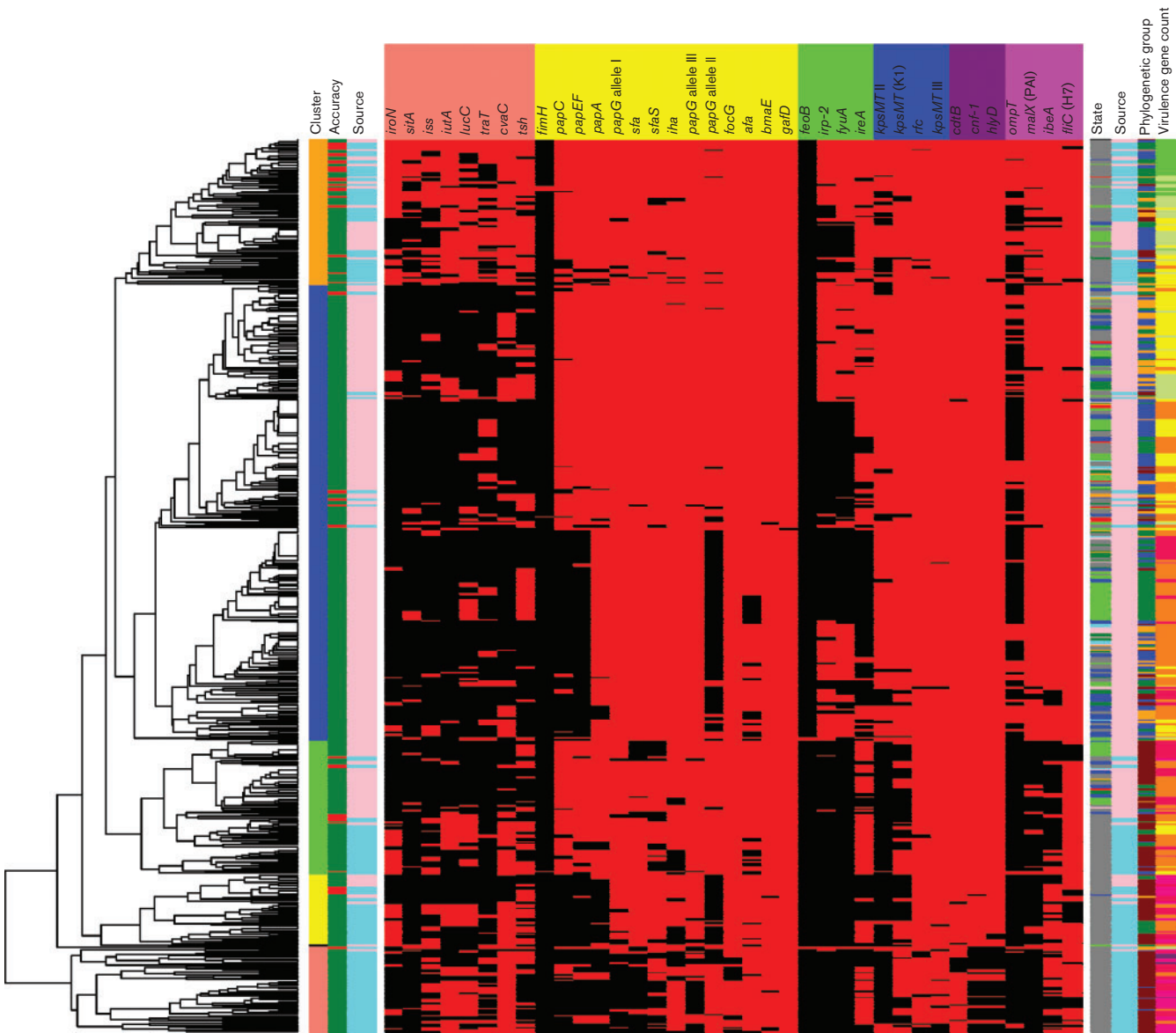


Fig. 1. Results of cluster and discriminant analyses, based on virulence genotype, plotted with each isolate's content of virulence genes, phylogenetic group, and state of origin. The left-most portion of this figure is the dendrogram resulting from the cluster analysis. Just to the right of the dendrogram is Column 1, which shows cluster membership. Orange, cluster 1; blue, cluster 2; lime, cluster 3; yellow, cluster 4; black, cluster 5 (consisting of a single APEC); salmon, cluster 6 (remaining clusters combined). Column 2 shows the accuracy of the prediction from the discriminant analysis. Correctly classified isolates are represented with the green bands, whereas incorrectly classified isolates are shown as red bands. Column 3 identifies an isolate as an APEC (pink) or UPEC (cyan). Columns 4–40 show the virulence genotype of each isolate tested. Each column in this group shows the results for a single gene. Black, gene present; red, gene absent. Column 41 shows the state of origin of the isolates: red, Alabama; green, Arkansas; blue, California; orange, Georgia; cyan, Indiana; pink, Minnesota; lime, Nebraska; purple, New York; black, North Dakota; grey, unknown. All UPEC originated from a single state and are also shown as grey. Column 42 again shows the APEC (pink) or UPEC (cyan) group membership. Column 43 shows the phylogenetic group of each isolate: blue, A; orange, B1; brown, B2; green, D. Column 44 shows the relative count of virulence genes in each isolate using the following colour code: dark green, 2–4 virulence genes; light green, 5–8; yellow, 9–12; orange, 13–16; red, 17–20; hot pink, 21–24; purple, 25–28.

Further, the data presented in the present study reveal that UPEC and APEC have similarities in their serogroups, virulence genotypes and assignments to phylogenetic

groups, supporting the hypothesis that poultry may be a vehicle for *E. coli* capable of causing human urinary tract disease. For instance, the most commonly occurring

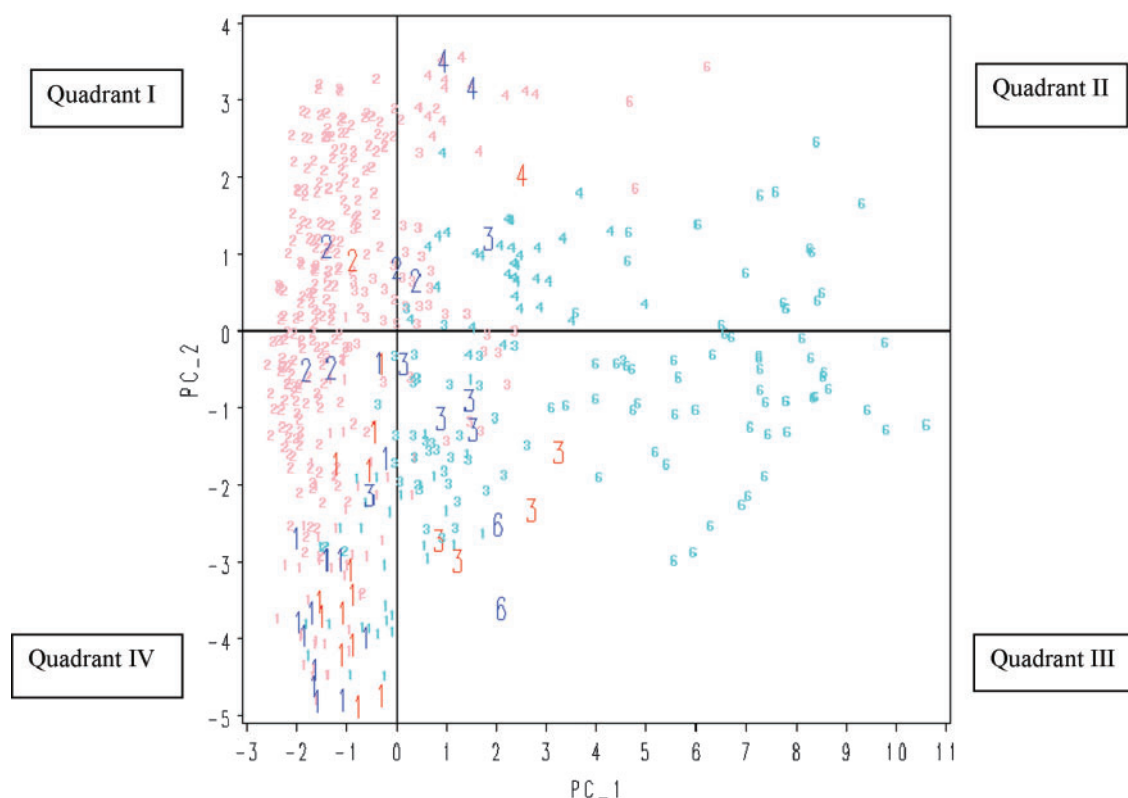


Fig. 2. PCA of the 724 APEC and UPEC isolates, based on virulence genotype, plotted with results of cluster and discriminant analyses. Each number in the plot represents an individual isolate. All the '1s' are members of cluster 1 (corresponding to the orange cluster from Fig. 1), all the '2s' are members of cluster 2 (corresponding to the blue cluster in Fig. 1), etc. Pink and red numbers identify APEC isolates; cyan and blue numbers identify UPEC. Large numbers indicate isolates that were misclassified as APEC or UPEC in the discriminant analysis.

serogroups among the UPEC were all shared with the APEC. The O2 serogroup, which was one of the more commonly occurring serogroups among the UPEC, was also among the more commonly occurring serogroups of the APEC. In fact, the O2 serogroup is considered to be among the three most commonly occurring serogroups of APEC (Sojka & Carnaghan, 1961). Also, an interesting finding of the present study was that the O11 serogroup occurred among the APEC. Manges *et al.* (2001) documented the emergence of UPEC clonal group A, a previously unrecognized clone of multidrug-resistant UPEC, as a cause of UTIs in women in three different states. That these UPEC emerged in three states and showed a high degree of genetic homogeneity, at least in the California isolates, suggested to the authors that these UPEC might have originated from contaminated food. In their study, two serogroups predominated among the clonal group A isolates from California, one of which was the O11 serogroup that was found among the APEC of the present study. It would be interesting to compare the genotypes of these California UPEC isolates to the APEC of the same serogroup to see if these isolates share other similarities.

In addition, results of the gene prevalence studies revealed some similarities between the two groups of isolates. Several APEC and UPEC strains shared identical genotypes for the virulence genes assayed (Table 6). In fact, when the isolates of this study were subjected to cluster analysis and PCA, and the results plotted (Figs 1 and 2), overlap between the UPEC and APEC isolates was detectable. So too, assignment of isolates to APEC and UPEC groups, based on a discriminant analysis of the virulence genotypes, resulted in misclassification of some isolates, especially when counts of virulence genes were low. A substantial number of both groups contained genes involved in the formation of different adhesins, including P (various *pap* genes and alleles) and type 1 fimbriae. APEC and UPEC isolates also contained many genes related to iron acquisition, including *fyuA* and *irp-2*, genes of the yersiniabactin operon (Oelschlaeger *et al.*, 2002a; Pelludat *et al.*, 1998; Schubert *et al.*, 1998, 2000, 2002); *iucC* and *iutA*, genes of the aerobactin operon (de Lorenzo *et al.*, 1986; de Lorenzo & Neilands, 1986); *iroN* of the salmochelin operon (Russo *et al.*, 2002; Hantke *et al.*, 2003); and *sitA* of the *sit* operon (Runyen-Janecky *et al.*, 2003). They also had a high prevalence of genes related to

complement resistance, such as *iss* (Binns *et al.*, 1979, 1982; Pfaff-McDonough *et al.*, 2000; Nolan *et al.*, 2003) and *traT* (Moll *et al.*, 1980; Binns *et al.*, 1982; Pfaff-McDonough *et al.*, 2000).

Some of the genes that were found to be widely distributed among both UPEC and APEC are known for their contribution to APEC virulence and have been localized to large, transmissible R plasmids, such as pTJ100 (Johnson *et al.*, 2002, 2004). Similar plasmids have been identified in UPEC (Sorsa *et al.*, 2003), and several of these pTJ100-associated genes have been found on the UPEC chromosome in PAIs (Dobrindt *et al.*, 2001, 2002; Oelschlaeger *et al.*, 2002a,b). In the present study, APEC and UPEC isolates were assessed for their possession of several of these pTJ100-related genes. All of the plasmid-related genes studied were found in over 60 % of the APEC examined (Table 5), suggesting that plasmids such as pTJ100 are widely distributed among APEC, and that APEC might be a reservoir of plasmid-mediated virulence genes transmissible to other bacteria. Although there were significant differences in distribution of the pTJ100-related genes between APEC and UPEC, except for *sitA*, 60 % or more of the UPEC contained *sitA*, *iss* and *traT*, and a third of the UPEC contained all the plasmid genes assessed except for *cvaC*. Unlike the situation that exists on pTJ100, where *iroN*, *iucC*, *iss*, *traT*, *tsh* and *sitA* are found in conjunction with *cvaC*, these genes in UPEC are rarely associated with *cvaC*. A possible explanation for this discrepancy might lie in the location of these genes in UPEC and APEC. In UPEC, these genes may be more likely found on non-ColV plasmids or within chromosomal PAIs (Dobrindt *et al.*, 2001, 2002; Oelschlaeger *et al.*, 2002a,b; Sorsa *et al.*, 2003) than on pTJ100-like plasmids. Further study will be needed to ascertain the location of these genes in UPEC.

There was also discernable overlap in APEC and UPEC isolates in their assignments to phylogenetic group, with some intriguing differences. As predicted by the literature, a majority of the UPEC fell into group B2 (Clermont *et al.*, 2000). In contrast, less than 20 % of the APEC were assigned to this group. However, a substantial number of both fell into group D, which is also known for its content of virulent ExPEC (Clermont *et al.*, 2000). Interestingly, about half of the APEC and 10.5 % of the UPEC were identified as belonging to group A, a group thought to be composed of commensal *E. coli* (Clermont *et al.*, 2000). Perhaps the isolates in group A are commensals, and their association with disease reflects the opportunistic nature of the infections from which they were obtained.

There were also notable differences between the two groups of isolates. For example, the O78 serogroup, which occurred frequently among the APEC of this study and is considered one of the three most commonly occurring serogroups in APEC (Sojka & Carnaghan, 1961), was not found among the UPEC studied here. Only one APEC strain was haemolytic, whereas 16 % of UPEC were. Also, several differences in the prevalence of the virulence genes across the two

groups of isolates were found. For instance, few APEC had the *cnf-1* gene, but about 28 % of UPEC contained it. Although a substantial number of isolates from both groups contained genes associated with pTJ100, these genes were significantly more likely to occur in APEC than they were in UPEC, except for *sitA*. Other differences in gene prevalence were also detected between these two groups of isolates (Table 5). These differences suggest that all the UPEC studied are likely not derived from APEC, but the similarities among several of the isolates suggest that it is possible that some are. Also, interpretation of these results must be tempered by the fact that a limited sample of UPEC, all isolated from the same geographic location, was examined in this study. A larger study, using alternate typing methods and UPEC isolates from widely different areas and types of UTIs, might prove beneficial to the overall interpretation of the data. However, a bias due to their origin from a single state was not readily evident, as the UPEC tested were quite diverse in terms of their serogroups, genotypes and phylogenetic groups. In fact, the UPEC, which all originated from North Dakota, were well represented in several clusters, whereas the bulk of the APEC, originating from multiple states, was found in a single cluster.

In sum, APEC and UPEC cause widely prevalent extraintestinal diseases. Their propensity for causing such diseases relates to their possession of many virulence-associated traits, including certain serogroups, adhesins, iron acquisition systems, toxins, protectins and invasins, that enable them to grow and cause disease in these host environments. Consequently, it seems reasonable that UPEC and APEC might show similar adaptations for an extraintestinal lifestyle, which, in turn, might enable APEC to cause extraintestinal disease in human beings. Further research will be necessary to determine if APEC can cause human UTIs or serve as a reservoir of virulence genes contributing to uropathogenesis.

ACKNOWLEDGEMENTS

We thank Dr Paul Carson (MeritCare Hospital, Fargo, ND) for the UPEC strains used in this study, and Dr James R. Johnson (VA Medical Center, Minneapolis, MN) for his advice and kind gift of many of the UPEC control strains and protocols used in the genotyping studies. Without their assistance, this study would not have been done.

REFERENCES

- Achtman, M., Kennedy, N. & Skurray, R. (1977). Cell-cell interactions in conjugating *Escherichia coli*: role of *traT* protein in surface exclusion. *Proc Natl Acad Sci U S A* **74**, 5104–5108.
- Barnes, H. J., Vaillancourt, J. P. & Gross, W. B. (2003). Colibacillosis. In *Diseases in Poultry*, pp. 631–652. Edited by Y. M. Saif. Ames, IA: Iowa State University Press.
- Binns, M. M., Davies, D. L. & Hardy, K. G. (1979). Cloned fragments of the plasmid ColV, I-K94 specifying virulence and serum resistance. *Nature* **279**, 778–781.

- Binns, M. M., Mayden, J. & Levine, R. P. (1982). Further characterization of complement resistance conferred on *Escherichia coli* by the plasmid genes *traT* of R100 and *iss* of ColV, I-K94. *Infect Immun* 35, 654–659.
- Bonacorsi, S. P. P., Clermont, O., Tinsley, C., Le Gall, I., Beaudoin, J. C., Elion, J., Nassif, X. & Bingen, E. (2000). Identification of regions of the *Escherichia coli* chromosome specific for neonatal meningitis-associated strains. *Infect Immun* 68, 2096–2101.
- Cavard, D. & Lazdunski, C. (1990). Colicin cleavage by OmpT protease during both entry into and release from *Escherichia coli* cells. *J Bacteriol* 172, 648–652.
- Chuba, P. J., Leon, M. A., Banerjee, A. & Palchaudhuri, S. (1989). Cloning and DNA sequence of plasmid determinant *iss*, coding for increased serum survival and surface exclusion, which has homology with lambda DNA. *Mol Gen Genet* 216, 287–292.
- Clermont, O., Bonacorsi, S. & Bingen, E. (2000). Rapid and simple determination of the *Escherichia coli* phylogenetic group. *Appl Environ Microbiol* 66, 4555–4558.
- de Lorenzo, V. & Neilands, J. B. (1986). Characterization of *iucA* and *iucC* genes of the aerobactin system of plasmid ColV-K30 in *Escherichia coli*. *J Bacteriol* 167, 350–355.
- de Lorenzo, V., Bindereif, A., Paw, B. H. & Neilands, J. B. (1986). Aerobactin biosynthesis and transport genes of plasmid ColV-K30 in *Escherichia coli* K-12. *J Bacteriol* 165, 570–578.
- Dobrindt, U., Blum-Oehler, G., Hartsch, T., Gottschalk, G., Ron, E. Z., Funfstuck, R. & Hacker, J. (2001). S-Fimbria-encoding determinant *sfa(I)* is located on pathogenicity island III(536) of uropathogenic *Escherichia coli* strain 536. *Infect Immun* 69, 4248–4256.
- Dobrindt, U., Blum-Oehler, G., Nagy, G., Schneider, G., Johann, A., Gottschalk, G. & Hacker, J. (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.
- Donnenberg, M. S. & Welch, R. (1996). Virulence determinants of uropathogenic *Escherichia coli*. In *Urinary Tract Infections: Molecular Pathogenesis and Clinical Management*, pp. 135–174. Edited by H. L. T. Mobley & J. W. Warren. Washington, DC: American Society for Microbiology.
- Eisen, M. B., Spellman, P. T., Brown, P. O. & Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* 95, 14863–14868.
- Fields, P. I., Blom, K., Hughes, H. J., Helsel, L. O., Feng, P. & Swaminathan, B. (1997). Molecular characterization of the gene encoding H antigen in *Escherichia coli* and development of a PCR-restriction fragment length polymorphism test for identification of *E. coli* O157:H7 and O157:NM. *J Clin Microbiol* 35, 1066–1070.
- Forbes, B. A., Sahm, D. F. & Weissfeld, A. S. (1998). *Bailey & Scott's Diagnostic Microbiology*. St Louis, MO: Mosby.
- Foxman, B. (2002). Epidemiology of urinary tract infections: incidence, morbidity, and economic costs. *Am J Med* 113 Suppl 1A, 5S–13S.
- Foxman, B., Zhang, L., Tallman, P. & 9 other authors (1997). Transmission of uropathogens between sex partners. *J Infect Dis* 175, 989–992.
- Gilson, L., Mahanty, H. K. & Kolter, R. (1987). Four plasmid genes are required for colicin V synthesis, export, and immunity. *J Bacteriol* 169, 2466–2470.
- Hacker, J., Goebel, W., Hof, H., Konig, W., Konig, B., Scheffer, J., Hughes, C. & Marre, R. (1988). Adhesins, serum resistance and cytotoxins of *E. coli* – genetic structure and role in pathogenicity. *Cell Biol* 24, 221–229.
- Hantke, K., Nicholson, G., Rabsch, W. & Winkelmann, G. (2003). Salmochelins, siderophores of *Salmonella enterica* and uropathogenic *Escherichia coli* strains, are recognized by the outer membrane receptor IroN. *Proc Natl Acad Sci U S A* 100, 3677–3682.
- Huberty, C. J. (1994). *Applied Discriminant Analysis*. New York: Wiley.
- Janben, T., Schwarz, C., Preikschat, P., Voss, M., Phillipp, H. C. & Wieler, L. H. (2001). Virulence associated genes in avian pathogenic *Escherichia coli* (APEC) isolated from internal organs of poultry having died from colibacillosis. *Int J Med Microbiol* 291, 371–378.
- Johnson, J. R. (1991). Virulence factors in *Escherichia coli* urinary tract infection. *Clin Microbiol Rev* 4, 80–128.
- Johnson, J. R. & Brown, J. J. (1996). A novel multiply primed polymerase chain reaction assay for identification of variant *papG* genes encoding the Gal(alpha 1-4)Gal-binding PapG adhesins of *Escherichia coli*. *J Infect Dis* 173, 920–926.
- Johnson, J. R. & Delavari, P. (2002). Concurrent fecal colonization with extraintestinal pathogenic *Escherichia coli* in a homosexual man with recurrent urinary tract infection and in his male sex partner. *Clin Infect Dis* 35, E65–E68.
- Johnson, J. R. & Stell, A. L. (2000). Extended virulence genotypes of *Escherichia coli* strains from patients with urosepsis in relation to phylogeny and host compromise. *J Infect Dis* 181, 261–272.
- Johnson, J. R., O'Bryan, T. T., Low, D. A., Ling, G., Delavari, P., Fasching, C., Russo, T. A., Carlino, U. & Stell, A. L. (2000). Evidence of commonality between canine and human extraintestinal pathogenic *Escherichia coli* strains that express *papG* allele III. *Infect Immun* 68, 3327–3336.
- Johnson, J. R., Delavari, P., Kuskowski, M. & Stell, A. L. (2001). Phylogenetic distribution of extraintestinal virulence-associated traits in *Escherichia coli*. *J Infect Dis* 183, 78–88.
- Johnson, T. J., Giddings, C. W., Horne, S. M. & 8 other authors (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.
- Johnson, J. R., Murray, A. C., Gajewski, A., Sullivan, M., Snippes, P., Kuskowski, M. A. & Smith, K. E. (2003). Isolation and molecular characterization of nalidixic acid-resistant extraintestinal pathogenic *Escherichia coli* from retail chicken products. *Antimicrob Agents Chemother* 47, 2161–2168.
- Johnson, T. J., Skyberg, J. & Nolan, L. K. (2004). Multiple antimicrobial resistance region of a putative virulence plasmid from an *Escherichia coli* isolate incriminated in avian colibacillosis. *Avian Dis* 48, 351–360.
- Levy, S. B., FitzGerald, G. B. & Macone, A. B. (1976). Spread of antibiotic-resistant plasmids from chicken to chicken and from chicken to man. *Nature* 260, 40–42.
- Linton, A. H., Howe, K., Bennett, P. M., Richmond, M. H. & Whiteside, E. J. (1977). The colonization of the human gut by antibiotic resistant *Escherichia coli* from chickens. *J Appl Bacteriol* 43, 465–469.
- Manges, A. R., Johnson, J. R., Foxman, B., O'Bryan T. T., Fullerton, K. E. & Riley, L. W. (2001). Widespread distribution of urinary tract infections caused by a multidrug-resistant *Escherichia coli* clonal group. *N Engl J Med* 345, 1007–1013.
- Maurer, J. J., Brown, T. P., Steffens, W. L. & Thayer, S. G. (1998). The occurrence of ambient temperature-regulated adhesins, curli, and the temperature-sensitive hemagglutinin *tsh* among avian *Escherichia coli*. *Avian Dis* 42, 106–118.
- McLachlan, G. J. (1992). *Discriminant Analysis and Statistical Pattern Recognition*. New York: Wiley.

- Moll, A., Manning, P. A. & Timmis, K. N. (1980). Plasmid-determined resistance to serum bactericidal activity: a major outer membrane protein, the *traT* gene product, is responsible for plasmid-specified serum resistance in *Escherichia coli*. *Infect Immun* **28**, 359–367.
- Morrison, D. F. (1976). *Multivariate Statistical Methods*, 2nd edn. New York: McGraw-Hill.
- Nolan, L. K., Horne, S. M., Giddings, C. W., Foley, S. L., Johnson, T. J., Lynne, A. M. & Skyberg, J. (2003). Resistance to serum complement, *iss*, and virulence of avian *Escherichia coli*. *Vet Res Comm* **27**, 101–110.
- Oelschlaeger, T. A., Dobrindt, U. & Hacker, J. (2002a). Virulence factors of uropathogens. *Curr Opin Urol* **12**, 33–38.
- Oelschlaeger, T. A., Dobrindt, U. & Hacker, J. (2002b). Pathogenicity islands of uropathogenic *E. coli* and the evolution of virulence. *Int J Antimicrob Agents* **19**, 517–521.
- Ojeniyi, A. A. (1989). Direct transmission of *Escherichia coli* from poultry to humans. *Epidemiol Infect* **103**, 513–522.
- Pelludat, C., Rakin, A., Jacobi, C. A., Schubert, S. & Heesemann, J. (1998). The yersiniabactin biosynthetic gene cluster of *Yersinia enterocolitica*: organization and siderophore-dependent regulation. *J Bacteriol* **180**, 538–546.
- Pfaff-McDonough, S. J., Horne, S. M., Giddings, C. W., Ebert, J. O., Doetkott, C., Smith, M. H. & Nolan, L. K. (2000). Complement resistance-related traits among *Escherichia coli* isolates from apparently healthy birds and birds with colibacillosis. *Avian Dis* **44**, 23–33.
- Provence, D. L. & Curtiss, R., III (1994). Isolation and characterization of a gene involved in hemagglutination by an avian pathogenic *Escherichia coli* strain. *Infect Immun* **62**, 1369–1380.
- Rodriguez-Siek, K. E., Giddings, C. W., Doetkott, C., Johnson, T. J. & Nolan, L. K. (2005). Characterizing the APEC pathotype. *Vet Res* **36**, 241–256.
- Runyen-Janecky, L. J., Reeves, S. A., Gonzales, E. G. & Payne, S. M. (2003). Contribution of the *Shigella flexneri* Sit, Iuc, and Feo iron acquisition systems to iron acquisition *in vitro* and in cultured cells. *Infect Immun* **71**, 1919–1928.
- Russo, T. A. & Johnson, J. R. (2000). Proposal for a new inclusive designation for extraintestinal pathogenic isolates of *Escherichia coli*: ExPEC. *J Infect Dis* **181**, 1753–1754.
- Russo, T. A. & Johnson, J. R. (2003). Medical and economic impact of extraintestinal infections due to *Escherichia coli*: focus on an increasingly important endemic problem. *Microbes Infect* **5**, 449–456.
- Russo, T. A., Carlino, U. B., Mong, A. & Jodush, S. T. (1999). Identification of genes in an extraintestinal isolate of *Escherichia coli* with increased expression after exposure to human urine. *Infect Immun* **67**, 5306–5314.
- Russo, T. A., Carlino, U. B. & Johnson, J. R. (2001). Identification of a new iron-regulated virulence gene, *ireA*, in an extraintestinal pathogenic isolate of *Escherichia coli*. *Infect Immun* **69**, 6209–6216.
- Russo, T. A., McFadden, C. D., Carlino-MacDonald, U. B., Beanan, J. M., Barnard, T. J. & Johnson, J. R. (2002). IroN functions as a siderophore receptor and is a urovirulence factor in an extraintestinal pathogenic isolate of *Escherichia coli*. *Infect Immun* **70**, 7156–7160.
- Salyers, A. A. & Whitt, D. D. (1994). *Escherichia coli* urinary tract infections. In *Bacterial Pathogenesis, a Molecular Approach*, p. 207. Washington, DC: American Society for Microbiology.
- Sanderson, K. E. & Zeigler, D. R. (1991). Storing, shipping, and maintaining records on bacterial strains. *Methods Enzymol* **204**, 248–264.
- SAS Institute, Inc. (2004). *SAS/STAT 9.1 User's Guide*. Cary, NC: SAS Institute, Inc.
- Schubert, S., Rakin, A., Karch, H., Carniel, E. & Heesemann, J. (1998). Prevalence of the 'high-pathogenicity island' of *Yersinia* species among *Escherichia coli* strains that are pathogenic to humans. *Infect Immun* **66**, 480–485.
- Schubert, S., Cuenca, S., Fischer, D. & Heesemann, J. (2000). High-pathogenicity island of *Yersinia pestis* in Enterobacteriaceae isolated from blood cultures and urine samples: prevalence and functional expression. *J Infect Dis* **182**, 1268–1271.
- Schubert, S., Picard, B., Gouriou, S., Heesemann, J. & Denamur, E. (2002). *Yersinia* high-pathogenicity island contributes to virulence in *Escherichia coli* causing extraintestinal infections. *Infect Immun* **70**, 5335–5337.
- Skyberg, J. A., Horne, S. M., Giddings, C. W., Wooley, R. E., Gibbs, P. S. & Nolan, L. K. (2003). Characterizing avian *Escherichia coli* isolates with multiplex polymerase chain reaction. *Avian Dis* **47**, 1441–1447.
- Snedecor, G. W. & Cochran, W. G. (1980). *Statistical Methods*. Ames, IA: Iowa State University Press.
- Sojka, W. J. & Carnaghan, R. B. A. (1961). *Escherichia coli* infection in poultry. *Res Vet Sci* **2**, 340–352.
- Sorsa, L. J., Dufke, S., Heesemann, J. & Schubert, S. (2003). Characterization of an *iroBCDEN* gene cluster on a transmissible plasmid of uropathogenic *Escherichia coli*: evidence for horizontal transfer of a chromosomal virulence factor. *Infect Immun* **71**, 3285–3293.
- van den Bogaard, A. E., Willems, R., London, N., Top, J. & Stobberingh, E. E. (2002). Antibiotic resistance of faecal enterococci in poultry, poultry farmers and poultry slaughterers. *J Antimicrob Chemother* **49**, 497–505.