

# Pathogenomics of the Virulence Plasmids of *Escherichia coli*

Timothy J. Johnson<sup>1\*</sup> and Lisa K. Nolan<sup>2</sup>

*Department of Veterinary and Biomedical Sciences, University of Minnesota, Saint Paul, Minnesota 55108,<sup>1</sup> and Department of Veterinary Microbiology and Preventive Medicine, Iowa State University, Ames, Iowa 50011<sup>2</sup>*

<b>INTRODUCTION</b> .....	<b>750</b>
<b>A Brief History of Bacterial Plasmids</b> .....	<b>750</b>
<b>Plasmid Incompatibility and Classification</b> .....	<b>751</b>
<b>Pathotypes and Plasmids</b> .....	<b>751</b>
<b>ETEC VIRULENCE PLASMIDS</b> .....	<b>753</b>
<b>Human ETEC Plasmids</b> .....	<b>753</b>
<b>Porcine ETEC Plasmids</b> .....	<b>756</b>
<b>K99 plasmids</b> .....	<b>756</b>
<b>987P plasmids</b> .....	<b>756</b>
<b>K88 plasmids</b> .....	<b>756</b>
<b>F18 plasmids</b> .....	<b>757</b>
<b>EAEC PLASMIDS</b> .....	<b>757</b>
<b>EIEC AND SHIGELLA INVASION PLASMIDS</b> .....	<b>758</b>
<b>EHEC AND STEC VIRULENCE PLASMIDS</b> .....	<b>760</b>
<b>EPEC VIRULENCE PLASMIDS</b> .....	<b>762</b>
<b>ExPEC PLASMIDS</b> .....	<b>762</b>
<b>ColV and ColBM Plasmids</b> .....	<b>763</b>
<b>Vir Plasmid</b> .....	<b>764</b>
<b>PHYLOGENY OF <i>E. COLI</i> PLASMIDS</b> .....	<b>765</b>
<b>RepFIB Plasmids</b> .....	<b>765</b>
<b>RepFIIA Plasmids</b> .....	<b>767</b>
<b>RepII Plasmids</b> .....	<b>768</b>
<b>CONCLUDING REMARKS</b> .....	<b>768</b>
<b>ACKNOWLEDGMENTS</b> .....	<b>770</b>
<b>REFERENCES</b> .....	<b>770</b>

## INTRODUCTION

Bacterial plasmids are self-replicating, extrachromosomal replicons that are key agents of change in microbial populations (1, 63). Naturally occurring plasmids are able to promote the dissemination of a variety of traits including drug resistance, virulence, and the metabolism of rare substances (63a). Recombinant plasmids have been essential to the field of molecular biology, but the wild-type plasmids from which these tools were derived are often underappreciated. *Escherichia coli*, perhaps the most-studied microorganism, has been found to possess a variety of plasmid types including those associated with virulence (91). Several types of *E. coli* virulence plasmids exist, including those essential for the virulence of enterotoxigenic *E. coli* (ETEC), enteroinvasive *E. coli* (EIEC), enteropathogenic *E. coli* (EPEC), enterohemorrhagic *E. coli* (EHEC), enteroaggregative *E. coli* (EAEC), and extraintestinal pathogenic *E. coli* (ExPEC). Despite the large number of plasmid types known to occur among *E. coli* strains, plasmids encoding virulence-associated traits fall almost exclusively within a single incompatibility family known

as IncF. Thanks to recent genome sequencing efforts, the evolution of these IncF-type virulence plasmids and the implications of their acquisition by a host *E. coli* strain are now beginning to be better understood.

## A Brief History of Bacterial Plasmids

E. L. Tatum and Joshua Lederberg first described their work involving genetic recombination in *Escherichia coli* in 1947 (180). In 1953, they described “sex in bacteria,” a body of work that was among the earliest to use the terms “episome” and “plasmid” to describe extranuclear structures capable of reproducing in an autonomous state (99). The earliest studies involving bacterial plasmids were focused on those encoding antimicrobial resistance, known as R factors (190). From those studies, it was found that plasmids possessed certain properties including autonomous replication, mobility, host range, and incompatibility with other plasmids (63a). Work involving the fertility factor, or “F factor,” in *E. coli* further broadened our understanding of plasmid transfer and replication. For example, studies of the F factor provided insight into the mechanisms of gene transfer between donor and recipient bacteria and led to the discovery that plasmids contain “replicons” responsible for the control of their own replication (82). Thanks to the extensive work of many researchers involving R plasmids, F factors, and colicin-encoding (Col) plasmids, the core attributes of plasmids were established in the 1960s. The

\* Corresponding author. Mailing address: Department of Veterinary and Biomedical Sciences, University of Minnesota, 1971 Commonwealth Ave., 205 Veterinary Science, Saint Paul, MN 55108. Phone: (612) 626-2542. Fax: (612) 625-5203. E-mail: joh04207@umn.edu.

early identification of an association between plasmids and the transfer of multidrug resistance (MDR) by several Japanese laboratories accelerated interest in plasmid genetics and led to an improved understanding of bacterial conjugation (50). For over 20 years, Pierre Fredericq studied Col factors, which he described as being episomal traits conferring a competitive advantage in microbial populations through the production of substances known as colicins. He discovered that some of these traits were self-transmissible while others were not. He also reported that a nontransmissible Col plasmid, when placed into the same cell as a conjugative plasmid, was transferred to recipient cells; such plasmids were deemed to be mobilizable, or capable of being moved from a donor to a recipient cell in the presence of another plasmid's conjugative apparatus. Fredericq also determined that these plasmids could be integrated into the host chromosome, and much subsequent work involved Hfr strains, which were shown in vitro to be capable of such integration (56–59, 96).

It was later found that properties conferred by bacterial plasmids, such as autonomous replication, transmissibility, stability, and drug resistance, could be localized to individual genetic modules within a plasmid. Such modules could be dissociated and transformed independently into recipient bacteria (63a). The recognition that insertion sequences (ISs) often flank these modules suggested mechanisms of underlying plasmid plasticity; that is, transposition events mediating deletions or introductions of single or large blocks of genes could result in plasmid remodeling (106).

In the 1970s, small high-copy-number plasmids of *E. coli* were modified for use as core tools in recombinant DNA technology. The work of Cohen et al. firmly established plasmids as tools for DNA cloning (33). The development of stable plasmid vectors and the use of restriction enzymes and antibiotics as selective agents contributed to the development of DNA cloning in *E. coli* (63a). Subsequently, other cloning vectors with broad host range and compatibility were developed to suit the needs of research involving other organisms. These fundamental concepts of plasmid biology continue to be used today toward developing the molecular tools needed to study bacteria of emerging interest.

### Plasmid Incompatibility and Classification

The first attempts to classify plasmids in the 1960s were derived from the discovery that certain plasmids could inhibit F-mediated conjugation when present in the same cell. Based on this property, plasmids were initially characterized as  $fi^+$  (fertility inhibition) or  $fi^-$ . These categories reflected similarities in conjugal transfer systems, where  $fi^+$  plasmids were those that were able to inhibit the transfer of a similar conjugative system (191). Because of its limitations, this scheme was replaced by incompatibility typing in the 1970s. This classification scheme was based upon the finding that some plasmids could coexist in the same bacterial host cell, while others could not (63a). Coresident plasmids are defined as being incompatible when they share the same replication mechanisms, and over time, one of these plasmids will be lost by the host strain due to the unstable nature of their coexistence. Since the plasmid replicon type determines the Inc group, the terms “Inc” and “Rep” are used interchangeably to describe plasmid

types (34). The work of Couturier et al. resulted in the development of an invaluable set of DNA probes specific for each known incompatibility group of the *Enterobacteriaceae* that allowed screening for replicon-associated regions that were unique to each known plasmid type (34). Based on these probes, others have devised PCR-based replicon typing schemes (25, 51, 67, 89, 107, 173), which identify the major Inc types occurring among the *Enterobacteriaceae*. Further refinements of replicon-based plasmid classification are likely, as the influx of genome sequencing data continues to increase and thus clarifies our understanding of plasmid structure and also because novel plasmid groups continue to be identified (186). Although recent reviews have acknowledged that as many as 26 incompatibility groups are known, this number will undoubtedly continue to increase in the future (63) (Table 1). Despite this ever-growing number of known plasmid replicon types, only a few have been associated with *E. coli* virulence. The vast majority of virulence-associated plasmids in *E. coli* belong to the F incompatibility group (91).

In an effort to find an alternative to replicon-based plasmid typing, Francia et al. and Garcillan-Barcia et al. recently proposed the use of conjugative relaxases as a plasmid-typing tool (55, 65). There are several potential advantages to using this typing scheme over traditional replicon-based typing. First, it is more extensive, covering a wider range of plasmids. Traditional plasmid-typing schemes have been limited to replication genes of the known incompatibility groups, although these groups are only the tip of the iceberg with regard to plasmid diversity. Relaxase-based typing can theoretically be applied to any plasmid with either a mobilization region or a type IV secretion system (T4SS). Also, most plasmids contain only one relaxase, whereas plasmids frequently contain multiple plasmid replicons (65). Finally, relaxases have been shown to evolve congruently and are thus stable markers of plasmid evolution. They were hypothesized to be under less evolutionary pressures than replication regions (65). Relaxase-based plasmid typing is still in its infancy but holds promise as a future classification schema. Since replicon-based typing is still the most widely used method for plasmid typing and holds historical importance, it will be used here for our comparisons.

### Pathotypes and Plasmids

Pathogenic *E. coli* strains have been divided into different pathotypes based upon the diseases that they cause, the virulence factors that they possess, and their host of isolation (90, 91). The types of disease caused by *E. coli* are grossly classified as being either intestinal or extraintestinal. Six pathotypes of intestinal pathogenic *E. coli* strains are recognized (Table 2), all of which possess virulence-associated plasmids (90, 91). ExPEC strains are a broad group of pathogens that colonize the extraintestinal compartment of animal and human hosts, resulting in such diverse conditions as urinary tract infections (UTIs), meningitis, peritonitis, and sepsis (171). Like the intestinal pathotypes, many ExPEC strains also contain virulence plasmids. As we will see in this review, the vast majority of these *E. coli* virulence plasmids have evolved from a single plasmid backbone type through the acquisition of traits that are essential for and specific to their respective pathotypes.

TABLE 1. List of known plasmid incompatibility (Inc) groups

Inc group	Description	Example in NCBI database	Size (bp) <sup>a</sup>	Source or GenBank accession no.
com9	Plasmids capable of carrying transfer and MDR functions	pIP71A	85,825	Sanger Institute
IncA/C	Broad-host-range plasmids capable of carrying transfer and MDR functions	pSN254	176,473	CP000604
IncB/O	Plasmids capable of carrying transfer, MDR, and virulence functions	pTP113	96,471	Sanger Institute
IncD	Phage-associated plasmids	None		
IncFIA	Plasmids capable of carrying transfer, MDR, and virulence functions	Plasmid F	99,159	AP001918
IncFIB	Plasmids capable of carrying transfer, MDR, and virulence functions	pO157	92,077	AF074613
IncFIC	Plasmids capable of carrying transfer, MDR, and virulence functions	Plasmid F	99,159	AP001918
IncFIIA	Plasmids capable of carrying transfer, MDR, and virulence functions	pR100	94,281	AP000342
IncFIV	Plasmids capable of carrying transfer, MDR, and virulence functions	pSU316	~77,000	P36REPA
IncFV	Plasmids capable of carrying transfer, MDR, and virulence functions	pED208	~90,000	AF411480
IncFVI	Plasmids capable of carrying transfer, MDR, and virulence functions	pSU212	ND	X55895
IncFVII	Plasmids capable of carrying transfer, MDR, and virulence functions	pSU221	ND	P2SINC
IncHI1	Plasmids capable of carrying transfer and MDR functions	pR27	180,641	AF250878
IncHI2	Plasmids capable of carrying transfer and MDR functions	pR478	274,762	BX664015
IncHI3	Plasmids capable of carrying transfer and MDR functions	Mip233	ND	AF192489
IncHII	Plasmids capable of carrying transfer and MDR functions	pHH1508a	ND	ECOTEHAB
IncI1	Plasmids capable of carrying transfer and MDR functions	pR64	120,826	AP005147
IncI2	Plasmids capable of carrying transfer and MDR functions	R721	75,582	AP002527
IncJ	Conjugative, self-transmitting, integrating elements	ICE R391	88,532	AY090559
IncK	Plasmids capable of carrying transfer and MDR functions	pR387	87,645	Sanger Institute
IncL/M	Plasmids capable of carrying transfer and MDR functions	pCTX-M3	89,468	AF550415
IncN	Broad-host-range plasmids capable of carrying transfer and MDR functions	R46	50,969	AY046276
IncP- $\alpha$	Broad-host-range plasmids capable of carrying transfer and MDR functions	RP4	60,099	L27758
IncP- $\beta$	Broad-host-range plasmids capable of carrying transfer and MDR functions	pB4	79,370	AJ431260
IncP- $\gamma$	Broad-host-range plasmids capable of carrying transfer and MDR functions	pQKH54	69,966	AM157767
IncP- $\delta$	Broad-host-range plasmids capable of carrying transfer and MDR functions	pEST4011	76,958	AY540995
IncP6 (IncG)	Broad-host-range plasmids capable of carrying transfer and MDR functions	Rms149	57,121	AJ877225
IncP7	Broad-host-range plasmids capable of carrying transfer and MDR functions	pCAR1	199,035	AB088420
IncP9	Broad-host-range plasmids capable of carrying transfer and MDR functions	pWWO	ND	WWODIRRPTA
IncQ1	Broad-host-range plasmids capable of carrying MDR functions	RSF1010	8,684	RSFRMRA
IncQ2	Broad-host-range plasmids capable of carrying MDR functions	pTC-F14	14,155	AF325537
IncR	Plasmids capable of carrying transfer and MDR functions	pK245	98,264	DQ449578
IncT	Large plasmids carrying transfer and DNA metabolism functions	pRts1	217,182	AP004237
IncU	Plasmids capable of carrying transfer and MDR functions	pFBAOT6	84,749	CR376602
IncV	Plasmids capable of carrying transfer and MDR functions	None		
IncW	Broad-host-range plasmids capable of carrying transfer and MDR functions	pR7K	39,792	AM901564
IncX1	Plasmids capable of carrying transfer and MDR functions	pOLA52	51,602	EU370913
IncX2	Plasmids capable of carrying transfer and MDR functions	pR6K	39,872	Sanger Institute
IncY	Phage-like plasmids	P1	94,481	AF234173

<sup>a</sup> ND, not determined.

Some of these pathotype-specific virulence plasmids are highly conserved; others are extremely diverse.

Plasmid replication, mobility, and stability have been subjected to intense scrutiny since the early days of cloning. Indeed, there are many who have dedicated their entire scientific

careers toward an improved understanding of plasmid biology. Today, there are journals (i.e., *Plasmid*) and scientific organizations (<http://www.ispb.org>) devoted to the study of plasmids and mobile genetic elements. The advent of rapid and affordable genome sequencing, coupled with sustained and intense

TABLE 2. *E. coli* pathotypes with virulence-associated plasmids

Pathotype	Disease	Virulence-associated plasmid(s)
ETEC	Watery diarrhea	Plasmids encoding colonization factors and toxins
EAEC	Persistent diarrhea	Plasmids encoding adherence factors and toxins
EIEC	Watery diarrhea, inflammatory colitis, or dysentery	Invasion plasmid
EHEC	Nonbloody diarrhea, HC, or HUS	Plasmids encoding toxins (pO157)
EPEC	Nonbloody diarrhea	EAF plasmid
ExPEC	UTI, sepsis, neonatal meningitis, or avian colibacillosis	ColV and Vir plasmids
Diffusely adherent <i>E. coli</i>	Nonbloody diarrhea	Plasmids encoding adhesins

interest in this field, has resulted in a great increase in our understanding of plasmid structure and biology. However, this rapid increase in knowledge comes at a cost, as many of these sequences get “lost in the shuffle” because of the large number of draft genome sequences in the NCBI database. Perhaps this is especially true for *E. coli* plasmids, as they are the most represented plasmids in the database. Here, the available virulence plasmid sequences from each *E. coli* pathotype are analyzed, and literature related to these sequences is reviewed in an effort to better understand *E. coli* virulence plasmid composition, evolution, and diversity.

### ETEC VIRULENCE PLASMIDS

ETEC strains are important agents of traveler’s diarrhea in humans, neonatal diarrhea in production animals, and post-weaning diarrhea in swine (39, 120, 185). Collectively, these diseases account for much human and animal misery and can be economically devastating in terms of lost productivity and health care costs. To cause disease, ETEC strains must first colonize the host’s intestinal epithelium. Intestinal colonization is mediated by pili/fimbriae acting as adhesins and promoting adherence (39, 120, 185). Adhesins of human and porcine ETEC strains are encoded by plasmids, as are many other known ETEC virulence factors (39, 120, 185). These adhesins will be referred to throughout as colonization factors (CFs). A second defining trait of ETEC is enterotoxin production, and these toxins can also be plasmid encoded. Highlighted below are the major ETEC types involved in human and porcine disease, their CFs and other virulence factors, and the plasmids encoding these traits.

#### Human ETEC Plasmids

Arguably, human diarrhea caused by ETEC is the most common disease caused by pathogenic *E. coli* strains. It is estimated that there are more than 650 million cases of ETEC infection each year, resulting in nearly 800,000 deaths (147). The majority of these cases occur in underdeveloped countries. Thus, ETEC strains pose a significant threat to the indigenous populations of these countries as well as travelers and military personnel visiting them (185). Human ETEC strains are acquired via the ingestion or handling of contaminated food and water. Following infection, a rapid onset of watery diarrhea ensues, which is usually self-limiting but can cause life-threatening dehydration.

For human ETEC strains to cause disease, they must attach to the small-intestinal epithelia, replicate on mucosal surfaces, evade host defenses, and cause damage to the host (185).

Human ETEC strains produce either heat-labile enterotoxin (LT), heat-stable enterotoxin (ST), or both. LT and ST traits occur uniformly among human ETEC isolates, with about 35% of isolates expressing ST, 35% expressing ST and LT (ST-LT), and the remainder expressing only LT (64). Greater than 90% of the strains expressing ST-LT have a CF, while significantly fewer strains expressing ST or LT alone also have CFs. Human ETEC strains belong to a large number of serogroups, a typing scheme based upon O-antigen cross-reactivity. The most commonly occurring ETEC serogroups are O6, O8, O25, O78, O128, and O153. These comprise ~60 to 70% of the isolates examined worldwide, with the remaining 30 to 40% belonging to a large number of different serogroups (195). While diversity does exist with regard to serogroup, it appears that a certain chromosomal background is required for an ETEC strain to be capable of causing disease.

Perhaps more important to ETEC strains than their chromosomes, however, are the plasmids that they possess. These plasmids encode CFs of ETEC strains, their toxins, and other adhesins. ETEC strains attach to the intestinal mucosa via CFs, which are proteinaceous structures with a high specificity for the host intestinal epithelia. Gaastra and Svennerholm described the broader human ETEC CF groups in an excellent review article (64). It is thought that CFs are involved in the initial attachment of ETEC to the host cells and that other ETEC virulence factors strengthen this attachment and promote tissue invasion (64). Human ETEC CFs can be either plasmid or chromosome encoded (185), typically by a polycistronic operon that includes the fimbrial subunit genes, chaperones, and ushers. However, the majority of human ETEC CFs are plasmid encoded and appear to have been horizontally acquired via flanking ISs and transposons. These CFs have undergone extensive evolutionary modification, resulting in a number of genetic variants. In fact, there are more than 20 known human ETEC CFs that are genetically distinct, with many also possessing distinct serological properties, further demonstrating that these CFs have the ability to move and rapidly evolve (64).

The diversity of the human ETEC CFs complicates our understanding of ETEC pathogenesis. In most of the studies that have characterized human ETEC populations for known CFs, substantial subpopulations of isolates that lacked known CFs were identified, suggesting that the extent of CF diversity is not fully appreciated and that novel CF types remain to be identified (195). Also, a complicating factor in an understanding of the role of CFs in ETEC pathogenesis is the evolving nature of CF nomenclature, with the result being that the literature can be confusing regarding the classification of the different human ETEC CFs. The original CF groups, known as

TABLE 3. Sequenced human ETEC plasmids

Plasmid	Strain	Size (bp)	Trait(s)	IS elements
pCoo	C921b-1	98,396	IncI1 replicon; IncFII replicon; CS1; <i>eatA</i>	IS1, IS3, IS2, IS66, IS91, IS100, IS186, IS630, IS679
pETEC_5	E24377A	5,033	Cryptic	
pETEC_6	E24377A	6,199	Streptomycin and sulfonamide resistance genes	
pETEC_35	E24377A	34,367	IncFII replicon; partial transfer and stability regions	
pETEC_73	E24377A	70,609	CS1; IncI1 replicon; R64-like thin pilus	IS1, IS2, IS3, IS21, IS66, IS186, IS481, IS629
pETEC_74	E24377A	74,224	<i>eatA</i> ; partial F transfer region; many IS elements	IS1, IS3, IS21, IS66, IS91, IS629, IS630
pETEC_80	E24377A	79,237	IncFII replicon; partial F transfer region; <i>ms</i> regulator; LT, CS3	IS1, IS2, IS3, IS21, IS30, IS66, IS629, IS1414
pH10407_66	H10407	66,681	IncFII replicon; F transfer region; LT	IS1, IS3, IS21, IS66
pH10407_95	H10407	94,797	CFA/I; <i>etpABC eatA</i>	IS1, IS2, IS21, IS30, IS66, IS629, ISSf14, IS911

CF antigen (CFA) groups I through IV, are the most prevalent CF types worldwide. In 1975, the first human ETEC CF was described because of its similarities to porcine CF type K88. This plasmid-encoded CF, named CFA/I, was isolated from archetypical human ETEC strain H10407 (47). Plasmids encoding these fimbriae are positively regulated by *trans*-acting regulatory genes on the same plasmids or other plasmids, such as *ms* and *cfar* (14, 26). Evans and Evans subsequently described CFA/II, the second described human CF (46). CFA/II was originally described as being a distinct CF type with similarities to CFA/I (46). However, it was later observed that CFA/II actually consists of three antigenically distinct components, CS1, CS2, and CS3 (172). These can occur as CS3 alone or in combination with CS1 or CS2 fimbriae (64). Similarly, CFA/IV was initially discovered to be a single CF type, but it was later found that the CFA/IV antigen has multiple components, expressing CS6 alone or in combination with CS5. Since the discovery of CFA types I through IV, many other antigenically and genetically distinct CF types have been discovered. Also, the nomenclature has shifted: CFA was replaced by putative CF (PCF), and PCF was subsequently replaced by colonization surface antigen (CS) type (64). Some of the original designations have been replaced by newer nomenclature (i.e., CFA/III is CS8, and the PCF types have all been replaced by CS types), whereas others have retained their original names (i.e., CFA/I) (64).

Plasmid pCoo was the first completely sequenced human ETEC CF-encoding plasmid (62). pCoo was isolated from human ETEC strain C921b-1, which is known to express CS1 and CS3 (143). Sequencing of this plasmid revealed that it was cointegrate in nature, containing regions sharing homology with RepI1 plasmid R64 from *Salmonella enterica* serovar Typhimurium (92, 94) and RepFIIA plasmid R100 from *Shigella* spp. (124). The composite regions of pCoo are separated by IS100-associated direct repeats, suggesting that a recombination event occurring between regions of two different plasmids resulted in the formation of the cointegrate pCoo (62). The key feature identified on this plasmid was the polycistronic *coo* operon, which contains four genes encoding the CS1 pilus. Additionally, pCoo encodes a predicted protein similar to the serine protease autotransporter *EatA*, which was implicated in ETEC virulence (141). The *coo* genes fall within the RepI1-like portion of pCoo, while *eatA* is within the RepFIIA-like portion. Interestingly, this cointegrate plasmid appears to be

stable, since an analysis of clinical CS1<sup>+</sup> isolates revealed that they all contain both cointegrate portions of pCoo (62).

ETEC strain H10407 is a prototypical CFA/I strain isolated from a patient in Bangladesh in 1973 (48). Oral challenge studies with this strain in human volunteers confirmed its ability to cause diarrhea (177). CFA/I is a single fimbrial structure that has been associated with ST enterotoxin on a single mobilizable plasmid (153). Strain H10407 (O78:K80:H11) was recently sequenced by the Wellcome Trust Sanger Institute. At the time of this review, its sequence was publicly available but not yet published ([http://www.sanger.ac.uk/Projects/E\\_coli\\_H10407/](http://www.sanger.ac.uk/Projects/E_coli_H10407/)). Like most ETEC strains, H10407 contains multiple plasmids possessing a variety of virulence-associated genes (Table 3). These plasmids include a 95-kb plasmid, pH10407\_95, encoding CFA/I, *EtpABC*, and *EatA*, and a 66-kb plasmid, pH10407\_66, encoding a conjugative transfer system and LT (Fig. 1). Both of these plasmids appear to possess a RepFIIA-like replication region. pH10407\_66 contains an F-like plasmid transfer region, while pH10407\_95 contains only remnants of this system.

Recently, the genome of archetypical ETEC strain E24377A was sequenced and analyzed (150). This strain belongs to the O139:H28 serotype and produces LT, ST, CS1, and CS3. Strain E24377A contains six plasmids, ranging in size from 5 to 80 kb. The CS1 antigen, as with strain E1392/75, is encoded on a single plasmid, pETEC\_73. A comparison of pETEC\_73 with pCoo (strain C921b-1) and pH10407\_95 (strain H10407) reveals that both pCoo and pETEC\_73 possess the CS1 operon within a RepI1 backbone region, while pH10407\_95 contains CFA/I within a RepFIIA backbone (Fig. 1). The RepI1 backbone regions of pCoo and pETEC\_73 include the RepI1 replication gene (*repZ*), genes encoding RepI1 stability (*sopAB*- and *psiAB*-like), and genes encoding the production of the R64-like thin pilus (Tra). The primary difference between pCoo and pETEC\_73 is the presence of a RepFIIA-like module in pCoo that is absent from pETEC\_73. Also, these plasmids differ slightly within the CS1-encoding region, as they are flanked by differing IS elements in the two plasmids. Thus, it appears that the CS1 operon was introduced into an ancestral RepI1 plasmid, and this occurred prior to pCoo's integration of RepFIIA backbone components and *eatA*.

CFA/I-encoding plasmid pH10407\_95 does not possess any RepI1-related regions but does contain a RepFIIA-like replicon and backbone (Fig. 1). This plasmid contains *eatA* (141);



FIG. 1. Linear comparison of RepFIIA and RepI1 virulence plasmids of *E. coli*. The top two maps are basic RepFIIA and RepI1 backbones, highlighted in orange (RepFIIA) and green (RepI1). These regions are highlighted throughout the figure to illustrate genetic load regions. The scale is in base pairs. Maps begin with either RepFIIA or RepI1 and are drawn to scale. ISs are marked below each map, and virulence genes of interest are marked and colored. The boxed region indicates ETEC plasmids that encode either CS1 or CFA/I. All maps were drawn with XPlasMap (<http://www.iayork.com/XPlasMap/index.shtml>).

the two-partner secretion locus *etpABC*, involved in adhesion to intestinal epithelia (54); and the genes encoding CFA/I. The presence of CFA/I on pH10407\_95, which is a RepFIIA plasmid, suggests that the ETEC CS operons have been acquired on multiple occasions on multiple plasmid backbones. Like pCoo, pH10407\_95 contains a truncated F transfer region. In fact, all three CF-encoding plasmids have incomplete transfer regions, leaving them to rely on other plasmids for mobilization and cotransfer (150, 153). This is a possibility for strain E24377A, which has apparently functional transfer regions on its coresident plasmids (Fig. 1). In addition to the core components and virulence factors mentioned above, many of the sequenced ETEC plasmids and other sequenced *E. coli* viru-

lence plasmids contain a group II intron-encoded reverse transcriptase/maturase in proximity to fimbrial operons or other horizontally acquired genetic regions. Plasmids containing this region include pETEC\_73, pETEC\_74, pETEC\_80, pCoo, pB171 (176), pUTI89 (31), p55989, and the K88<sup>+</sup> and K99<sup>+</sup> plasmids. This intron, E.c.I4, has apparently been inserted into IS629 and IS911, accounting for its mobility and distribution among *E. coli* and *Shigella* sp. strains (112). This intron is widely distributed among *E. coli* populations, but its specific function and importance are not known. Whatever the case, it appears that this intron was inserted into these plasmid backbones prior to the evolution of these pathotype-associated plasmids.

Despite the known genetic diversity of human ETEC plasmids with regard to CF types, only a few these plasmids have been sequenced. Therefore, our full understanding of the evolution of these plasmids and its impact on virulence capabilities remains to be elucidated. Future comparative genomic studies will aid in our understanding of this dynamic and interesting group of plasmids that have relevance to human and animal health.

### Porcine ETEC Plasmids

ETEC strains are routinely isolated from cases of diarrheal disease occurring in neonatal and postweaning pigs. These diseases account for significant losses to producers worldwide (39). Interestingly, porcine ETEC CFs are distinct from human CFs, and the binding specificities of animal and human ETEC strains are often quite different (120). However, both human and porcine ETEC strains possess ST and LT together with their fimbrial adhesins (CFs). In addition to these traits, porcine ETEC strains may also harbor toxin-encoding genes such as EAST1 (120). All of these traits can be plasmid encoded. Porcine ETEC strains are perhaps the best-studied ETEC type, and there are several commonly occurring CFs implicated in ETEC-caused porcine disease. Porcine ETEC strains differ among themselves in the CF types that they contain, which relates to the age of the animal that they infect (39).

**K99 plasmids.** Neonatal diarrhea in piglets is caused primarily by ETEC strains possessing K99 fimbriae (also known as F5 fimbriae) (39). The K99 antigen, encoded by a ~78-kb conjugative plasmid, was identified by Smith and Linggood (169). The K99 antigen is expressed as mannose-resistant fimbriae of ~9 nm in diameter (39). K99 fimbriae mediate bacterial binding to the small intestinal glycolipid ganglioside of newborn pigs. The K99 trait is transmissible, and these plasmids have been found to be highly conserved among K99<sup>+</sup> isolates with regard to size and restriction patterns (81). The K99 operon (*fanABCDEFGH*) includes positive regulators, a major pilin subunit, an usher, a chaperone, and pilus assembly and elongation genes (5, 97, 154, 163). The expression of K99 is temperature dependent, inhibited at lower temperatures, and activated at body temperature (187). Bradley demonstrated that K99<sup>+</sup>-encoding plasmids possess a repressible conjugation system similar to that of the F plasmid (17), and Harnett and Gyles showed that the K99 and STa genetic machineries were linked to a single plasmid of ~80 MDa (70). Conversely, later work by Harnett and Gyles showed that STa and STb occurred on heterogeneous plasmids among porcine ETEC strains independent of the K99 plasmid (71).

Currently, no completed K99 plasmid sequences are available. However, draft sequencing has been performed on the plasmid content of a porcine ETEC K99<sup>+</sup> isolate (<http://www.umn.edu/~joh04207>). In this isolate, the K99 operon was located on a conjugative plasmid containing the RepFIA and RepFIIA replicons and the porcine attaching-and-effacing (A/E)-associated (*paa*) gene (3). This strain also harbored a separate RepII plasmid (our unpublished results). These results are in agreement with the results of a previously reported replicon typing study using K99<sup>+</sup> isolates (107) and a pilus typing study (17).

**987P plasmids.** 987P, or F6, fimbriae are also found in porcine ETEC strains that cause neonatal diarrhea in pigs, with 987P acting to mediate adhesion to intestinal cells (38). 987P pili are subject to phase variation, and this is dependent on growth conditions (80, 121, 122). The operon encoding 987P fimbriae can be located on plasmids or on the bacterial chromosome, although this has been subject to debate (27). Plasmids containing the 987P operon range in size from 35 to 40 MDa (159). The 987P fimbrial gene cluster contains eight genes, *fasABCDEFGH*, adjacent to a *Tn1681*-like transposon containing genes encoding the heat-stable enterotoxin ST1a (160). No 987P-encoding plasmids have been sequenced, and little is known regarding the genetic makeup of these plasmids.

**K88 plasmids.** Postweaning diarrhea in pigs is also caused by ETEC (120). This disease usually occurs within 1 week after weaning, resulting in decreased weight gain and often death. The classical ETEC type implicated in postweaning diarrhea is characterized by the possession of the K88 CF, but recently, ETEC types possessing the F18 CF and other novel CFs have emerged as porcine pathogens. Complicating the control of this disease is a rapid increase in the percentage of multidrug-resistant porcine ETEC strains, which is thought to be due to the extensive use of antibiotics as growth promoters and therapeutic agents (15, 41, 114). While vaccination strategies have been effective in preventing neonatal diarrhea, those aimed at protecting weaned pigs have had mixed results (49). These results are likely multifactorial due to difficulties in stimulating a protective mucosal response and the vast diversity of ETEC strains capable of causing this disease. Classically, however, the key virulence factors that have been implicated in ETEC-caused postweaning diarrhea are hemolysins and K88 or F18 fimbriae (39).

The K88 (F4) antigen was first described in 1961, the first CF of the porcine ETEC types to be discovered (137). K88<sup>+</sup> ETEC strains are the most common etiological agents of neonatal and postweaning diarrhea in pigs, and these strains belong primarily to the O8, O45, O138, O141, O147, O149, and O157 serogroups (65). The K88 antigen is encoded by transmissible plasmids (136) and has been associated with the ability to utilize raffinose as a sole carbon source (170). Shipley et al. showed that the genes encoding K88 production and raffinose fermentation were linked to a ~75-kb nonconjugative plasmid (162). Some larger variants of the K88 plasmid also exist, which do contain a functional transfer region. K88 antigen has been associated with enteritis and edema disease in swine, and it was found that these strains possessed an enhanced ability to adhere to porcine intestinal epithelia (120). K88 fimbriae, encoding adherence to the epithelial mucosa, are classified into three genetic variants based upon their antigenic regions, K88ab, K88ac, and K88ad, with K88ac being the most common variant found among porcine ETEC strains (49). These three variants all bind to carbohydrates or glycoconjugants present on intestinal epithelial cells, intestinal mucus, or red blood cells. However, they have different binding specificities for porcine tissue. The operon encoding the K88 CF includes 10 different genes (*faeABCDEFGHIJ*) (78). These include regulatory genes (*faeAB*), a major subunit (*faeG*), minor subunits (*faeCFHIJ*), and ushers and chaperones (*faeDE*). The mapping of this operon showed that it is surrounded by ISSs, with multiple copies of *IS1* separating *faeA* and *faeB*, the 5' end of the

operon being flanked by IS91, and the 3' end being flanked by IS629 (78). Bradley demonstrated that K88<sup>+</sup> plasmids possessed conjugation systems resembling those of the IncI1 incompatibility group (17). Mainil et al. demonstrated that many K88 plasmid-containing isolates possess an F-type replicon, but only F-type replicons were probed in this study, leaving the possibility that K88<sup>+</sup> plasmids might belong to either the IncF or IncI1 group (108).

No completed sequence of a K88-encoding plasmid is currently available, but draft sequencing has been performed on plasmid preparations of K88ab<sup>+</sup> and K88ac<sup>+</sup> isolates (<http://www.umn.edu/~joh04207>). The K88ab and K88ac operons from these plasmids are highly similar to one another and are located on plasmids with a similar arrangement and a similar core backbone. These sequences are also identical to the K88ab operon described by Huisman et al. (78). These operons occur in close proximity to a raffinose fermentation operon, as previously described in the literature (162), and they appear to be located on a RepFIIA plasmid backbone. The virulence and accessory components of these plasmids have evolved in an IS-mediated fashion typical for this plasmid type.

**F18 plasmids.** As mentioned above, F18<sup>+</sup> ETEC strains are important agents of porcine postweaning diarrhea. The F18 fimbrial adhesins, which are plasmid encoded, display a zigzag pattern when imaged with electron microscopy (123). In addition to their occurrence among porcine ETEC strains causing postweaning diarrhea, F18-encoding plasmids are also found among verotoxin-producing isolates implicated in edema diseases of pigs (29, 30, 37, 73). Typically, F18<sup>+</sup> isolates also possess other plasmids encoding STa and STb. Olasz et al. studied prototypical F18ac<sup>+</sup> porcine ETEC O147 strain Ec2173 (134). They found that the F18-encoding genes *fedABCEF* appear to be located on a 200-kb transmissible plasmid, whereas STa and STb were encoded on a separate plasmid. The F18-encoding plasmid in this study, pF18, was also found to contain a hemolysin determinant. A later study involving a 120-kb plasmid known as pTC identified that it encoded STa, STb, and tetracycline resistance (134). Later analysis of pTC found that it contained a putative pathogenicity island (PAI) with a 10-kb fragment containing genes encoding the ST enterotoxins (52). Mainil et al. later identified the AIDA adhesin as being an additional component of the F18 plasmid (110), and Fekete et al. demonstrated that F18-encoding plasmids possessed a single known plasmid replicon, RepFIC (51). Those authors also observed that variants of the F18 adhesin exist: F18ab-encoding plasmids, which were highly variable in their sizes within a population examined, and F18ac-encoding plasmids, which were found to be of a more consistent size.

While no completed F18-encoding plasmid sequences are available, draft sequencing has been performed on the plasmid complement of an F18<sup>+</sup>, multidrug-resistant ETEC strain, UMNVDL (<http://www.umn.edu/~joh04207>). Within this strain's plasmid complement are at least seven plasmids, ranging in size from ~1 to over 120 kbp. The largest plasmid sequenced contained drug resistance determinants on a Rep1 plasmid backbone. This plasmid contains a *sul3*-associated class 1 integron with adjacent genes encoding resistance to macrolides, streptomycin-spectinomycin, mercury, silver, and quaternary ammonium compounds similar to the novel region described by Liu et al. (103). A second plasmid harbors the

*sepA* gene, which was previously implicated in porcine postweaning diarrhea (66). This plasmid contains a RepFII-type plasmid backbone with an intact F transfer region. A third plasmid of at least 60 kb in size has a RepFIIA/RepFIC-like backbone and contains the F18 fimbrial operon. Additionally, this plasmid possesses a hemolysin operon and the *aidA* adhesin gene, which was shown to play a role in porcine ETEC pathogenesis (152). Data from the sequencing of these plasmids agree with previous observations regarding plasmid sizes, numbers, replicon types, and virulence factors of F18<sup>+</sup> porcine ETEC strains. The completion of these and additional F18-related plasmid sequences will aid in our understanding of essential components of these strains.

## EAEC PLASMIDS

EAEC strains are the most recently described of the *E. coli* intestinal pathotypes (77). These bacteria were first described by Nataro et al. in 1987, based upon their distinct aggregative adherence phenotype, which is seen as a brick-like pattern when the bacteria adhere to cultured HEp-2 cells (129). EAEC strains are considered to be an emergent diarrheal pathotype implicated in traveler's diarrhea and affecting immunocompromised children in developing countries (77). In fact, EAEC strains are second only to ETEC strains as being the most common agent of traveler's diarrhea. It is thought that food and water are the most likely means of transmission (77). Epidemiological studies involving this strain have demonstrated that EAEC virulence is heterogeneous, complex, and likely dependent on multiple bacterial factors and host immune status (126). EAEC pathogenesis is thought to involve three primary steps. First, the bacteria adhere to the intestinal mucosa using aggregative adherent fimbriae (AAF). Second, the bacteria produce a mucus-mediated biofilm on the enterocyte surface. Finally, the bacteria release toxins that affect the inflammatory response, intestinal secretion, and mucosal toxicity (77). Aspects of each of these steps involve plasmid-encoded traits.

A primary virulence factor of EAEC is that encoding the aggregative adherence phenotype (72). This trait was found to be associated with AAF (127) and is localized to a 55- to 65-MDa plasmid, termed the "pAA plasmid" (129). Like ETEC CFs, allelic variants of AAF have been identified. AAF from prototypical EAEC strain 17-2 (127) is genetically distinct from AAF from prototypical strain O42 (126), and their respective allelic variants are named types AAF/I and AAF/II. Other allelic variants of AAF have been described, including AAF/III from prototypical strain 55589 (9) and AAF/IV from strain C1010-00 (16). All the identified AAF allelic types appear to be plasmid encoded, and most of the strains analyzed tend to possess only a single AAF allelic type (72). AAF genes are regulated by an AraC-like transcriptional activator, AggR, and strains containing AggR have been termed "typical" EAEC strains (131). The AAF regulon contains both fimbrial genes and a regulator linked to one another on the pAA-type plasmid. There is evidence that AggR is a global regulator of EAEC virulence, as it exhibits effects on a number of chromosomal virulence factors as well (125). The major AAF pilins regulated by AggR include *aggA* (AAF/I), *aafA* (AAF/II), and *agg3* (AAF/III) (9, 35, 131). AggR also regulates the expression



of *aap*, a dispersin that is highly prevalent among EAEC isolates and facilitates the movement of EAEC across the intestinal mucosa for subsequent aggregation and adherence (77). This dispersin is exported out of the EAEC cell via the anti-aggregation protein transporter system, encoded by the genes *aatPABCD* (8). This ABC transporter system is highly prevalent among EAEC populations, highly conserved, and regulated by *AggR* (77). While few studies have involved large numbers of EAEC isolates, recent work by Jenkins et al. found that two groups of EAEC exist based upon gene clustering. They are distinguished by the presence or absence of genes encoded on plasmid pAA and en bloc sets of genes located on genomic islands near the *pheU* and *glyU* loci (83). The definition of "typical" versus "atypical" EAEC strains has thus been supported by such results, with typical EAEC strains possessing pAA-associated genes and certain chromosomal islands, apparently coinherited.

The EAEC plasmids also encode toxins such as the plasmid-encoded toxins Pet and EAST1 (45). Pet appears to belong to the serine protease autotransporter family and has been shown to confer cytoskeletal rearrangements, suggesting a role for Pet in EAEC pathogenesis (24, 188). EAST1 has been found to activate guanylate cyclase, resulting in ion secretion (128). However, relatively few EAEC strains actually possess the genes encoding Pet and EAST1, so their role in EAEC pathogenesis may be limited (36).

Three EAEC plasmids have been completely sequenced: pO42, belonging to AAF/II<sup>+</sup> strain O42; 55989p, belonging to AAF/III<sup>+</sup> strain 55989; and pO86A1, containing a novel AAF-like operon. All three of these plasmids are F-type plasmids with stability, maintenance, and transfer regions (Fig. 2). Plasmid 55989p is considerably smaller than plasmids pO42 and pO86A1, which is due to truncations in the F transfer region. This plasmid also differs from pO86A1 and pO42 in that it contains a RepFIC replication region instead of RepFIIA, although all three plasmids also contain a second replication region known as RepFIB (Fig. 2). All three plasmids encode their respective AAF types, and each contains the AAF regulatory gene *aggR*. While the AAF types possess considerable genetic diversity, *aggR* is generally highly conserved among the plasmid sequences available. A phylogenetic comparison based upon a nucleotide alignment of available *aggR* sequences revealed that *aggR* genes from AAF types I and III appear to be most closely related, whereas other AAF types are more divergent (Fig. 3). Also, sharing nucleotide similarity with *aggR* is the AraC-type transcriptional regulator *rms* of human ETEC plasmid types (26).

The features common to all three sequenced EAEC plasmids are the AAF operons, *aggR*, and *aatPABCD* (Fig. 2). In all three plasmids, these sequences are present on a RepFIB/FIIA-type backbone. Each of these plasmids also has unique regions not present in the other two sequenced plasmids, including the *pet* gene in pO42, the *ipd* gene in pO86A1 encoding an extracellular serine protease, and the Ets iron transport system in pO42 (Fig. 2). The acquisition of Eit by pO42 is particularly interesting because it was previously found only within ExPEC ColV and ColBM plasmids on a RepFIB/FIIA plasmid backbone (87, 88). Although the EAEC plasmids share a common plasmid backbone and core EAEC-associated virulence genes, the gross genetic composition and synteny of

these three plasmids are quite different from one another. This would suggest that a significant amount of gene shuffling and rearrangement has occurred since the introduction of their virulence-associated module or that this module has been introduced on different occasions.

### EIEC AND SHIGELLA INVASION PLASMIDS

EIEC and *Shigella* strains are described for their ability to cause shigellosis or bacillary dysentery in human hosts (140). Worldwide, this disease is estimated to cause nearly 1 million deaths per year from over 150 million cases (95). EIEC strains are so named for their ability to invade HeLa cells, and in vivo, they can infect human mucosa of the colon and invade M cells, macrophages, and epithelial cells. These interactions with the intestinal mucosa can result in a watery diarrhea that may contain mucous and blood.

EIEC and *Shigella* strains are characterized by their possession of a large plasmid, pINV, which encodes the ability to invade host cells (140). The Inv plasmid is perhaps the most mosaic of the *E. coli* virulence plasmids; in fact, nearly one-third of the Inv plasmid encodes IS elements. Previous analysis of the first two sequenced Inv plasmids revealed a 30-kb region responsible for entry into epithelial cells (22). This region included transcriptional activators (*virB* and *mxiEab*), effectors (*ipaADCB*, *ipgB1D*, and *icsB*), chaperones (*ipgACE* and *spa15*), components of the needle complex (*mxiGHJMD*), and inner membrane protein-encoding genes (*mxiA*, *spa24*, *spa9*, *spa29*, and *spa40*). While some other components of the pInv type III secretion system are scattered throughout the plasmid, the expression of Inv-encoded virulence factors is globally regulated by VirB and MxiE (22, 100).

Multilocus sequence analysis of Inv plasmids was previously performed by using the *ipgD*, *mxiA*, and *mxiC* genes (98). This analysis resulted in the separation of the Inv plasmid into two distinct but closely related clusters, and the two Inv plasmid types identified from these clusters shared different incompatibility properties, even though they all contain the same RepFIIA plasmid replicon. This is likely reflected by historical difficulties in classifying the IncFIIA incompatibility group and its known genetic diversity (156). Recent genome sequencing efforts have resulted in the completed sequences of eight Inv-type plasmids, and these results further underscore the genetic diversity of these plasmids and the presence of a variety of Inv evolutionary intermediates (Fig. 4). Overall, the Inv plasmids appear to have evolved via a series of deletion/acquisition events, with a high level of nucleotide homology (>95%) within their conserved genetic regions. The lack of polymorphic sites within these regions suggests a recent ancestry among these plasmids. Based upon comparative analysis of the eight sequenced Inv plasmids, the core components of these plasmids lie within syntenic blocks of DNA containing a RepFIIA-like plasmid replicon and the *ospB*, *ospD2D1*, *sopAB*, *ospC1D3*, *ipaH<sub>9,8</sub>*, and *traI* genes (Fig. 4). Gross comparisons and sequence alignments of these plasmids suggest that the *Shigella* Inv plasmids have evolved from an EIEC Inv plasmid predecessor via a series of deletions and additions of colinear blocks of DNA (Fig. 4 and 5). It appears that the EIEC Inv ancestral plasmid contained a functional F-type transfer region, the *sopAB* stability genes adjacent to a group II intron-

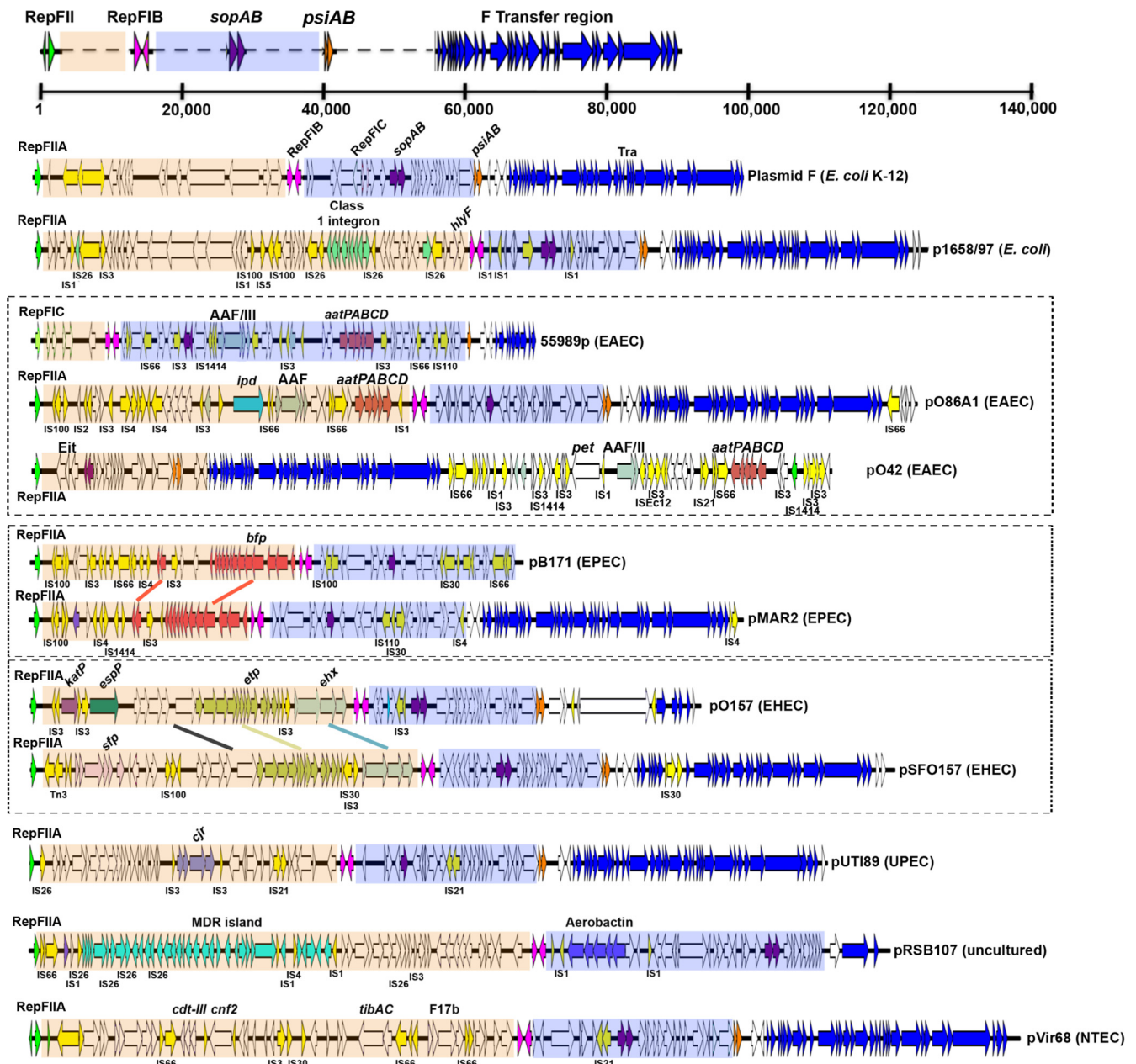


FIG. 2. Linear comparison of RepFIB/FIIA virulence plasmids of *E. coli*. The top map is the basic RepFIB/FIIA backbone, with genetic load regions highlighted in orange or blue. These regions are highlighted throughout the figure to illustrate genetic load regions. The scale is in base pairs. Maps begin with either RepFIIA or RepFIC and are drawn to scale. ISs are marked below each map, and virulence genes of interest are marked and colored. Boxed regions indicate EAEC plasmids (top), EPEC plasmids (middle), and EHEC plasmids (bottom).

encoding reverse transcriptase, and a RepFII-like replicon. All of the sequenced *Shigella* Inv plasmids contain similar remnants of the F transfer region, including its flanking regions and a truncated *traI*, suggesting that a deletion of this region occurred after its introduction into *Shigella* species. In addition to an intact F transfer region, EIEC p53638\_226 also contains a fimbrial operon and a phosphoglycerate transport (Pgt) system that are not present in the *Shigella* Inv plasmids, suggesting that these elements were later acquired by EIEC plasmids or lost by *Shigella* plasmids. p53638\_226 also contains a genetic region encoding iron- and temperature-regulated sensitivity to

colicin Js and the enterotoxin SenB (165) bounded by ISs. This region has been retained by some of the *Shigella* Inv plasmids but lost in others (Fig. 4).

In addition to the loss of portions of the F transfer region, *Shigella* plasmids appear to have evolved from an EIEC plasmid ancestor through the acquisition of type III secretion system components. Of the *Shigella* Inv plasmids sequenced, pSF5 and pWR501 are most closely related to one another based upon alignments of conserved gene sequences (Fig. 5). However, pWR501 contains nearly 90 kb of genetic information more than pSF5, including *ipaH*<sub>2,5</sub>, *shET2*, *ipaJ*, *virB*,

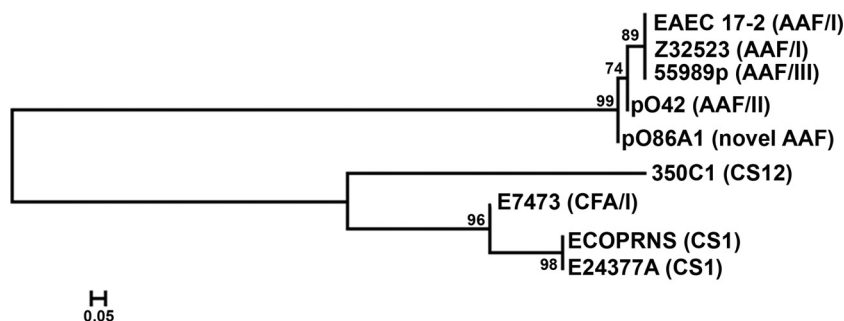


FIG. 3. Evolutionary relationships of sequenced *aggR* (EAEC) and *rms* (ETEC) genes. The evolutionary history was inferred using the neighbor-joining method (157). Bootstrap consensus trees were inferred from 500 replicates, and branches corresponding to partitions reproduced in fewer than 50% of bootstrap replicates were collapsed. Bootstrap confidence values greater than 50% are listed to the left of the nodes. The tree is drawn to scale proportional to evolutionary distances. Distances were computed using the maximum composite likelihood method (179) and are in units of numbers of base substitutions per site. Phylogenetic analyses were conducted by use of MEGA4 (178).

*ipaADCB*, *ipgCBI ADEF*, *mxiGHJKL MEDCA*, *virAG*, and *ushA*. Many of these genes are components of the pInv type III secretion system. Falling in between the apparent minimal *Shigella* Inv plasmid pSF5 and *Shigella* Inv plasmids pCP301 and pWR501, which contain the largest genetic loads, are intermediates characterized by genetic rearrangements, additions, and deletions (Fig. 4). Data from sequence analysis of conserved plasmid genes (Fig. 5) are congruent with data from global comparisons of the sequenced Inv plasmids (Fig. 4); that is, pCP301, pWR501, and pSF5 have similar arrangements of colinear DNA segments, even though pSF5 lacks regions that pCP301 and pWR501 have apparently acquired. In contrast, pSB4\_227 and pBS512\_211 have similar genetic arrangements that are different from pCP301, pWR501, and pSF5, suggesting an independent evolution. pSs046 is divergent from both of the above-described clusters, although it has an arrangement resembling portions of pCP301 and pWR501. Overall, the Inv plasmids have undergone a stepwise evolution characterized by additions and deletions of large blocks of DNA within an apparently short evolutionary time frame. This evolution has been shaped by an abundance of mobile elements within this plasmid type.

#### EHEC AND STEC VIRULENCE PLASMIDS

EHEC strains are a subset of Shiga toxin-producing *E. coli* (STEC) strains responsible for hemorrhagic colitis (HC) and hemolytic-uremic syndrome (HUS) in humans (197). Although *E. coli* strains of several serogroups are members of the EHEC pathotype, *E. coli* strains of the O157 serogroup are responsible for the most severe cases of these diseases (91, 128). EHEC strains are typically known for their ability to produce Shiga toxin and to induce A/E lesions in the host's gut epithelium. While these strains are pathogenic for humans, they can reside peacefully as reservoirs of infection in many production animal types including cattle, swine, and poultry (197). Disease outbreaks in humans are usually associated with the ingestion of some type of food product such as undercooked beef, fresh vegetables, unpasteurized milk and cider, and salami. While the annual health care burden associated with EHEC-caused infections is not overwhelming (\$0.3 to \$0.7 billion per year in the United States), the severity of the disease is alarming. In young children and the elderly particularly, this disease can

progress to HUS, a development that is accompanied by a dramatic increase in rates of morbidity and mortality (197).

The primary virulence determinants of EHEC strains are chromosomally encoded. These include a number of variants in terms of the Shiga toxin and the locus of enterocyte effacement PAI encoding A/E lesion production. However, plasmids may play an important role in the pathogenesis of O157-caused disease. Plasmid pO157 is found in 99 to 100% of clinical O157:H7 isolates from humans (102, 138, 151). However, the role of pO157 in EHEC pathogenesis has not been clearly defined. Some reports have correlated pO157 with hemolytic activity and adherence to intestinal epithelial cells, but the absence of a reliable model of human infection has hindered the progress of our overall understanding of pO157 related to EHEC pathogenesis (23).

pO157, from U.S. HC outbreak-associated strain EDL933, was the first such plasmid sequenced (23). The authors of that study found that pO157 was a 92-kb, F-like plasmid containing previously identified EHEC virulence factors including a hemolysin operon (*ehx*), a type II secretion system (*etpC* to *etpO*), an extracellular protease (*espP*), and a *toxB* homolog. Shortly after the EDL933 pO157 sequence was published, a second pO157 sequence from a strain implicated in an outbreak in Japan was published (111). This plasmid was found to be highly similar to pO157 from strain EDL933, differing essentially only at the single nucleotide polymorphism (SNP) level. Subsequent sequences of multiple pO157 plasmids have been obtained through genome sequencing efforts; these sequences further support initial comparisons suggesting that the pO157 plasmids are essentially identical. More sensitive means have subsequently been applied to study the evolution of plasmid pO157. By using pO157 from the Sakai strain as a reference for the resequencing and mapping of pO157 from multiple strains, Zhang et al. found that differences between pO157 plasmids ranged from few to many at the SNP level (198). Such results reinforce the idea that SNP analysis can be applied on a plasmid genome scale as a more sensitive means for examining the evolution of closely related plasmid genomes.

While O157:H7 strains have classically been characterized because of their inability to ferment sorbitol, sorbitol-fermenting O157:H7 strains have recently emerged as important etiological agents of diarrhea and HUS (21). Brunder et al. pub-

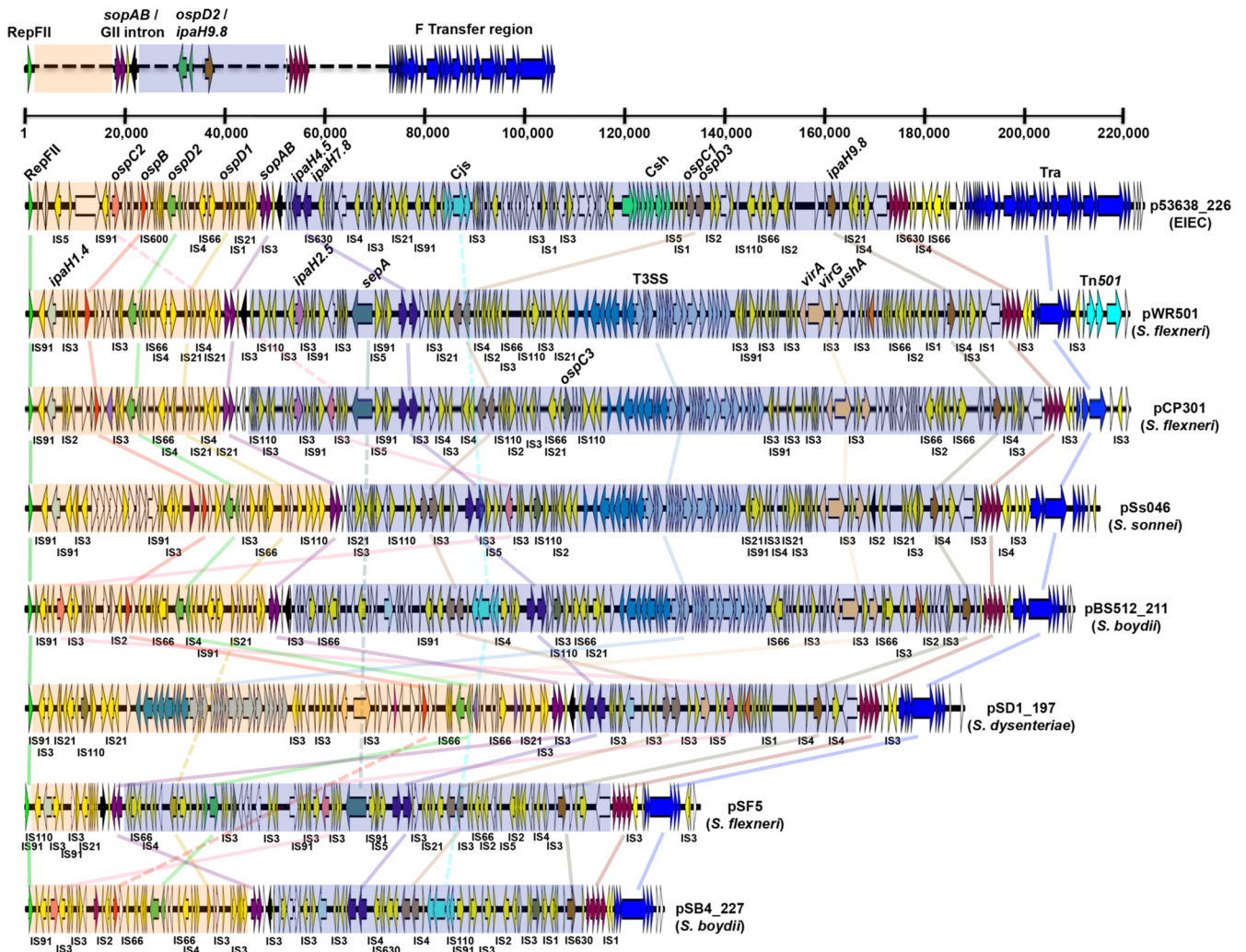


FIG. 4. Linear comparison of Inv virulence plasmids of EIEC/*Shigella* strains. The top map is the basic RepFIB/FIIA backbone, with genetic load regions highlighted in orange or blue. These regions are highlighted throughout the figure to illustrate genetic load regions. The scale is in base pairs. Maps begin with RepFIIA and are drawn to scale. ISs are marked below each map, and virulence genes of interest are marked and colored. Colored lines follow genes of interest throughout the diagram to illustrate gene presence and overall synteny.

lished the first plasmid sequence from a sorbitol-fermenting O157:H<sup>-</sup> isolate (20). This plasmid, pSFO157, was significantly larger than pO157, at 121 kb in size. This size difference was due primarily to the presence of an intact F transfer region and *sfp* fimbriae in pSFO157 (absent in pO157), while pSFO157 was found to lack the *katP*, *espP*, and *toxB* genes present in pO157 (Fig. 2). Similar to the EIEC/*Shigella* plasmids, it was hypothesized that pSFO157 is an ancestor of pO157, which has evolved via reductive evolution involving a loss-of-transfer function (20). It is also thought that the acquisition of these plasmids played a major role in the evolution of the hypothesized ancestral O55 strain, A3, to the extant O157 human pathogens (53, 193).

Large hemolysin-encoding plasmids are found in the majority of STEC isolates, including those not belonging to the O157 serogroup (133). For example, STEC strains belonging to the O113 serogroup can cause sporadic cases of disease that are indistinguishable from some O157-caused diseases. A study by Newton et al. compared the completed sequence of pO113, an

STEC plasmid, to plasmids of O157 isolates (133). Between the two plasmid types, genes shared included the *ehx* (*hly*) hemolysin operon, *espP*, and *iha* (133). However, pO113 also contained genes sharing similarity with the IncI1 transfer region, several putative adhesins and toxins, but it lacked the *toxB* region found in pO157. Those authors also found that pO113 is highly conserved among the O113 STEC isolates that they examined. Furthermore, analysis of the *ehxA* virulence gene and the *repA* gene of the RepFIB replicon demonstrated the evolutionary divergence of plasmids pO157 and pO113 from a common ancestor. Interestingly, phylogenetic analyses using *ehxA*, a virulence factor, and *repA*, a replication gene, were incongruent. This could reflect differences in selective pressures between virulence genes and constitutive genes, but it also emphasizes the difficulties in examining the phylogeny of plasmid genomes, which contain a high degree of plasticity and mobility. Also interesting in this regard is work by Suzuki et al. examining the genomic signatures of all published genomes compared to the signature of pO157 using Mahalanobis dis-

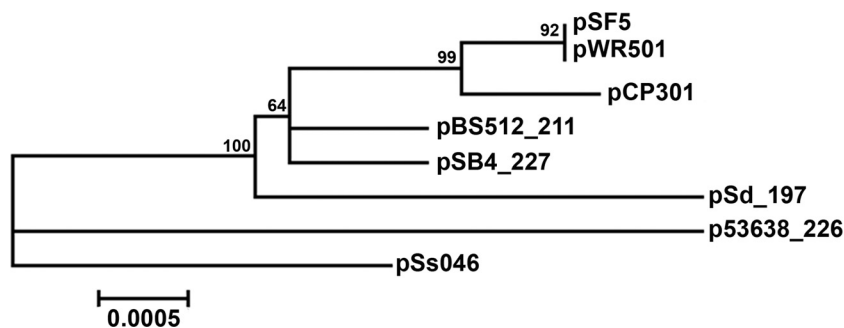


FIG. 5. Evolutionary relationships of concatenated *ospD2* and *traI* genes from Inv plasmids. The evolutionary history was inferred using the neighbor-joining method (157). Bootstrap consensus trees were inferred from 500 replicates, and branches corresponding to partitions reproduced in fewer than 50% of bootstrap replicates were collapsed. Bootstrap confidence values greater than 50% are listed to the left of the nodes. The tree is drawn to scale proportional to evolutionary distances. Distances were computed using the maximum composite likelihood method (179) and are in units of numbers of base substitutions per site. Phylogenetic analyses were conducted by use of MEGA4 (178).

tance (175). They used this approach to generate a hypothesis about the evolutionary history of pO157. By using this approach, it was determined that pO157 shared the closest genomic signature with the *Yersinia pestis* chromosome—closer than its natural host EHEC O157:H7 chromosome. pO157 also shared similar signatures with *Y. pestis* plasmid pCD1. Overall, these results led those authors to hypothesize that pO157 was acquired from *Y. pestis*, which is supported by experimental evidence demonstrating the *in vivo* transfer of genes and plasmids between *E. coli* and *Y. pestis* (75).

#### EPEC VIRULENCE PLASMIDS

EPEC is a common etiological agent of acute and persistent diarrhea in infants aged less than 5 years (28). EPEC strains were first described in the 1940s for their association with infantile diarrhea during summer outbreaks in developed countries. While this phenomenon has apparently subsided in developed countries, EPEC still presents a major problem in developing countries, where frequent outbreaks can have mortality rates approaching 30% (28). The EPEC pathotype was so named by Neter et al. in 1955 while describing primary intestinal pathogens not typically present in the feces of healthy individuals (132). The primary serogroups identified among EPEC isolates include O26, O55, O86, O111, O114, O119, O125, O126, O127, O128ab, O142, and O158 (28, 184). Signs and symptoms caused by EPEC infection include diarrhea, vomiting, fever, and malaise (28). EPEC strains are identified by their trademark A/E histopathology (32). In the gut, they can also cause a decrease in the number and size of microvilli and a loss of the mucosal surface layer. Another property exhibited by EPEC is localized adherence to human cell lines such as HEp-2 (130). Studies by Baldini et al. (6, 7) and McConnell et al. (115) demonstrated that localized adherence could be attributed to a 60-MDa plasmid, and these plasmids were thus named EPEC adherence factor (EAF) plasmids. By using human volunteer challenge studies, a wild-type EPEC strain was found to be significantly more likely to cause diarrhea than was its EAF plasmid-cured derivative, further suggesting a role for the EAF plasmid in EPEC pathogenesis (101). Subsequent studies have shown that the EAF plasmid is responsible for localized adherence but is not required for the formation of A/E lesions (93). Later work iden-

tified a chromosomal PAI, the locus of enterocyte effacement PAI, responsible for the production of A/E lesions (116). Two categories of EPEC have been described: typical EPEC, containing the EAF plasmid, and atypical EPEC, lacking this plasmid. It is currently not completely understood if atypical EPEC strains are less virulent than typical EPEC strains or if they possess an alternative set of virulence factors accounting for the absence of the EAF plasmid.

Three completed EAF plasmid sequences are publicly available. The first plasmid sequenced, pB171, was described by Tobe et al. (182). This plasmid was isolated from EPEC strain B171-8, belonging to the O111 serogroup and implicated in human diarrhea. This plasmid was found to be 68,817 bp in size and to possess the RepFIB and RepFIIA replicons. Within this plasmid was a cluster of genes encoding bundle-forming pili (*bfp* genes), which are responsible for the localized adherence patterns exhibited by EPEC, and the *perABC* genes, involved in the transcriptional activation of *bfp* and other chromosomally encoded virulence factors. The sequencing of a second plasmid, pMAR7, revealed a high degree of colinearity between pB171 and pMAR7 (19) (Fig. 2). pMAR7 was derived from pMAR2 from EPEC strain E2348/69 belonging to the O127:H6 serotype. The primary differences between pMAR7/pMAR2 and pB171 is the presence of an intact F transfer region in pMAR7/pMAR2, which is completely absent from pB171, and 16 ORFs in pB171 not present in pMAR7/pMAR2, which are mostly mobile elements (19). pMAR2 was sequenced as a part of a recent genome sequencing effort involving strain E2438/69, and pMAR7 and pMAR2 were reported to be identical outside of three SNPs and two single-base insertions/deletions within intergenic regions (79). Although the sequenced EAF plasmids shared very strong similarities with one another, further plasmids from isolates representing different EPEC clonal types should be examined to better understand the overall evolution of the EAF plasmid.

#### ExPEC PLASMIDS

Some *E. coli* strains can cause disease in extraintestinal locations. Like the intestinal pathotypes, ExPEC strains carry a distinct set of virulence genes enabling their extraintestinal life-style and pathogenicity (84). Infections due to ExPEC

carry with them a high cost in terms of both human and animal morbidity and mortality and annual costs to the human health care system and to animal owners. A great deal of work has focused on the virulence mechanisms and molecular characterization of ExPEC, leading to a better appreciation of the great diversity within this pathotype. Within this group are ExPEC strains that are adapted for life in the urinary tract (uropathogenic *E. coli* [UPEC]), bloodstream and meninges of neonates (neonatal meningitis *E. coli*), or animal hosts such as birds (avian pathogenic *E. coli* [APEC]). Although we know much about the molecular mechanisms of chromosomally encoded virulence factors of human ExPEC (85), ExPEC virulence factors can also be carried on plasmids. In particular, colicin-encoding plasmids, particularly ColV plasmids, have long been associated with ExPEC virulence (168). In addition to the ColV plasmid, Smith and Huggins identified a second plasmid type associated with invasive *E. coli* of human and animal origin known as the Vir plasmid (166–168). For an excellent and prescient review of ColV plasmids and their properties, see a review by Waters and Crosa (192). Here, a brief history of the ColV and Vir plasmids will be given, and the impact of genome sequencing on our understanding of these plasmid types will be examined.

#### ColV and ColBM Plasmids

ColV plasmids have a history in the scientific literature dating back to the discovery of the phenomenon of “principle V” by Gratia in 1925, who described a substance produced by a “colibacillus strain” capable of lysing other cells (68). Nearly 40 years later, Nagel de Zwaig demonstrated that the colicin V phenotype was transferable (119). It was later found that this transmissibility was linked to F-type plasmids and that these plasmids conferred a virulence phenotype (192). However, work by Quackenbush and Falkow (148) and Williams and Warner (194) demonstrated that colicin V itself, the namesake of the ColV plasmid, was not responsible for the virulence phenotype conferred by these plasmids. Their findings indicated that other traits encoded by ColV plasmids must be responsible for their contributions to virulence. Indeed, Binns et al. reported that iron acquisition mechanisms and serum resistance were linked to a ColV plasmid (11, 12). Iron acquisition in ColV-containing strains was initially attributed to the presence of the aerobactin siderophore system, and serum resistance was attributed to the presence of the *iss* and *traT* genes (12, 194), both of which Johnson et al. confirmed as occurring on ColV plasmids of APEC (86). Dozois et al. described an additional iron acquisition operon on ColV plasmids (42), and Provence and Curtiss described *tsh*, the first serine protease autotransporter described for the *Enterobacteriaceae* (146), which Dozois et al. later reported to be linked to ColV plasmids (42). In addition to ColV production, iron acquisition, and adherence, ColV plasmids have been associated with resistance to chlorine and disinfectants (74), growth in human urine (164), improved growth under acidic pH conditions (149), bacteriophage resistance (76), and the establishment of avian colibacillosis (164, 196) and murine septicemia, meningitis (2), and UTI (164).

In 2006, Johnson et al. published the first complete sequence of a ColV plasmid, pAPEC-O2-ColV, from an APEC isolate

belonging to the O2 serogroup (88). Analysis of its sequence revealed that many of the virulence factors known to contribute to APEC pathogenesis were located within a plasmid-encoded PAI. This sequence was interpreted along with gene prevalence data in an effort to identify core ColV components. The results of that study confirmed the hypothesis set forth by Waters and Crosa (192), that “constant” and “variable” portions of the ColV plasmid PAI exist. The constant region contained the RepFIB replicon (145); the aerobactin operon (189); the Sit iron and manganese transport system (199); a putative outer membrane protease gene, *ompT*; an avian hemolysin gene, *hlyF* (118); a novel ABC transport system known as Ets (88); the salmochelin siderophore system (69); and *iss* (12). The variable portion of this PAI contained the temperature-sensitive hemagglutinin gene *tsh* (43) and another novel transport system known as Eit (88).

Since the completion of the sequence of pAPEC-O2-ColV, three additional ColV plasmid sequences have become available (Fig. 6). Like pAPEC-O2-ColV, pAPEC-1 was isolated from a prototypical APEC isolate belonging to the O78 serogroup (117). pEcoS88 originated from an emergent O45:K1:H7 ExPEC strain causing neonatal meningitis (183). pCVM29188\_146 originated from an *S. enterica* subsp. *enterica* serovar Kentucky isolate from retail chicken breast (60). Despite the diverse sources from which these plasmids originated, their core genetic constitution remains remarkably conserved (Fig. 6). Similar to the EIEC/*Shigella* plasmids, it appears that ColV plasmids have evolved primarily via additions and deletions of large blocks of DNA, since the regions common to the plasmids share very high nucleotide similarities (>95%) and very few SNPs. As mentioned above, the core components of these plasmids are found surrounding the RepFIB replicon, and they include *hlyF* and *ompT*, salmochelin and *iss*, Sit and aerobactin, and a colicin-encoding region. These components are common to all sequenced ColV-type virulence plasmids and are highly syntenic. ColV plasmids also contain an F transfer region and regions encoding plasmid maintenance and stability (Fig. 6). In pAPEC-1 and pAPEC-O103-ColBM, the F transfer region is truncated but in a different manner in each plasmid. Remaining variations in ColV plasmid sizes (ranging from 80 to 180 kb in the literature) reflect other genetic additions or deletions (192).

Besides ColV, other colicin-encoding plasmids were previously identified among ExPEC strains. These plasmids were shown to encode colicins B and M but were not known to possess virulence factors (113, 158). In 2006, Johnson et al. described an APEC plasmid, pAPEC-O1-ColBM, encoding colicins B and M in combination with a PAI similar to that of pAPEC-O2-ColV (87). Such “ColBM” plasmids appear to have evolved from ColV plasmids, since the ColBM regions are embedded with remnants of a truncated ColV operon (Fig. 6). To date, four ColBM plasmids have been sequenced, including three from APEC strains (pAPEC-O1-ColBM, pVM01, and pAPEC-O103-ColBM) (Table 3) (87, 181) and one from an MDR *E. coli* strain isolated from a coastal environment (pSMS35\_130) (Table 3) (61). Interestingly, three ColV/ColBM plasmids (pAPEC-O103-ColBM, pSMS35\_130, and pCVM29188\_146) have acquired MDR-encoding regions, conferring resistance to drugs such as tetracycline, sulfisoxazole, ampicillin, streptomycin, and trimethoprim (60, 61). The

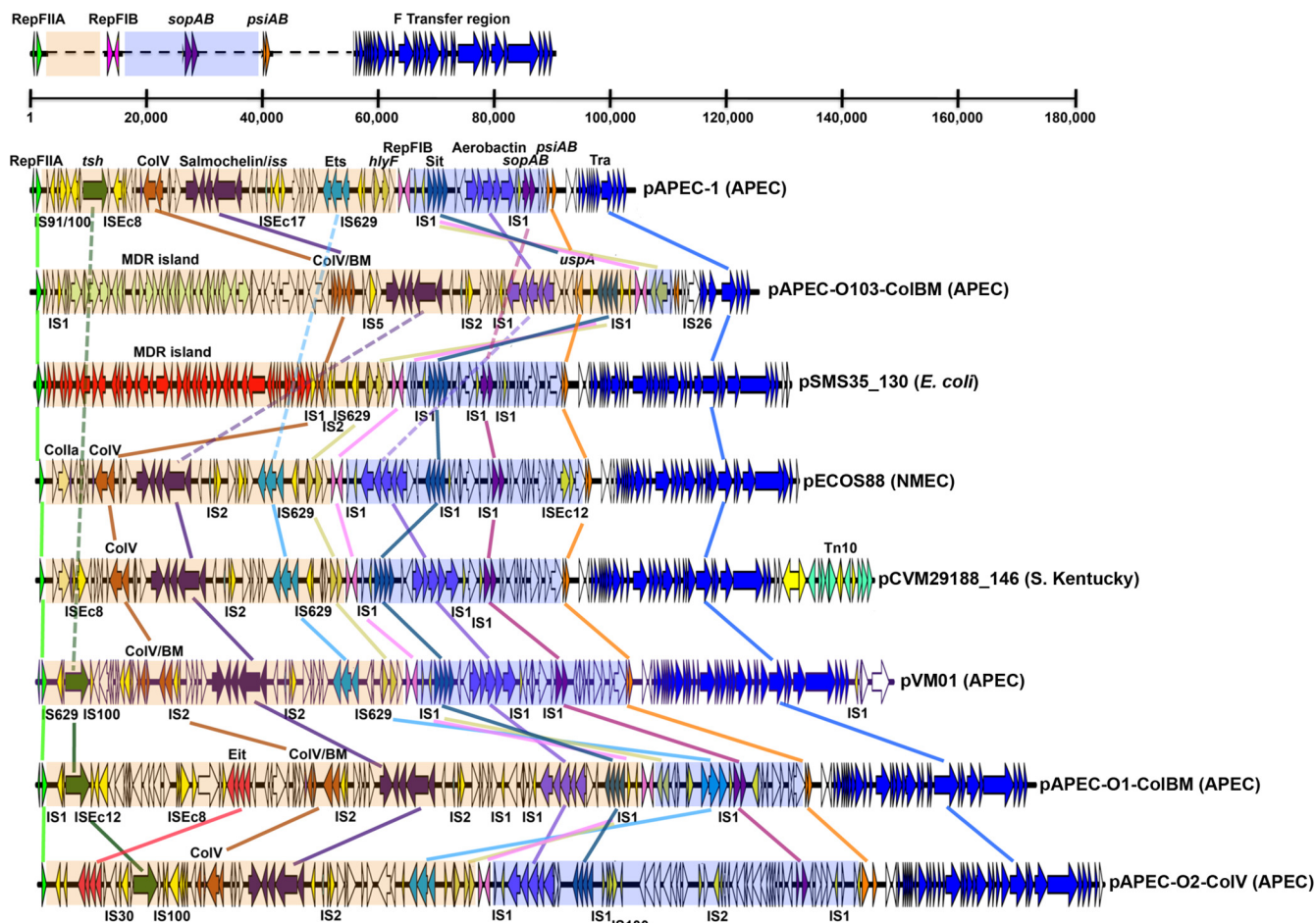


FIG. 6. Linear comparison of ColV and ColBM virulence plasmids of ExPEC. The top map is the basic RepFIB/FIIA backbone, with genetic load regions highlighted in orange or blue. These regions are highlighted throughout the figure to illustrate genetic load regions. The scale is in base pairs. Maps begin with RepFIIA and are drawn to scale. ISs are marked below each map, and virulence genes of interest are marked and colored. Colored lines follow genes of interest throughout the diagram to illustrate gene presence and overall synteny. NMEC, neonatal meningitis *E. coli*.

linkage of MDR-encoding regions with the ColV-encoded PAI is particularly disturbing, since it provides a means for the selection of highly virulent strains through the use of antibiotics. pCVM29188\_146 harbors typical ColV-associated virulence genes but also contains Tn10 downstream of the F transfer region, an atypical region for accessory components (Fig. 6). pSMS35\_130 was found to contain an MDR-encoding island encoding high-level resistance to at least six antimicrobial agents in total. While pSMS35\_130 harbors an MDR-encoding region and *hlyF*, *ompT*, and *sitABCD*, it is an atypical ColV/ColBM plasmid because it lacks many of the other typical ColV PAI virulence factors, including *iss*, *iroBCDEN*, and the salmochelin and aerobactin operons. In contrast, pAPEC-O103-ColBM possesses both the typical ColV PAI virulence factors and an MDR island. Still, both of these plasmids contain only minimal accessory components compared to the ColBM plasmids pAPEC-O1-ColBM and pVM01, which are from typical APEC isolates and contain a larger number of additional uncharacterized hypothetical genes (Fig. 6). While the ColV and ColBM plasmids might be considered a monophyletic group of plasmids with a highly conserved core genetic

makeup, diversity within this group is rampant due to apparently frequent IS-mediated recombinational events. This is supported by our analysis of the eight sequenced ColV and ColBM plasmids, where none are identical with regard to gene content. As a result, the noncore components of these sequenced plasmids are highly heterogeneous and result in a large ColV plasmid “pangenome,” or a collective set of all genes present in ColV plasmids. Most of these accessory components appear to have been introduced into their respective plasmids via recombination events mediated by IS elements such as *IS1*.

### Vir Plasmid

Necrotoxicogenic *Escherichia coli* (NTEC) strains are *E. coli* strains that produce cytotoxic necrotizing factor (CNF) (40). CNFs are strong toxins that cause tissue damage and host disease. NTEC strains are responsible for various diseases of humans and animals including UTI, septicemia, and diarrhea (4, 13, 18, 39, 40, 135, 174). Thus, NTEC strains are both intestinal and extraintestinal pathogens. A subgroup of NTEC

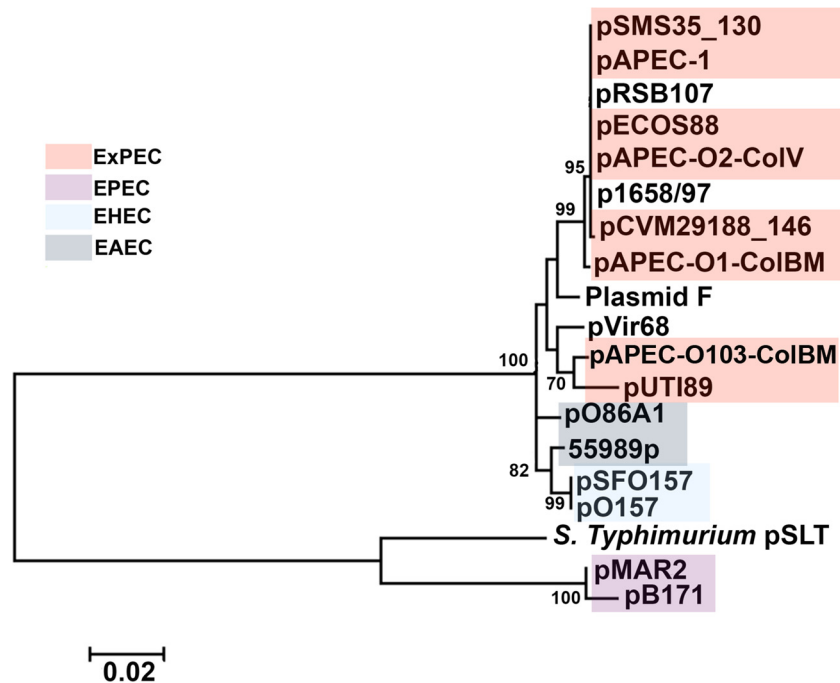


FIG. 7. Evolutionary relationships of the RepFIB *repA* gene from RepFIB/FIIA plasmids. The evolutionary history was inferred using the neighbor-joining method (157). Bootstrap consensus trees were inferred from 500 replicates, and branches corresponding to partitions reproduced in fewer than 50% of bootstrap replicates were collapsed. Bootstrap confidence values greater than 50% are listed to the left of the nodes. The tree is drawn to scale proportional to evolutionary distances. Distances were computed using the maximum composite likelihood method (179) and are in units of numbers of base substitutions per site. Phylogenetic analyses were conducted by use of MEGA4 (178). Colored boxes represent the host pathotype source of the plasmid sequenced.

strains are known as NTEC-2, characterized by their possession of the CNF variant known as CNF-2. The defining traits of NTEC-2 are encoded on a virulence plasmid. In 1974, Smith first described this virulence plasmid that later came to be known as Vir (166). The Vir plasmid was described for its association with a surface antigen, ability to produce the Vir cytotoxin, and lethality in chickens. In 1980, Lopez-Alvarez and Gyles described the Vir plasmid in *E. coli* strains causing human and bovine septicemia (104). Lopez-Alvarez et al. also determined that the NTEC-2 Vir plasmid was transmissible to avirulent *E. coli* recipients and that it was F-like in nature (105). Later, it was shown that these plasmids encoded CNF-2, which was responsible for the cytopathology and lethality in mice caused by certain NTEC-2 strains (40). The Vir plasmid was also found to encode an F17-like fimbria known as F17b, although this fimbrial operon was not found on all *cnf2*<sup>+</sup> isolates examined (10, 44, 109).

Previous work identified and sequenced some apparent core components of the NTEC-2 Vir plasmid, including regions encoding CNF-2 (139), type III cytolethal distending toxin (142), and F17b fimbriae (44). A completed sequence of an NTEC-2 Vir plasmid is available (unpublished data) (GenBank accession number CP001162). Analysis of this plasmid, pVir68, revealed that it is 138,362 bp in size and contains RepFIB and RepFIIA replicons. As expected, pVir68 contains *cdtABC*, *cnf2*, and the F17b fimbrial operon, all of which are identical to previously published sequences (Fig. 2). pVir68 also contains a functional F transfer region and plasmid stability and maintenance regions. Furthermore, this plasmid contains two previ-

ously unidentified regions: a putative novel fimbrial operon and the *tibAC* adhesin/invasin genes, shown elsewhere to induce bacterial aggregation and biofilm formation (161). Overall, the NTEC-2 Vir plasmid appears to have evolved in a typical fashion compared to other virulence plasmids of RepFIB/FIIA origin, with the acquisition of virulence factors via recombination events occurring within a basic RepFIB/FIIA backbone.

## PHYLOGENY OF *E. COLI* PLASMIDS

### RepFIB Plasmids

Gross plasmid comparisons can be used in combination with sequence analysis in an effort to understand how genomes have evolved. By using all of the sequenced RepFIB/FIIA plasmids, some insights can be achieved (Fig. 2, 7, and 8). First, all RepFIB/FIIA plasmids possess some core components, including their two respective replicons, the *sopAB* and *psiAB* regions, ensuring their stability, and the F transfer region, ensuring their self-transmissibility (Fig. 2). It is evident that two “genetic load” regions exist, directly upstream and downstream of the RepFIB locus. The upstream genetic load region tends to contain most of the acquired virulence factors for each plasmid type and tends to have the greater load, whereas the downstream genetic load region contains the stability regions, with most genetic acquisitions occurring in between RepFIB and *psiAB*. Exceptions to this observation are pO42, which has acquired much of its virulence-associated genetic load down-



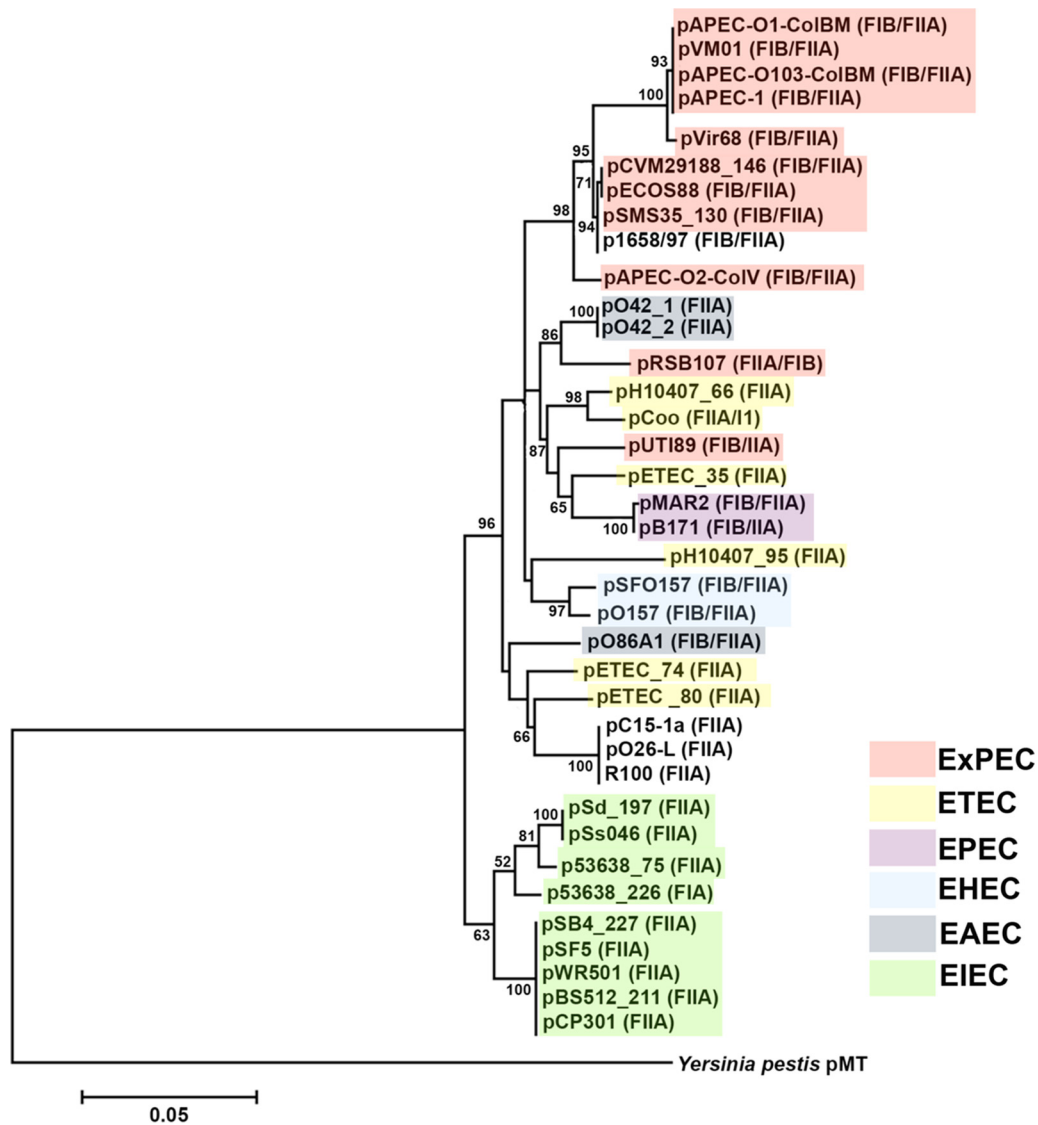


FIG. 8. Evolutionary relationships of the RepFIIA *repA1* gene from RepFIB/FIIA and RepFIIA plasmids. The evolutionary history was inferred by using the neighbor-joining method (157). Bootstrap consensus trees were inferred from 500 replicates, and branches corresponding to partitions reproduced in fewer than 50% of bootstrap replicates were collapsed. Bootstrap confidence values greater than 50% are listed to the left of the nodes. The tree is drawn to scale proportional to evolutionary distances. Distances were computed by using the maximum composite likelihood method (179) and are in units of numbers of base substitutions per site. Phylogenetic analyses were conducted by use of MEGA4 (178). Colored boxes represent the host pathotype source of the plasmid sequenced.

stream of the F transfer region and upstream of the RepFIIA replicon, and pCVM29188\_146, which has acquired *Tn10* in the same location. Nevertheless, acquisitions within this region appear to be highly unusual. RepFIB/FIIA plasmids have the ability to acquire both virulence factors and MDR-encoding islands, as illustrated by plasmids pAPEC-O103-ColBM, pSMS35\_130, pRSB107, and pCVM29188\_146 (Fig. 2). The NCBI database also contains numerous examples of possible evolutionary intermediates of these plasmids. For example, p1658/97 (200) is a plasmid from a human clinical outbreak strain that contains the *hlyF* gene adjacent to the RepFIB region, an arrangement that has been found only for ColV and ColBM plasmids. However, this plasmid lacks other ColV-associated genes and instead contains a class 1 integron. A

second example of a possible evolutionary intermediate is pRSB107, which was isolated from a sewage treatment plant (176). This plasmid contains the aerobactin operon, which is a typical ColV component, but lacks other ColV-associated genes and instead contains a large MDR-encoding island (Fig. 2). As the representation of F-type plasmids in the NCBI database increases, so will examples such as those highlighted above.

The underlying mechanisms controlling the apparently stable genetic makeup of these plasmids (i.e., the clearly defined genetic load regions and conserved core components) are not completely understood, nor is it completely understood what drives the evolution of the genetic load regions. Since these regions surround the RepFIB replication gene and its adjacent

site-specific recombinase, both of which are conserved among these plasmids, it is possible that this region plays a role in such evolution. Perhaps the RepFIB site-specific recombinase acts upon IS elements present within these regions during plasmid replication and segregation. Also, some of the sequenced RepFIB/FIIA plasmids have far fewer IS elements than others, suggesting that the ancestral RepFIB/FIIA plasmid backbone lacked IS elements. Thus, these plasmids have been greatly affected by the acquisition of IS elements over time, which in turn introduced the virulence factors and resistance-encoding genes that now characterize these plasmid types. Also interesting is the loss of a functional F transfer region by a number of sequenced virulence plasmids. Intuition would suggest that these plasmids do not want to render themselves transfer deficient; thus, a truncation in the F transfer region would seem undesirable. However, when one considers that most wild-type *E. coli* strains harbor multiple plasmids that might contain more than one T4SS, the truncation of the F transfer region might be more easily explained. Perhaps multiple T4SSs in the same bacterium actually interfere with the conjugative process. In this case, it would be beneficial for the duplicate transfer regions to be deleted. This would result in the formation of plasmids with truncated transfer regions, still capable of conjugation and dissemination via the T4SSs of other plasmids within the host complement.

The alignment and phylogenetic analysis of the RepFIB *repA* gene shows that it is quite conserved (Fig. 7). As expected, the ColV/ColBM plasmids cluster together with pRSB107 and p1658/97, both mentioned above as being possible ColV/ColBM intermediates. EHEC, EPEC, and EAEC plasmids also cluster with members of their own respective groups. Interestingly, EPEC plasmids are distinct from the other RepFIB/FIIA plasmids and actually clustered closest to outgroup plasmid pSLT from *S. Typhimurium*, suggesting that EPEC plasmids do not share a recent ancestry with other RepFIB/FIIA plasmids. An alignment of RepFIIA *repA1* replication genes from the RepFIB/FIIA plasmids suggests that the addition of RepFIB to RepFIIA to form the RepFIB/RepFIIA plasmid type might have occurred on multiple occasions (Fig. 8). The resulting dendrogram from this alignment depicts RepFIIA and RepFIB/FIIA plasmids clustering together within multiple lineages, suggesting that plasmids within these lineages shared a common RepFIIA ancestor lacking the RepFIB replicon in each case and that RepFIB was subsequently acquired during multiple recombinational events. A second possibility is that of reductive evolution, where the shared ancestor contained both replicons and evolved via the loss of the RepFIB replicon by RepFIIA plasmids. However, the former hypothesis is better supported by a similar situation in the *Salmonella* virulence plasmids, with some plasmids containing both RepFIB and RepFIIA and others containing only RepFIIA (155). The evolution of the RepFIC replicon from RepFIIA, or vice versa, suggests that RepFIB was present on these plasmids prior to the change within the RepFIIA region. This can be observed for plasmid 55989p, where RepFIC resides in a location similar to that of RepFIIA, and RepFIB is also present (Fig. 2). However, the lack of a sequenced RepFIB<sup>+</sup> RepFIC<sup>-</sup> RepFIIA<sup>-</sup> plasmid makes it difficult to draw conclusions on the overall evolution of this plasmid type.

With the ColV and ColBM plasmids, the presence of either

colicin-encoding operon is incongruent with the alignments of their replication genes (Fig. 2, 5, 7, and 8); that is, multiple phylogenetic groups emerge based upon replication gene alignments that contain both ColV and ColBM plasmids. This supports the idea that these plasmids have arisen from a RepFIB/FIIA ancestral plasmid on multiple occasions. This is also supported by the findings by Christenson and Gordon showing that ColBM plasmids have arisen on at least three separate occasions (31). The coevolution of similar plasmid types on multiple occasions also provides further evidence that the possession of a particular set of core virulence factors provides the host bacterium with the tools that it needs to cause a specific disease type. Since there are fewer examples of EAEC, EHEC, and EPEC plasmids in the database, it is difficult to determine if they have evolved in a fashion similar to that of the ColV and ColBM plasmids. However, the examples available for these pathotypes thus far suggest that multiple plasmids from each of these pathotypes share a common ancestor (Fig. 7).

### RepFIIA Plasmids

From a genetic backbone standpoint, the RepFIIA plasmids are more variable than the RepFIB/FIIA plasmids (Fig. 1). The core components of these plasmids still include the RepFIIA region, *psiAB*, and *sopAB* but in a less conserved manner. The *sopAB* and *psiAB* genes are not always present on every sequenced RepFIIA plasmid, and their genetic arrangement is sometimes in a reverse orientation within the plasmid. Still, these plasmids generally contain these core elements and have a genetic load region between the RepFIIA replicon and *psiAB*. The larger number of sequenced RepFIIA plasmids provides us with a glimpse into the possible range of this genetic load, with plasmids ranging in size from 35 kb to greater than 100 kb. As with the RepFIB/FIIA plasmids, the RepFIIA plasmids appear to have evolved from an ancestral RepFIIA<sup>+</sup> *sopAB*<sup>+</sup> *psiAB*<sup>+</sup> *Tra*<sup>+</sup> plasmid that acquired IS elements, which subsequently facilitated the acquisition of virulence factors and drug resistance-encoding genes. Like the RepFIB/FIIA plasmids, some RepFIIA plasmids have truncated F transfer regions, rendering them apparently unable to self-transfer. Also worth noting is the presence of three plasmids in ETEC strain E24377A (pETEC\_35, pETEC\_74, and pETEC\_80) with similar RepFIIA replication regions. The presence of multiple plasmids with similar replicons challenges the idea that they are incompatible with one another and should not coexist in a stable manner. While the stability of these plasmids in their wild-type host is not known, Rasko et al. noted that an analysis of strain E24377A suggests that ETEC is a "pathovar in flux" (150). Quite possibly, the plasmid sequences of strain E24377A in the NCBI database are simply a snapshot in time of these plasmids (or cointegrates thereof) in flux. This is particularly evident when E24377A's plasmids are compared to those of strains C921b-1 (pCoo) and H10407 (pH10407\_95 and pH10407\_66). pCoo forms a cointegrate between the RepFIIA and RepI1 plasmid backbones and contains *eatA* and the CS1 operon (Fig. 1), while pH10407\_95 is a RepFIIA plasmid containing *eatA*, CS1, *aatPABCD* (typically found among EAEC plasmids), and CFA/I. In E24377A, CS1 is encoded on a RepI1 backbone, and *eatA* is encoded on a

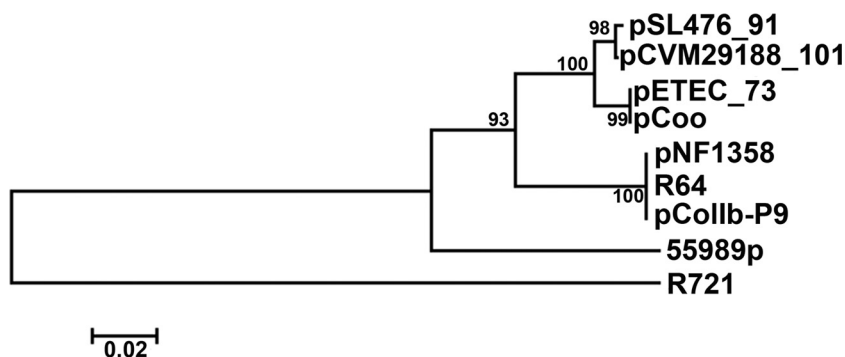


FIG. 9. Evolutionary relationships of the RepIa *repZ* gene from sequenced RepI1 plasmids. The evolutionary history was inferred by using the neighbor-joining method (157). Bootstrap consensus trees were inferred from 500 replicates, and branches corresponding to partitions reproduced in fewer than 50% of bootstrap replicates were collapsed. Bootstrap confidence values greater than 50% are listed to the left of the nodes. The tree is drawn to scale proportional to evolutionary distances. Distances were computed by using the maximum composite likelihood method (179) and are in units of numbers of base substitutions per site. Phylogenetic analyses were conducted by use of MEGA4 (178).

separate RepFIIA plasmid. Certainly, the RepFIIA plasmids are in constant flux, and recombination with their coresident plasmids is inevitable, as illustrated here. Given that there are more than 20 known CF types in human ETEC strains alone, it will be very interesting to see how additional sequenced ETEC plasmids compare to what is currently available.

Overall, the alignment of RepFIIA's *repA1* genes from the sequenced plasmids distinguishes the EIEC plasmids from other *E. coli* virulence and MDR plasmids (Fig. 8). Typical MDR-encoding plasmids (pC15-1a and pO26-L) seem to be related to the prototypical MDR-encoding plasmid R100. As expected, p1658/97 seems to be a relative of the ColV and ColBM plasmids, since it contains *hlyF* and a RepFIB region most similar to these plasmids, and possibly represents a predecessor to such plasmids. ETEC plasmids belong to multiple lineages, further emphasizing their diversity and likely emergence on multiple occasions.

### RepI1 Plasmids

Since several *E. coli* virulence plasmids form cointegrates with RepI1 plasmid sequences or contain solely a RepI1 backbone, they should also be mentioned here. Fewer completed plasmid sequences belonging to the RepI1 group are available (Fig. 9). pSL476\_91 is a 91.4-kb plasmid from an *S. enterica* subsp. *enterica* serovar Heidelberg isolate containing a functional transfer region and genes encoding Collb activity and immunity. pCVM29188\_101 is a 101-kb plasmid isolated from *S. enterica* subsp. *enterica* serovar Kentucky containing a transfer region, a composite transposon-like element (containing *bla*<sub>CMY-2</sub>-*blc-sugE*), and Collb activity and immunity genes. Other plasmids closely related to the RepI1-Collb plasmids according to RepI1 alignments are pCoo and pETEC\_73, CS1-encoding human ETEC plasmids. These plasmids lack the Collb genes but still possess a RepI1 backbone with Tra genes. Their similarity based upon the *repZ* gene alignment (RepI1) suggests that an ancestral CS1-containing RepI1 unit merged with a RepFIIA plasmid to form cointegrate plasmids pCoo and pETEC\_73. Apparently divergent from the above-mentioned plasmids, but sharing common ancestry with one another, are pCollb-P9, a 93.4-kb prototypical Collb plasmid

with a Tra region (90); pNF1358, a 99.3-kb plasmid similar in content to pCVM29188\_101; and R64, a 120-kb prototypical RepI1 MDR-encoding plasmid (96). Finally, most divergent from the above-mentioned plasmids is 55989p, an EAEC AAF/III RepFIB/FIC plasmid, with its RepFIC region sharing homology with RepI1. RepI1 plasmids represent an additional plasmid type capable of possessing and disseminating *E. coli* virulence factors.

### CONCLUDING REMARKS

With over 40 *E. coli* plasmid sequences now completed and multiple plasmid sequences available for each *E. coli* pathotype, the pathogenomics of these diverse plasmids are only now beginning to be fully understood (Table 4). Comparisons made at the plasmid genome-wide level emphasize their plasticity, which has been daunting to analyze, particularly with regard to plasmid evolution. Despite this diversity, *E. coli* virulence plasmids are restricted to a few plasmid backbones that have conservation and synteny with regard to their core components. These plasmids contain distinct genetic load regions, which appear to evolve via IS-mediated site-specific recombination. There are now numerous examples in the database of *E. coli* virulence plasmids that have also acquired MDR-encoding islands. The ease with which this might occur is particularly disturbing, as is the means by which these elements might disseminate. Certainly, these plasmid sequences are only a snapshot view of bacterial genomes in flux. As illustrated with ETEC plasmid comparisons, multiple plasmids within a single bacterial host are under constant pressure and are more amenable to recombination than their host chromosomes. Therefore, rapid changes in bacterial populations brought forth via changes in plasmid content should be of primary concern to those involved in human health and food safety.

The advent of pyrosequencing, its application to bacterial plasmid complements, and the evolving comparative tools for plasmid analysis will facilitate more meaningful comparisons in the future. Efforts are currently under way to improve the ability to perform meaningful annotation and analyses of bacterial plasmids, including a plasmid-focused annotation and analysis site (<http://www.theseed.org>) and an *E. coli* plasmid

TABLE 4. Sequenced *E. coli* virulence-associated plasmids

Plasmid	Host species or <i>E. coli</i> pathotype of origin	Host strain	Size (bp)	Replicon(s)	Key component(s) <sup>a</sup>	GenBank accession no. or source
pSMS35_130	<i>E. coli</i>	SMS-3-5	130,440	FIB, FIIA	<i>sit, hlyF, ompT</i> , ColBM	CP000971
p55989	EAEC	55989	72,482	FIB, FIC	AAF/III, <i>aggR, aatPABCD</i>	CU928159
p042	EAEC	42	113,346	FIIA	AAF/II, <i>aggR, aatPABCD, pet</i> , Tra	Sanger Institute
pO157_EDL933	EHEC	EDL933	92,077	FIB, FIIA	<i>hly, katP, etp, toxB, aspP</i>	AF074613
pO157_TW14588	EHEC	TW14588	92,381	FIB, FIIA	<i>hly, katP, etp, toxB, aspP</i>	NZ_DS999999
pO157_Sakai	EHEC	Sakai	92,721	FIB, FIIA	<i>hly, katP, etp, toxB, aspP</i>	AB011549
pO157_EC4115	EHEC	EC4115	94,644	FIB, FIIA	<i>hly, katP, etp, toxB, aspP</i>	CP001163
pO157_EC4206	EHEC	EC4206	94,707	FIB, FIIA	<i>hly, katP, etp, toxB, aspP</i>	ABHK02000007
pO157_EC4045	EHEC	EC4045	94,753	FIB, FIIA	<i>hly, katP, etp, toxB, aspP</i>	ABHL02000008
pO157_EC4042	EHEC	EC4042	95,362	FIB, FIIA	<i>hly, katP, etp, toxB, aspP</i>	ABHM02000004
pO86A1	EHEC	DIJ1	120,730	FIB, FIIA	Novel AAF, <i>aggR, aatPABCD</i>	AB255435
pSFO157	EHEC	3072/96	121,239	FIB, FIIA	<i>hly, katP, etp, toxB, aspP, sfp</i> , Tra	AF401292
pO113	EHEC	EH41	165,548	I1	<i>ehxABD, espP, saa, iha</i>	AY258503
pB171	EPEC	B171	68,817	FIB, FIIA	<i>bfp, per, gadC</i>	AB024946
pMAR7 (pMAR2)	EPEC	E2348/69	101,558	FIB, FIIA	<i>bfp, per, gadC</i> , Tra	DQ388534
pETEC_5	EPEC	E24377A	5,033	ND		CP000801
pETEC_6	EPEC	E24377A	6,199	ND	Str, Sul	CP000798
pETEC_35	EPEC	E24377A	34,367	FIIA	Tra	CP000796
pH10407_66	EPEC	H10407	66,681	FIIA	LT	Sanger Institute
pETEC_73	EPEC	E24377A	70,609	I1	CS1	CP000797
pETEC_74	EPEC	E24377A	74,224	FIIA	<i>eatA</i>	CP000799
pETEC_80	EPEC	E24377A	79,237	FIIA	CS3, LT, <i>rns</i>	CP000795
pH10407_95	EPEC	H10407	94,797	FIIA	CFA/I, <i>etpABC, eatA</i>	Sanger Institute
pCoo	EPEC	C921b-1	98,396	FIIA, I1	CS1, <i>eatA</i>	CR942285
pAPEC-1	ExPEC	chi7122	103,275	FIB, FIIA	Aerobactin, <i>sit, hlyF, ompT, ets, iss, salmochelin</i> , ColV, <i>tsh</i>	CP000836
pAPEC-O103-ColBM	ExPEC	O103	124,705	FIB, FIIA	Aerobactin, <i>sit, hlyF, ompT, iss, salmochelin</i> , ColBM	CP001232
pECOS88	ExPEC	S88	133,853	FIB, FIIA	Aerobactin, <i>sit, hlyF, ompT, ets, iss, salmochelin</i> , ColV	CU928146
pCVM29188_146	<i>S. Kentucky</i>	CVM29188	146,811	FIB, FIIA	Aerobactin, <i>sit, hlyF, ompT, ets, iss, salmochelin</i> , ColV, <i>tsh</i>	CP001122
pVM01	ExPEC	E3	151,002	FIB, FIIA	Aerobactin, <i>sit, hlyF, ompT, ets, iss, salmochelin</i> , ColBM, <i>tsh</i>	EU330199
pAPEC-O1-ColBM	ExPEC	APEC O1	174,241	FIB, FIIA	Aerobactin, <i>sit, hlyF, ompT, ets, iss, salmochelin</i> , ColBM, <i>tsh</i>	DQ381420
pAPEC-O2-ColV	ExPEC	A2363	184,501	FIB, FIIA	Aerobactin, <i>sit, hlyF, ompT, ets, iss, salmochelin</i> , ColV, <i>tsh</i>	AY545598
pVir68	NTEC-2	Vir68	138,362	FIB, FIIA	<i>cnf2, cdt-III, hlyABC</i>	CP001162
pSB4_227	<i>Shigella boydii</i>	Sb227	126,697	FIIA	Inv plasmid lacking TTSS	CP000037
pSF5	<i>S. flexneri</i>	Serotype 5	136,694	FIIA	Inv plasmid lacking TTSS	AY879342
pSD1_197	<i>Shigella dysenteriae</i>	Sd197	182,726	FIIA	Inv plasmid with TTSS, lacking some components	CP000035
pBS512_211	<i>S. boydii</i>	CDC3083-94	210,919	FIIA	Inv plasmid with TTSS, lacking some components	CP001062
pSs046	<i>Shigella sonnei</i>	Ss046	214,396	FIIA	Inv plasmid with TTSS	CP000039
pCP301	<i>S. flexneri</i>	301	221,618	FIIA	Inv plasmid with TTSS	AF386526
pWR100 (pWR501)	<i>S. flexneri</i>	M90T	221,851	FIIA	Inv plasmid with TTSS	AF348706
p53638_226	EIEC	53638	225,683	FIIA	Backbone of Inv plasmid, lacking TTSS; has novel fimbrial operon and Pgt system	CP001064
pUTI89	UPEC	UTI89	114,230	FIB, FIIA	Tra	CP000244
p1ESCUM	UPEC	UMN026	122,301	FIB, FIIA	Tra	CU928148

<sup>a</sup> TTSS, type 3 secretion system.

genome database dedicated to the implementation of user-friendly comparative genomics tools (<http://www.ecoli.cvm.iastate.edu>). The modular and plastic nature of bacterial plasmids surely makes the study of their evolution more complicated than that of bacterial chromosomes. However, the future use of unique and multifactorial approaches (i.e., replicon typing, relaxase typing, multilocus sequence analysis, and analysis of whole-genome content and arrangement) for these problems will aid in our future ability to understand virulence plasmid evolution. The use of such improved tools will also enhance the consistency of plasmid nomenclature and annotation. This, combined with additional plasmid sequences and a refinement of plasmid-typing protocols, will undoubtedly enable a better understanding of the evolution of these dynamic elements.

#### ACKNOWLEDGMENTS

Preparation of this review was supported by the University of Minnesota College of Veterinary Medicine, the University of Minnesota Supercomputing Institute, and grant EFO062666 from the National Science Foundation.

We give special thanks to Richard Isaacson (University of Minnesota) for critique of this review.

#### REFERENCES

- Actis, L. A., M. E. Tolmasky, and J. H. Crosa. 1999. Bacterial plasmids: replication of extrachromosomal genetic elements encoding resistance to antimicrobial compounds. *Front. Biosci.* **4**:D43–D62.
- Aguero, M., G. de la Fuente, E. Vivaldi, and F. Cabello. 1989. ColV increases the virulence of *Escherichia coli* K1 strains in animal models of neonatal meningitis and urinary infection. *Med. Microbiol. Immunol.* **178**: 211–216.
- An, H., J. M. Fairbrother, C. Desautels, and J. Harel. 1999. Distribution of a novel locus called *paa* (porcine attaching and effacing associated) among enteric *Escherichia coli*. *Adv. Exp. Med. Biol.* **473**:179–184.
- Ansuini, A., P. Candotti, G. Vecchi, V. Falbo, F. Minelli, and A. Caprioli. 1994. Necrotogenic *E. coli* in rabbits and horses. *Vet. Res.* **134**:608.
- Bakker, D., C. E. Vader, B. Roosendaal, F. R. Mooi, B. Oudega, and F. K. de Graaf. 1991. Structure and function of periplasmic chaperone-like proteins involved in the biosynthesis of K88 and K99 fimbriae in enterotoxigenic *Escherichia coli*. *Mol. Microbiol.* **5**:875–886.
- Baldini, M. M., J. B. Kaper, M. M. Levine, D. C. Candy, and H. W. Moon. 1983. Plasmid-mediated adhesion in enteropathogenic *Escherichia coli*. *J. Pediatr. Gastroenterol. Nutr.* **2**:534–538.
- Baldini, M. M., J. B. Kaper, M. M. Levine, and H. W. Moon. 1983. Molecular nature of adhesion in enteropathogenic *Escherichia coli*. *Lancet* **ii**:218.
- Baudry, B., S. J. Savarino, P. Vial, J. B. Kaper, and M. M. Levine. 1990. A sensitive and specific DNA probe to identify enteroaggregative *Escherichia coli*, a recently discovered diarrheal pathogen. *J. Infect. Dis.* **161**:1249–1251.
- Bernier, C., P. Gounon, and C. Le Bouguenec. 2002. Identification of an aggregative adhesion fimbria (AAF) type III-encoding operon in enteroaggregative *Escherichia coli* as a sensitive probe for detecting the AAF-encoding operon family. *Infect. Immun.* **70**:4302–4311.
- Bertin, Y., C. Martin, E. Oswald, and J. P. Girardeau. 1996. Rapid and specific detection of F17-related pilin and adhesin genes in diarrheic and septicemic *Escherichia coli* strains by multiplex PCR. *J. Clin. Microbiol.* **34**:2921–2928.
- Binns, M., J. Mayden, and R. Levine. 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.
- Binns, M. M., D. L. Davies, and K. G. Hardy. 1979. Cloned fragments of the plasmid ColV, I-K94 specifying virulence and serum resistance. *Nature* **279**:778–781.
- Blanco, J. E., M. Blanco, and J. Blanco. 1995. Enterotoxigenic, verotoxigenic, and necrotogenic *Escherichia coli* in food and clinical samples. Role of animals as reservoirs of strains pathogenic for humans. *Microbiologia* **11**:97–110.
- Bodero, M. D., E. A. Harden, and G. P. Munson. 2008. Transcriptional regulation of subclass 5b fimbriae. *BMC Microbiol.* **8**:180.
- Boerlin, P., R. Travis, C. L. Gyles, R. Reid-Smith, N. Janecko, H. Lim, V. Nicholson, S. A. McEwen, R. Friendship, and M. Archambault. 2005. Antimicrobial resistance and virulence genes of *Escherichia coli* isolates from swine in Ontario. *Appl. Environ. Microbiol.* **71**:6753–6761.
- Boisen, N., C. Struve, F. Scheutz, K. A. Krogh, and J. P. Nataro. 2008. New adhesin of enteroaggregative *Escherichia coli* related to the Afa/Dra/AAF family. *Infect. Immun.* **76**:3281–3292.
- Bradley, D. E. 1985. Transfer systems of K88 and K99 plasmids. *Plasmid* **13**:118–128.
- Breitschwerdt, E. B., C. DebRoy, A. M. Mexas, T. T. Brown, and A. K. Remick. 2001. Isolation of necrotogenic *Escherichia coli* from a dog with hemorrhagic pneumonia. *J. Am. Vet. Med. Assoc.* **226**:2016–2019.
- Brinkley, C., V. Burland, R. Keller, D. J. Rose, A. T. Boutin, S. A. Klink, F. R. Blattner, and J. B. Kaper. 2006. Nucleotide sequence analysis of the enteropathogenic *Escherichia coli* adherence factor plasmid pMAR7. *Infect. Immun.* **74**:5408–5413.
- Brunder, W., H. Karch, and H. Schmidt. 2006. Complete sequence of the large virulence plasmid pSFO157 of the sorbitol-fermenting enterohemorrhagic *Escherichia coli* O157:H– strain 3072/96. *Int. J. Med. Microbiol.* **296**:467–474.
- Brunder, W., A. S. Khan, J. Hacker, and H. Karch. 2001. Novel type of fimbriae encoded by the large plasmid of sorbitol-fermenting enterohemorrhagic *Escherichia coli* O157:H–. *Infect. Immun.* **69**:4447–4457.
- Buchrieser, C., P. Glaser, C. Rusniok, H. Nedjari, H. D’Hauteville, F. Kunst, P. Sansonetti, and C. Parsot. 2000. The virulence plasmid pWR100 and the repertoire of proteins secreted by the type III secretion apparatus of *Shigella flexneri*. *Mol. Microbiol.* **38**:760–771.
- Burland, V., Y. Shao, N. T. Perna, G. Plunkett, H. J. Sofia, and F. R. Blattner. 1998. The complete DNA sequence and analysis of the large virulence plasmid of *Escherichia coli* O157:H7. *Nucleic Acids Res.* **26**:4196–4204.
- Canizalez-Roman, A., and F. Navarro-Garcia. 2003. Fodrin CaM-binding domain cleavage by Pet from enteroaggregative *Escherichia coli* leads to actin cytoskeletal disruption. *Mol. Microbiol.* **48**:947–958.
- Carattoli, A., A. Bertini, L. Villa, V. Falbo, K. L. Hopkins, and E. J. Threlfall. 2005. Identification of plasmids by PCR-based replicon typing. *J. Microbiol. Methods* **63**:219–228.
- Caron, J., L. M. Coffield, and J. R. Scott. 1989. A plasmid-encoded regulatory gene, *ms*, required for expression of the CS1 and CS2 adhesins of enterotoxigenic *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **86**:963–967.
- Casey, T. A., R. A. Schneider, and E. A. Dean-Nystrom. 1993. Identification of plasmid and chromosomal copies of 987P pilus genes in enterotoxigenic *Escherichia coli* 987. *Infect. Immun.* **61**:2249–2252.
- Chen, H. D., and G. Frankel. 2005. Enteropathogenic *Escherichia coli*: unravelling pathogenesis. *FEMS Microbiol. Rev.* **29**:83–98.
- Cheng, D., H. Sun, J. Xu, and S. Gao. 2005. Prevalence of fimbrial colonization factors F18ab and F18ac in *Escherichia coli* isolates from weaned piglets with edema and/or diarrhea in China. *Vet. Microbiol.* **110**:35–39.
- Cheng, D., H. Sun, J. Xu, and S. Gao. 2006. PCR detection of virulence factor genes in *Escherichia coli* isolates from weaned piglets with edema disease and/or diarrhea in China. *Vet. Microbiol.* **115**:320–328.
- Christenson, J. K., and D. M. Gordon. 2009. Evolution of colicin BM plasmids: the loss of the colicin B activity gene. *Microbiology* **155**:1645–1655.
- Clarke, S. C., R. D. Haigh, P. P. Freestone, and P. H. Williams. 2003. Virulence of enteropathogenic *Escherichia coli*, a global pathogen. *Clin. Microbiol. Rev.* **16**:365–378.
- Cohen, S. N., A. C. Chang, H. W. Boyer, and R. B. Helling. 1973. Construction of biologically functional bacterial plasmids in vitro. *Proc. Natl. Acad. Sci. USA* **70**:3240–3244.
- Couturier, M., F. Bex, P. L. Bergquist, and W. K. Maas. 1988. Identification and classification of bacterial plasmids. *Microbiol. Rev.* **52**:375–395.
- Czczulin, J. R., S. Balepur, S. Hicks, A. Phillips, R. Hall, M. H. Kothary, F. Navarro-Garcia, and J. P. Nataro. 1997. Aggregative adherence fimbria II, a second fimbrial antigen mediating aggregative adherence in enteroaggregative *Escherichia coli*. *Infect. Immun.* **65**:4135–4145.
- Czczulin, J. R., T. S. Whittam, I. R. Henderson, F. Navarro-Garcia, and J. P. Nataro. 1999. Phylogenetic analysis of enteroaggregative and diffusely adherent *Escherichia coli*. *Infect. Immun.* **67**:2692–2699.
- da Silva, A. S., G. F. Valadares, M. P. Penatti, B. G. Brito, and D. da Silva Leite. 2001. *Escherichia coli* strains from edema disease: O serogroups, and genes for Shiga toxin, enterotoxins, and F18 fimbriae. *Vet. Microbiol.* **80**:227–233.
- Dean, E. A., S. C. Whipp, and H. W. Moon. 1989. Age-specific colonization of porcine intestinal epithelium by 987P-piliated enterotoxigenic *Escherichia coli*. *Infect. Immun.* **57**:82–87.
- DebRoy, C., and C. W. Maddox. 2001. Identification of virulence attributes of gastrointestinal *Escherichia coli* isolates of veterinary significance. *Anim. Health Res. Rev.* **2**:129–140.
- De Rycke, J., A. Milon, and E. Oswald. 1999. Necrotic *Escherichia coli* (NTEC): two emerging categories of human and animal pathogens. *Vet. Res.* **30**:221–233.
- Do, N. T., H. P. Cu, N. N. Nguyen, X. H. Nguyen, X. T. Au, T. H. Van, N. Q. Vu, and D. J. Trott. 2006. Antimicrobial resistance phenotypes of ETEC isolates from piglets with diarrhea in North Vietnam. *Ann. N. Y. Acad. Sci.* **1081**:543–545.
- Dozois, C. M., F. Daigle, and R. Curtiss III. 2003. Identification of patho-

- gen-specific and conserved genes expressed in vivo by an avian pathogenic *Escherichia coli* strain. Proc. Natl. Acad. Sci. USA **100**:247–252.
43. Dozois, C. M., M. Dho-Moulin, A. Bree, J. M. Fairbrother, C. Desautels, and R. Curtiss III. 2000. Relationship between the Tsh autotransporter and pathogenicity of avian *Escherichia coli* and localization and analysis of the *tsh* genetic region. Infect. Immun. **68**:4145–4154.
  44. el Mazouari, K., E. Oswald, J. P. Hernalsteens, P. Lintermans, and H. De Greve. 1994. F17-like fimbriae from an invasive *Escherichia coli* strain producing cytotoxic necrotizing factor type 2 toxin. Infect. Immun. **62**:2633–2638.
  45. Eslava, C., F. Navarro-García, J. R. Czczulin, I. R. Henderson, A. Cravioto, and J. P. Nataro. 1998. Pet, an autotransporter enterotoxin from enteroaggregative *Escherichia coli*. Infect. Immun. **66**:3155–3163.
  46. Evans, D. G., and D. J. Evans, Jr. 1978. New surface-associated heat-labile colonization factor antigen (CFA/II) produced by enterotoxigenic *Escherichia coli* of serogroups O6 and O8. Infect. Immun. **21**:638–647.
  47. Evans, D. G., R. P. Silver, D. J. Evans, Jr., D. G. Chase, and S. L. Gorbach. 1975. Plasmid-controlled colonization factor associated with virulence in *Escherichia coli* enterotoxigenic for humans. Infect. Immun. **12**:656–667.
  48. Evans, D. J., Jr., and D. G. Evans. 1973. Three characteristics associated with enterotoxigenic *Escherichia coli* isolated from man. Infect. Immun. **8**:322–328.
  49. Fairbrother, J. M., E. Nadeau, and C. L. Gyles. 2005. *Escherichia coli* in postweaning diarrhea in pigs: an update on bacterial types, pathogenesis, and prevention strategies. Anim. Health Res. Rev. **6**:17–39.
  50. Falkow, S. 1975. Infectious multiple drug resistance. Pion, London, United Kingdom.
  51. Fekete, P. Z., J. Gerardin, E. Jacquemin, J. G. Mainil, and B. Nagy. 2002. Replicon typing of F18 fimbriae encoding plasmids of enterotoxigenic and verotoxigenic *Escherichia coli* strains from porcine postweaning diarrhoea and oedema disease. Vet. Microbiol. **85**:275–284.
  52. Fekete, P. Z., G. Schneider, F. Olasz, G. Blum-Oehler, J. H. Hacker, and B. Nagy. 2003. Detection of a plasmid-encoded pathogenicity island in F18+ enterotoxigenic and verotoxigenic *Escherichia coli* from weaned pigs. Int. J. Med. Microbiol. **293**:287–298.
  53. Feng, P., K. A. Lampel, H. Karch, and T. S. Whittam. 1998. Genotypic and phenotypic changes in the emergence of *Escherichia coli* O157:H7. J. Infect. Dis. **177**:1750–1753.
  54. Fleckenstein, J. M., K. Roy, J. F. Fischer, and M. Burkitt. 2006. Identification of a two-partner secretion locus of enterotoxigenic *Escherichia coli*. Infect. Immun. **74**:2245–2258.
  55. Francia, M. V., A. Varsaki, M. P. Garcillan-Barcia, A. Latorre, C. Drainas, and F. de la Cruz. 2004. A classification scheme for mobilization regions of bacterial plasmids. FEMS Microbiol. Rev. **28**:79–100.
  56. Fredericq, P., and G. Cornelis. 1978. Formation of a cointegrate between phage phi80 and Col, Trp plasmid. J. Gen. Microbiol. **105**:347–349.
  57. Fredericq, P., and E. Dehalle. 1969. Recombination of R factors with Col factors and chromosomal genes in *Escherichia coli*. J. Gen. Microbiol. **55**:xi–xii.
  58. Fredericq, P., and E. Delhalle. 1972. Recombinations between R and colicinogenic factors in *Escherichia coli* K12. I. Production of recombinants. Ann. Inst. Pasteur (Paris) **122**:909–921.
  59. Fredericq, P., V. Kromeyer, and M. Kettner. 1971. Transferable colicinogenic factors as mobilizing agents for extrachromosomal streptomycin resistance. Z. Allg. Mikrobiol. **11**:11–17. (In German.)
  60. Fricke, W. F., P. F. McDermott, M. K. Mammel, S. Zhao, T. J. Johnson, D. A. Rasko, P. J. Fedorka-Cray, A. Pedrosa, J. M. Whichard, J. E. Leclerc, D. G. White, T. A. Cebular, and J. Ravel. 2009. Antimicrobial resistance-conferring plasmids with similarity to virulence plasmids from avian pathogenic *Escherichia coli* strains in *Salmonella enterica* serovar Kentucky isolates from poultry. Appl. Environ. Microbiol. **75**:5963–5971.
  61. Fricke, W. F., M. S. Wright, A. H. Lindell, D. M. Harkins, C. Baker-Austin, J. Ravel, and R. Stepanauskas. 2008. Insights into the environmental resistance gene pool from the genome sequence of the multidrug-resistant environmental isolate *Escherichia coli* SMS-3-5. J. Bacteriol. **190**:6779–6794.
  62. Froehlich, B., J. Parkhill, M. Sanders, M. A. Quail, and J. R. Scott. 2005. The pCoo plasmid of enterotoxigenic *Escherichia coli* is a mosaic cointegrate. J. Bacteriol. **187**:6509–6516.
  63. Frost, L. S., R. Leplae, A. O. Summers, and A. Toussaint. 2005. Mobile genetic elements: the agents of open source evolution. Nat. Rev. Microbiol. **3**:722–732.
  - 63a. Funnell, B. E., and G. J. Phillips (ed.). 2004. Plasmid biology. ASM Press, Washington, DC.
  64. Gaastra, W., and A. M. Svennerholm. 1996. Colonization factors of human enterotoxigenic *Escherichia coli* (ETEC). Trends Microbiol. **4**:444–452.
  65. Garcillan-Barcia, M. P., M. V. Francia, and F. de la Cruz. 2009. The diversity of conjugative relaxases and its application in plasmid classification. FEMS Microbiol. Rev. **33**:657–687.
  66. Goswami, P. S., C. L. Gyles, R. M. Friendship, C. Poppe, G. K. Kozak, and P. Boerlin. 2008. Effect of plasmid pTENT2 on severity of porcine postweaning diarrhoea induced by an O149 enterotoxigenic *Escherichia coli*. Vet. Microbiol. **131**:400–405.
  67. Gotz, A., R. Pukall, E. Smit, E. Tietze, R. Prager, H. Tschape, J. D. van Elsas, and K. Smalla. 1996. Detection and characterization of broad-host-range plasmids in environmental bacteria by PCR. Appl. Environ. Microbiol. **62**:2621–2628.
  68. Gratia, A. 1925. Sur un remarquable exemple d'antagisme entre deux souches de colibacille. C. R. Soc. Biol. **93**:1041.
  69. Hantke, K., G. Nicholson, W. Rabsch, and G. Winkelmann. 2003. Salmochelins, siderophores of *Salmonella enterica* and uropathogenic *Escherichia coli* strains, are recognized by the outer membrane receptor IroN. Proc. Natl. Acad. Sci. USA **100**:3677–3682.
  70. Harnett, N. M., and C. L. Gyles. 1983. Enterotoxigenicity of bovine and porcine *Escherichia coli* of O groups 8, 9, 20, 64, 101, and X46. Am. J. Vet. Res. **44**:1210–1214.
  71. Harnett, N. M., and C. L. Gyles. 1985. Enterotoxin plasmids in bovine and porcine enterotoxigenic *Escherichia coli* of O groups 9, 20, 64 and 101. Can. J. Comp. Med. **49**:79–87.
  72. Harrington, S. M., E. G. Dudley, and J. P. Nataro. 2006. Pathogenesis of enteroaggregative *Escherichia coli* infection. FEMS Microbiol. Lett. **254**:12–18.
  73. Helgersson, A. F., V. Sharma, A. M. Dow, R. Schroeder, K. Post, and N. A. Cornick. 2006. Edema disease caused by a clone of *Escherichia coli* O147. J. Clin. Microbiol. **44**:3074–3077.
  74. Hicks, S. J., and R. J. Rowbury. 1986. Virulence plasmid-associated adhesion of *Escherichia coli* and its significance for chlorine resistance. J. Appl. Bacteriol. **61**:209–218.
  75. Hinnebusch, B. J., M. L. Rosso, T. G. Schwan, and E. Carniel. 2002. High-frequency conjugative transfer of antibiotic resistance genes to *Yersinia pestis* in the flea midgut. Mol. Microbiol. **46**:349–354.
  76. Hu, Y. 1992. The effect of plasmids on the resistance of *E. coli* to phages. Wei Sheng Wu Xue Bao **32**:456–458. (In Chinese.)
  77. Huang, D. B., A. Mohanty, H. L. DuPont, P. C. Okhuysen, and T. Chiang. 2006. A review of an emerging enteric pathogen: enteroaggregative *Escherichia coli*. J. Med. Microbiol. **55**:1303–1311.
  78. Huisman, T. T., D. Bakker, P. Klaassen, and F. K. de Graaf. 1994. Leucine-responsive regulatory protein, IS1 insertions, and the negative regulator FaeA control the expression of the *fae* (K88) operon in *Escherichia coli*. Mol. Microbiol. **11**:525–536.
  79. Iguchi, A., N. R. Thomson, Y. Ogura, D. Saunders, T. Ooka, I. R. Henderson, D. Harris, M. Asadulghani, K. Kurokawa, P. Dean, B. Kenny, M. A. Quail, S. Thurston, G. Dougan, T. Hayashi, J. Parkhill, and G. Frankel. 2009. Complete genome sequence and comparative genome analysis of enteropathogenic *Escherichia coli* O127:H6 strain E2348/69. J. Bacteriol. **191**:347–354.
  80. Isaacson, R. E., B. Nagy, and H. W. Moon. 1977. Colonization of porcine small intestine by *Escherichia coli*: colonization and adhesion factors of pig enteropathogens that lack K88. J. Infect. Dis. **135**:531–539.
  81. Isaacson, R. E., and G. L. Start. 1992. Analysis of K99 plasmids from enterotoxigenic *Escherichia coli*. FEMS Microbiol. Lett. **69**:141–146.
  82. Jacob, F., and S. Brenner. 1963. On the regulation of DNA synthesis in bacteria: the hypothesis of the replicon. C. R. Hebd. Seances Acad. Sci. **256**:298–300. (In French.)
  83. Jenkins, C., C. van Ijperen, E. G. Dudley, H. Chart, G. A. Willshaw, T. Cheasty, H. R. Smith, and J. P. Nataro. 2005. Use of a microarray to assess the distribution of plasmid and chromosomal virulence genes in strains of enteroaggregative *Escherichia coli*. FEMS Microbiol. Lett. **253**:119–124.
  84. Johnson, J. R., and T. A. Russo. 2002. Extraintestinal pathogenic *Escherichia coli*: “the other bad E coli.” J. Lab. Clin. Med. **139**:155–162.
  85. Johnson, J. R., and T. A. Russo. 2005. Molecular epidemiology of extraintestinal pathogenic (uropathogenic) *Escherichia coli*. Int. J. Med. Microbiol. **295**:383–404.
  86. Johnson, T. J., C. W. Giddings, S. M. Horne, P. S. Gibbs, R. E. Wooley, J. Skyberg, P. Olah, R. Kercher, J. S. Sherwood, S. L. Foley, and L. K. Nolan. 2002. Location of increased serum survival gene and selected virulence traits on a conjugative R plasmid in an avian *Escherichia coli* isolate. Avian Dis. **46**:342–352.
  87. Johnson, T. J., S. J. Johnson, and L. K. Nolan. 2006. Complete DNA sequence of a ColBM plasmid from avian pathogenic *Escherichia coli* suggests that it evolved from closely related ColV virulence plasmids. J. Bacteriol. **188**:5975–5983.
  88. Johnson, T. J., K. E. Siek, S. J. Johnson, and L. K. Nolan. 2006. DNA sequence of a ColV plasmid and prevalence of selected plasmid-encoded virulence genes among avian *Escherichia coli* strains. J. Bacteriol. **188**:745–758.
  89. Johnson, T. J., Y. M. Wannemuehler, S. J. Johnson, C. M. Logue, D. G. White, C. Doetkott, and L. K. Nolan. 2007. Plasmid replicon typing of commensal and pathogenic *Escherichia coli* isolates. Appl. Environ. Microbiol. **73**:1976–1983.
  90. Kaper, J. B. 2005. Pathogenic *Escherichia coli*. Int. J. Med. Microbiol. **295**:355–356.

91. Kaper, J. B., J. P. Nataro, and H. L. Mobley. 2004. Pathogenic *Escherichia coli*. *Nat. Rev. Microbiol.* **2**:123–140.
92. Kim, S. R., and T. Komano. 1997. The plasmid R64 thin pilus identified as a type IV pilus. *J. Bacteriol.* **179**:3594–3603.
93. Knutton, S., M. M. Baldini, J. B. Kaper, and A. S. McNeish. 1987. Role of plasmid-encoded adherence factors in adhesion of enteropathogenic *Escherichia coli* to HEp-2 cells. *Infect. Immun.* **55**:78–85.
94. Komano, T., T. Yoshida, K. Narahara, and N. Furuya. 2000. The transfer region of Inc11 plasmid R64: similarities between R64 tra and *Legionella icm/dot* genes. *Mol. Microbiol.* **35**:1348–1359.
95. Kotloff, K. L., J. P. Winickoff, B. Ivanoff, J. D. Clemens, D. L. Swerdlow, P. J. Sansonetti, G. K. Adak, and M. M. Levine. 1999. Global burden of *Shigella* infections: implications for vaccine development and implementation of control strategies. *Bull. W. H. O.* **77**:651–666.
96. Krcmery, V., P. Fredericq, B. Wiedemann, and C. Hurwitz. 1971. Mobilization of extrachromosomal determinants for streptomycin resistance by transferable colicinogenic factors. *Ann. N. Y. Acad. Sci.* **182**:118–122.
97. Kuehn, M. J., D. J. Ogg, J. Kihlberg, L. N. Slonim, K. Flemmer, T. Bergfors, and S. J. Hultgren. 1993. Structural basis of pilus subunit recognition by the PapD chaperone. *Science* **262**:1234–1241.
98. Lan, R., B. Lumb, D. Ryan, and P. R. Reeves. 2001. Molecular evolution of large virulence plasmid in *Shigella* clones and enteroinvasive *Escherichia coli*. *Infect. Immun.* **69**:6303–6309.
99. Lederberg, J., and E. L. Tatum. 1953. Sex in bacteria; genetic studies, 1945–1952. *Science* **118**:169–175.
100. Le Gall, T., M. Mavris, M. C. Martino, M. L. Bernardini, E. Denamur, and C. Parsot. 2005. Analysis of virulence plasmid gene expression defines three classes of effectors in the type III secretion system of *Shigella flexneri*. *Microbiology* **151**:951–962.
101. Levine, M. M., J. P. Nataro, H. Karch, M. M. Baldini, J. B. Kaper, R. E. Black, M. L. Clements, and A. D. O'Brien. 1985. The diarrheal response of humans to some classic serotypes of enteropathogenic *Escherichia coli* is dependent on a plasmid encoding an enteroadhesiveness factor. *J. Infect. Dis.* **152**:550–559.
102. Levine, M. M., J. G. Xu, J. B. Kaper, H. Lior, V. Prado, B. Tall, J. Nataro, H. Karch, and K. Wachsmuth. 1987. A DNA probe to identify enterohemorrhagic *Escherichia coli* O157:H7 and other serotypes that cause hemorrhagic colitis and hemolytic uremic syndrome. *J. Infect. Dis.* **156**:175–182.
103. Liu, J., P. Keelan, P. M. Bennett, and V. I. Enne. 2009. Characterization of a novel macrolide efflux gene, *mef(B)*, found linked to *sul3* in porcine *Escherichia coli*. *J. Antimicrob. Chemother.* **63**:423–426.
104. Lopez-Alvarez, J., and C. L. Gyles. 1980. Occurrence of the Vir plasmid among animal and human strains of invasive *Escherichia coli*. *Am. J. Vet. Res.* **41**:769–774.
105. Lopez-Alvarez, J., C. L. Gyles, P. L. Shipley, and S. Falkow. 1980. Genetic and molecular characteristics of Vir plasmids of bovine septicemic *Escherichia coli*. *J. Bacteriol.* **141**:758–769.
106. Mahillon, J., and M. Chandler. 1998. Insertion sequences. *Microbiol. Mol. Biol. Rev.* **62**:725–774.
107. Mainil, J. G., F. Bex, P. Dreze, A. Kaackenbeeck, and M. Couturier. 1992. Replicon typing of virulence plasmids of enterotoxigenic *Escherichia coli* isolates from cattle. *Infect. Immun.* **60**:3376–3380.
108. Mainil, J. G., G. Daube, E. Jacquemin, P. Pohl, and A. Kaackenbeeck. 1998. Virulence plasmids of enterotoxigenic *Escherichia coli* isolates from piglets. *Vet. Microbiol.* **62**:291–301.
109. Mainil, J. G., J. Gerardin, and E. Jacquemin. 2000. Identification of the F17 fimbrial subunit- and adhesin-encoding (*f17A* and *f17G*) gene variants in necrotizing *Escherichia coli* from cattle, pigs and humans. *Vet. Microbiol.* **73**:327–335.
110. Mainil, J. G., E. Jacquemin, P. Pohl, A. Kaackenbeeck, and I. Benz. 2002. DNA sequences coding for the F18 fimbriae and AIDA adhesin are localised on the same plasmid in *Escherichia coli* isolates from piglets. *Vet. Microbiol.* **86**:303–311.
111. Makino, K., K. Ishii, T. Yasunaga, M. Hattori, K. Yokoyama, C. H. Yut-sudo, Y. Kubota, Y. Yamaichi, T. Iida, K. Yamamoto, T. Honda, C. G. Han, E. Ohtsubo, M. Kasamatsu, T. Hayashi, S. Kuhara, and H. Shinagawa. 1998. Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an enterohemorrhagic *Escherichia coli* O157:H7 derived from Sakai outbreak. *DNA Res.* **5**:1–9.
112. Martinez-Abarca, F., and N. Toro. 2000. Group II introns in the bacterial world. *Mol. Microbiol.* **38**:917–926.
113. Massie, B., D. Dea, and A. Sasarman. 1984. A detailed physical and genetic map of two R. ColBM IncFIII plasmids. *Plasmid* **11**:105–108.
114. Maynard, C., J. M. Fairbrother, S. Bekal, F. Sanschagrin, R. C. Levesque, R. Brousseau, L. Masson, S. Lariviere, and J. Harel. 2003. Antimicrobial resistance genes in enterotoxigenic *Escherichia coli* O149:K91 isolates obtained over a 23-year period from pigs. *Antimicrob. Agents Chemother.* **47**:3214–3221.
115. McConnell, M. M., H. Chart, S. M. Scotland, H. R. Smith, G. A. Willshaw, and B. Rowe. 1989. Properties of adherence factor plasmids of enteropathogenic *Escherichia coli* and the effect of host strain on expression of adherence to HEp-2 cells. *J. Gen. Microbiol.* **135**:1123–1134.
116. McDaniel, T. K., K. G. Jarvis, M. S. Donnenberg, and J. B. Kaper. 1995. A genetic locus of enterocyte effacement conserved among diverse enterobacterial pathogens. *Proc. Natl. Acad. Sci. USA* **92**:1664–1668.
117. Mellata, M., J. W. Touchman, and R. Curtiss. 2009. Full sequence and comparative analysis of the plasmid pAPEC-1 of avian pathogenic *E. coli* chi7122 (O78:K80:H9). *PLoS ONE* **4**:e4232.
118. Morales, C., M. D. Lee, C. Hofacre, and J. J. Maurer. 2004. Detection of a novel virulence gene and a Salmonella virulence homologue among *Escherichia coli* isolated from broiler chickens. *Foodborne Pathog. Dis.* **1**:160–165.
119. Nagel De Zwaig, R. 1966. Association between colicinogenic and fertility factors. *Genetics* **54**:381–390.
120. Nagy, B., and P. Z. Fekete. 2005. Enterotoxigenic *Escherichia coli* in veterinary medicine. *Int. J. Med. Microbiol.* **295**:443–454.
121. Nagy, B., H. W. Moon, and R. E. Isaacson. 1976. Colonization of porcine small intestine by *Escherichia coli*: ileal colonization and adhesion by pig enteropathogens that lack K88 antigen and by some acapsular mutants. *Infect. Immun.* **13**:1214–1220.
122. Nagy, B., H. W. Moon, and R. E. Isaacson. 1977. Colonization of porcine intestine by enterotoxigenic *Escherichia coli*: selection of pilated forms in vivo, adhesion of pilated forms to epithelial cells in vitro, and incidence of a pilus antigen among porcine enteropathogenic *E. coli*. *Infect. Immun.* **16**:344–352.
123. Nagy, B., S. C. Whipp, H. Imberechts, H. U. Bertschinger, E. A. Dean-Nystrom, T. A. Casey, and E. Salajka. 1997. Biological relationship between F18ab and F18ac fimbriae of enterotoxigenic and verotoxigenic *Escherichia coli* from weaned pigs with oedema disease or diarrhoea. *Microb. Pathog.* **22**:1–11.
124. Nakaya, R., A. Nakamura, and Y. Murata. 1960. Resistance transfer agents in *Shigella*. *Biochem. Biophys. Res. Commun.* **3**:654–659.
125. Nataro, J. P. 2005. Enteroaggregative *Escherichia coli* pathogenesis. *Curr. Opin. Gastroenterol.* **21**:4–8.
126. Nataro, J. P., Y. Deng, S. Cookson, A. Cravioto, S. J. Savarino, L. D. Guers, M. M. Levine, and C. O. Tacket. 1995. Heterogeneity of enteroaggregative *Escherichia coli* virulence demonstrated in volunteers. *J. Infect. Dis.* **171**:465–468.
127. Nataro, J. P., Y. Deng, D. R. Maneval, A. L. German, W. C. Martin, and M. M. Levine. 1992. Aggregative adherence fimbriae I of enteroaggregative *Escherichia coli* mediate adherence to HEp-2 cells and hemagglutination of human erythrocytes. *Infect. Immun.* **60**:2297–2304.
128. Nataro, J. P., and J. B. Kaper. 1998. Diarrheagenic *Escherichia coli*. *Clin. Microbiol. Rev.* **11**:142–201.
129. Nataro, J. P., J. B. Kaper, R. Robins-Browne, V. Prado, P. Vial, and M. M. Levine. 1987. Patterns of adherence of diarrheagenic *Escherichia coli* to HEp-2 cells. *Pediatr. Infect. Dis. J.* **6**:829–831.
130. Nataro, J. P., I. C. Scaletsky, J. B. Kaper, M. M. Levine, and L. R. Trabulsi. 1985. Plasmid-mediated factors conferring diffuse and localized adherence of enteropathogenic *Escherichia coli*. *Infect. Immun.* **48**:378–383.
131. Nataro, J. P., D. Yikang, D. Yingkan, and K. Walker. 1994. AggR, a transcriptional activator of aggregative adherence fimbria I expression in enteroaggregative *Escherichia coli*. *J. Bacteriol.* **176**:4691–4699.
132. Neter, E., O. Westpahl, O. Luderitz, R. M. Gino, and E. A. Gorzynski. 1955. Demonstration of antibodies against enteropathogenic *Escherichia coli* in sera of children of various ages. *Pediatrics* **16**:801–808.
133. Newton, H. J., J. Sloan, D. M. Bulach, T. Seemann, C. C. Allison, M. Tauschek, R. M. Robins-Browne, J. C. Paton, T. S. Whittam, A. W. Paton, and E. L. Hartland. 2009. Shiga toxin-producing *Escherichia coli* strains negative for locus of enterocyte effacement. *Emerg. Infect. Dis.* **15**:372–380.
134. Olasz, F., P. Z. Fekete, G. Blum-Oehler, Z. Boldogkoi, and B. Nagy. 2005. Characterization of an F18+ enterotoxigenic *Escherichia coli* strain from post weaning diarrhoea of swine, and of its conjugative virulence plasmid pTC. *FEMS Microbiol. Lett.* **244**:281–289.
135. Orden, J. A., G. Dominguez-Bernal, S. Martinez-Pulgarin, M. Blanco, J. E. Blanco, A. Mora, J. Blanco, J. Blanco, and R. de la Fuente. 2007. Necrotoxic *Escherichia coli* from sheep and goats produce a new type of cytotoxic necrotizing factor (CNF3) associated with the *eae* and *ehxA* genes. *Int. Microbiol.* **10**:47–55.
136. Orskov, I., and F. Orskov. 1966. Episome-carried surface antigen K88 of *Escherichia coli*. I. Transmission of the determinant of the K88 antigen and influence on the transfer of chromosomal markers. *J. Bacteriol.* **91**:69–75.
137. Orskov, I., F. Orskov, W. J. Sojka, and J. M. Leach. 1961. Simultaneous occurrence of *E. coli* B and L antigens in strains from diseased swine. Influence of cultivation temperature. Two new *E. coli* K antigens: K87 and K88. *Acta Pathol. Microbiol. Scand.* **53**:404–422.
138. Ostroff, S. M., P. I. Tarr, M. A. Neill, J. H. Lewis, N. Hargrett-Bean, and J. M. Kobayashi. 1989. Toxin genotypes and plasmid profiles as determinants of systemic sequelae in *Escherichia coli* O157:H7 infections. *J. Infect. Dis.* **160**:994–998.
139. Oswald, E., M. Sugai, A. Labigne, H. C. Wu, C. Fiorentini, P. Boquet, and A. D. O'Brien. 1994. Cytotoxic necrotizing factor type 2 produced by viru-

- lent *Escherichia coli* modifies the small GTP-binding proteins Rho involved in assembly of actin stress fibers. *Proc. Natl. Acad. Sci. USA* **91**:3814–3818.
140. **Parsot, C.** 2005. *Shigella* spp. and enteroinvasive *Escherichia coli* pathogenicity factors. *FEMS Microbiol. Lett.* **252**:11–18.
  141. **Patel, S. K., J. Dotson, K. P. Allen, and J. M. Fleckenstein.** 2004. Identification and molecular characterization of EatA, an autotransporter protein of enterotoxigenic *Escherichia coli*. *Infect. Immun.* **72**:1786–1794.
  142. **Peres, S. Y., O. Marches, F. Daigle, J. P. Nougayrede, F. Herault, C. Tasca, J. De Rycke, and E. Oswald.** 1997. A new cytolethal distending toxin (CDT) from *Escherichia coli* producing CNF2 blocks HeLa cell division in G2/M phase. *Mol. Microbiol.* **24**:1095–1107.
  143. **Perez-Casal, J., J. S. Swartley, and J. R. Scott.** 1990. Gene encoding the major subunit of CS1 pili of human enterotoxigenic *Escherichia coli*. *Infect. Immun.* **58**:3594–3600.
  144. Reference deleted.
  145. **Picken, R. N., A. J. Mazaitis, S. Saadi, and W. K. Maas.** 1984. Characterization of the basic replicons of the chimeric R/Ent plasmid pCG86 and the related Ent plasmid P307. *Plasmid* **12**:10–18.
  146. **Provence, D. L., and R. Curtiss III.** 1994. Isolation and characterization of a gene involved in hemagglutination by an avian pathogenic *Escherichia coli* strain. *Infect. Immun.* **62**:1369–1380.
  147. **Qadri, F., A. M. Svennerholm, A. S. Faruque, and R. B. Sack.** 2005. Enterotoxigenic *Escherichia coli* in developing countries: epidemiology, microbiology, clinical features, treatment, and prevention. *Clin. Microbiol. Rev.* **18**:465–483.
  148. **Quackenbush, R., and S. Falkow.** 1979. Relationship between colicin V activity and virulence in *Escherichia coli*. *Infect. Immun.* **24**:562–564.
  149. **Raja, N., M. Goodson, W. C. Chui, D. G. Smith, and R. J. Rowbury.** 1991. Habituation to acid in *Escherichia coli*: conditions for habituation and its effects on plasmid transfer. *J. Appl. Bacteriol.* **70**:59–65.
  150. **Rasko, D. A., M. J. Rosovitz, G. S. Myers, E. F. Mongodin, W. F. Fricke, P. Gajer, J. Crabtree, M. Sebahia, N. R. Thomson, R. Chaudhuri, I. R. Henderson, V. Sperandio, and J. Ravel.** 2008. The pangenome structure of *Escherichia coli*: comparative genomic analysis of *E. coli* commensal and pathogenic isolates. *J. Bacteriol.* **190**:6881–6893.
  151. **Ratnam, S., S. B. March, R. Ahmed, G. S. Bezanson, and S. Kasatiya.** 1988. Characterization of *Escherichia coli* serotype O157:H7. *J. Clin. Microbiol.* **26**:2006–2012.
  152. **Ravi, M., M. Ngeleka, S. H. Kim, C. Gyles, F. Berthiaume, M. Mourez, D. Middleton, and E. Simko.** 2007. Contribution of AIDA-I to the pathogenicity of a porcine diarrheagenic *Escherichia coli* and to intestinal colonization through biofilm formation in pigs. *Vet. Microbiol.* **120**:308–319.
  153. **Reis, M. H., M. Heloiza, T. Afonso, L. R. Trabulsi, A. J. Mazaitis, R. Maas, and W. K. Maas.** 1980. Transfer of a CFA/I-ST plasmid promoted by a conjugative plasmid in a strain of *Escherichia coli* of serotype O128ac:H12. *Infect. Immun.* **29**:140–143.
  154. **Roosendaal, B., and F. K. de Graaf.** 1989. The nucleotide sequence of the *fanD* gene encoding the large outer membrane protein involved in the biosynthesis of K99 fimbriae. *Nucleic Acids Res.* **17**:1263.
  155. **Rychlik, I., D. Gregorova, and H. Hradecka.** 2006. Distribution and function of plasmids in *Salmonella enterica*. *Vet. Microbiol.* **112**:1–10.
  156. **Saadi, S., W. K. Maas, D. F. Hill, and P. L. Bergquist.** 1987. Nucleotide sequence analysis of RepFIC, a basic replicon present in IncFI plasmids P307 and F, and its relation to the RepA replicon of IncFII plasmids. *J. Bacteriol.* **169**:1836–1846.
  157. **Saitou, N., and M. Nei.** 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**:406–425.
  158. **Säsärman, A., B. Massie, M. Zollinger, H. Gagne-Tellier, F. Shareck, S. Garzon, and R. Morisset.** 1980. Naturally occurring R.ColBM plasmids belonging to the IncFIII incompatibility group. *J. Gen. Microbiol.* **119**:475–483.
  159. **Schifferli, D. M., E. H. Beachey, and R. K. Taylor.** 1990. The 987P fimbrial gene cluster of enterotoxigenic *Escherichia coli* is plasmid encoded. *Infect. Immun.* **58**:149–156.
  160. **Schifferli, D. M., E. H. Beachey, and R. K. Taylor.** 1991. Genetic analysis of 987P adhesion and fimbriation of *Escherichia coli*: the *fas* genes link both phenotypes. *J. Bacteriol.* **173**:1230–1240.
  161. **Sherlock, O., R. M. Vejborg, and P. Klemm.** 2005. The TibA adhesin/invasin from enterotoxigenic *Escherichia coli* is self recognizing and induces bacterial aggregation and biofilm formation. *Infect. Immun.* **73**:1954–1963.
  162. **Shiple, P. L., C. L. Gyles, and S. Falkow.** 1978. Characterization of plasmids that encode for the K88 colonization antigen. *Infect. Immun.* **20**:559–566.
  163. **Simons, B. L., P. Rathman, C. R. Malij, B. Oudega, and F. K. de Graaf.** 1990. The penultimate tyrosine residue of the K99 fibrillar subunit is essential for stability of the protein and its interaction with the periplasmic carrier protein. *FEMS Microbiol. Lett.* **55**:107–112.
  164. **Skyberg, J. A., T. J. Johnson, J. R. Johnson, C. Clabots, C. M. Logue, and L. K. Nolan.** 2006. Acquisition of avian pathogenic *Escherichia coli* plasmids by a commensal *E. coli* isolate enhances its abilities to kill chicken embryos, grow in human urine, and colonize the murine kidney. *Infect. Immun.* **74**:6287–6292.
  165. **Smajs, D., and G. M. Weinstock.** 2001. The iron- and temperature-regulated *cjyBC* genes of *Shigella* and enteroinvasive *Escherichia coli* strains code for colicin Js uptake. *J. Bacteriol.* **183**:3958–3966.
  166. **Smith, H. W.** 1974. A search for transmissible pathogenic characters in invasive strains of *Escherichia coli*: the discovery of a plasmid-controlled toxin and a plasmid-controlled lethal character closely associated, or identical, with colicine V. *J. Gen. Microbiol.* **83**:95–111.
  167. **Smith, H. W.** 1978. Transmissible pathogenic characteristics of invasive strains of *Escherichia coli*. *J. Am. Vet. Med. Assoc.* **173**:601–607.
  168. **Smith, H. W., and M. B. Huggins.** 1976. Further observations on the association of the colicine V plasmid of *Escherichia coli* with pathogenicity and with survival in the alimentary tract. *J. Gen. Microbiol.* **92**:335–350.
  169. **Smith, H. W., and M. A. Linggood.** 1971. Observations on the pathogenic properties of the K88, Hly and Ent plasmids of *Escherichia coli* with particular reference to porcine diarrhoea. *J. Med. Microbiol.* **4**:467–485.
  170. **Smith, H. W., and Z. Parsell.** 1975. Transmissible substrate-utilizing ability in enterobacteria. *J. Gen. Microbiol.* **87**:129–140.
  171. **Smith, J. L., P. M. Fratamico, and N. W. Gunther.** 2007. Extraintestinal pathogenic *Escherichia coli*. *Foodborne Pathog. Dis.* **4**:134–163.
  172. **Smyth, C. J.** 1982. Two mannose-resistant haemagglutinins on enterotoxigenic *Escherichia coli* of serotype O6:K15:H16 or H– isolated from travellers' and infantile diarrhoea. *J. Gen. Microbiol.* **128**:2081–2096.
  173. **Sobecky, P. A., T. J. Mincer, M. C. Chang, and D. R. Helinski.** 1997. Plasmids isolated from marine sediment microbial communities contain replication and incompatibility regions unrelated to those of known plasmid groups. *Appl. Environ. Microbiol.* **63**:888–895.
  174. **Starci, M., A. L. Stell, J. van der Goot, H. G. C. J. M. Hendriks, C. van Vorstenbosch, L. van Dijk, and W. Gastra.** 2002. Haemolytic *Escherichia coli* isolated from dogs with diarrhea have characteristics of both uropathogenic and necrotogenic strains. *Vet. Microbiol.* **85**:361–377.
  175. **Suzuki, H., M. Sota, C. J. Brown, and E. M. Top.** 2008. Using Mahalanobis distance to compare genomic signatures between bacterial plasmids and chromosomes. *Nucleic Acids Res.* **36**:e147.
  176. **Szczepanowski, R., S. Braun, V. Riedel, S. Schneiker, I. Krahn, A. Puhler, and A. Schluter.** 2005. The 120 592 bp IncF plasmid pRSB107 isolated from a sewage-treatment plant encodes nine different antibiotic-resistance determinants, two iron-acquisition systems and other putative virulence-associated functions. *Microbiology* **151**:1095–1111.
  177. **Tacket, C. O., G. Losonsky, H. Link, Y. Hoang, P. Guesry, H. Hilpert, and M. M. Levine.** 1988. Protection by milk immunoglobulin concentrate against oral challenge with enterotoxigenic *Escherichia coli*. *N. Engl. J. Med.* **318**:1240–1243.
  178. **Tamura, K., J. Dudley, M. Nei, and S. Kumar.** 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* **24**:1596–1599.
  179. **Tamura, K., M. Nei, and S. Kumar.** 2004. Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc. Natl. Acad. Sci. USA* **101**:11030–11035.
  180. **Tatum, E. L., and J. Lederberg.** 1947. Gene recombination in the bacterium *Escherichia coli*. *J. Bacteriol.* **53**:673–684.
  181. **Tivendale, K. A., A. H. Noormohammadi, J. L. Allen, and G. F. Browning.** 2009. The conserved portion of the putative virulence region contributes to virulence of avian pathogenic *Escherichia coli*. *Microbiology* **155**:450–460.
  182. **Tobe, T., T. Hayashi, C. G. Han, G. K. Schoolnik, E. Ohtsubo, and C. Sasakawa.** 1999. Complete DNA sequence and structural analysis of the enteropathogenic *Escherichia coli* adherence factor plasmid. *Infect. Immun.* **67**:5455–5462.
  183. **Touchon, M., C. Hoede, O. Tenailon, V. Barbe, S. Baeriswyl, P. Bidet, E. Bingen, S. Bonacorsi, C. Bouchier, O. Bouvet, A. Calteau, H. Chiapello, O. Clermont, S. Cruveiller, A. Danchin, M. Diard, C. Dossat, M. E. Karoui, E. Frapy, L. Garry, J. M. Ghigo, A. M. Gilles, J. Johnson, C. Le Bouguenec, M. Lescat, S. Mangenot, V. Martinez-Jehanne, I. Matic, X. Nassif, S. Oztas, M. A. Petit, C. Pichon, Z. Rouy, C. S. Ruf, D. Schneider, J. Tourret, B. Vacherie, D. Vallenet, C. Medigue, E. P. Rocha, and E. Denamur.** 2009. Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. *PLoS Genet.* **5**:e1000344.
  184. **Trabulsi, L. R., R. Keller, and T. A. Tardelli Gomes.** 2002. Typical and atypical enteropathogenic *Escherichia coli*. *Emerg. Infect. Dis.* **8**:508–513.
  185. **Turner, S. M., A. Scott-Tucker, L. M. Cooper, and I. R. Henderson.** 2006. Weapons of mass destruction: virulence factors of the global killer enterotoxigenic *Escherichia coli*. *FEMS Microbiol. Lett.* **263**:10–20.
  186. **Van der Auwera, G. A., J. E. Krol, H. Suzuki, B. Foster, R. Van Houdt, C. J. Brown, M. Mergeay, and E. M. Top.** 2009. Plasmids captured in *C. metallidurans* CH34: defining the PromA family of broad-host-range plasmids. *Antonie van Leeuwenhoek* **96**:193–204.
  187. **van der Woude, M. W., M. Braster, H. W. van Verseveld, and F. K. de Graaf.** 1990. Control of temperature-dependent synthesis of K99 fimbriae. *FEMS Microbiol. Lett.* **56**:183–188.
  188. **Villaseca, J. M., F. Navarro-Garcia, G. Mendoza-Hernandez, J. P. Nataro, A. Cravioto, and C. Eslava.** 2000. Pet toxin from enteroaggregative *Escherichia coli* produces cellular damage associated with fodrin disruption. *Infect. Immun.* **68**:5920–5927.



189. Warner, P. J., P. H. Williams, A. Bindereif, and J. B. Neilands. 1981. ColV plasmid-specific aerobactin synthesis by invasive strains of *Escherichia coli*. *Infect. Immun.* **33**:540–545.
190. Watanabe, T. 1963. Infective heredity of multiple drug resistance in bacteria. *Bacteriol. Rev.* **27**:87–115.
191. Watanabe, T., H. Nishida, C. Ogata, T. Arai, and S. Sato. 1964. Episome-mediated transfer of drug resistance in *Enterobacteriaceae*. VII. Two types of naturally occurring R factors. *J. Bacteriol.* **88**:716–726.
192. Waters, V. L., and J. H. Crosa. 1991. Colicin V virulence plasmids. *Microbiol. Rev.* **55**:437–450.
193. Wick, L. M., W. Qi, D. W. Lacher, and T. S. Whittam. 2005. Evolution of genomic content in the stepwise emergence of *Escherichia coli* O157:H7. *J. Bacteriol.* **187**:1783–1791.
194. Williams, P., and P. Warner. 1980. ColV plasmid-mediated, colicin V-independent iron uptake system of invasive strains of *Escherichia coli*. *Infect. Immun.* **29**:411–416.
195. Wolf, M. K. 1997. Occurrence, distribution, and associations of O and H serogroups, colonization factor antigens, and toxins of enterotoxigenic *Escherichia coli*. *Clin. Microbiol. Rev.* **10**:569–584.
196. Wooley, R. E., L. K. Nolan, J. Brown, P. S. Gibbs, C. W. Giddings, and K. S. Turner. 1993. Association of K-1 capsule, smooth lipopolysaccharides, *traT* gene, and colicin V production with complement resistance and virulence of avian *Escherichia coli*. *Avian Dis.* **37**:1092–1096.
197. Yoon, J. W., and C. J. Hovde. 2008. All blood, no stool: enterohemorrhagic *Escherichia coli* O157:H7 infection. *J. Vet. Sci.* **9**:219–231.
198. Zhang, W., W. Qi, T. J. Albert, A. S. Motiwala, D. Alland, E. K. Hyttia-Trees, E. M. Ribot, P. I. Fields, T. S. Whittam, and B. Swaminathan. 2006. Probing genomic diversity and evolution of *Escherichia coli* O157 by single nucleotide polymorphisms. *Genome Res.* **16**:757–767.
199. Zhou, D., W. D. Hardt, and J. E. Galan. 1999. *Salmonella typhimurium* encodes a putative iron transport system within the centisome 63 pathogenicity island. *Infect. Immun.* **67**:1974–1981.
200. Zienkiewicz, M., I. Kern-Zdanowicz, M. Golebiewski, J. Zylinska, P. Mieczkowski, M. Gniadkowski, J. Bardowski, and P. Ceglowski. 2007. Mosaic structure of p1658/97, a 125-kilobase plasmid harboring an active amplicon with the extended-spectrum beta-lactamase gene *bla*<sub>SHV-5</sub>. *Antimicrob. Agents Chemother.* **51**:1164–1171.

**Timothy J. Johnson** received his Ph.D. in 2004 from North Dakota State University in Fargo, North Dakota. His postdoctoral work was performed at the College of Veterinary Medicine at Iowa State University (2004 to 2007). Since 2007, he has worked as an Assistant Professor of Microbiology within the Department of Veterinary and Biomedical Sciences, College of Veterinary Medicine, University of Minnesota. Dr. Johnson's research interests involve genomics-based approaches towards understanding the evolution of virulence and antimicrobial resistance in enteric bacteria of production animals and humans. He is particularly interested in the mobile genetic elements contained by these organisms and the means by which they are disseminated among bacterial populations. His research also strives toward the development of means to control diseases of poultry, swine, and humans.



**Lisa K. Nolan** received D.V.M. and Ph.D. degrees from the University of Georgia. Her early career was spent at North Dakota State University, where she rose through the ranks from Assistant to Full Professor and served as the founding Director of the Great Plains Institute of Food Safety. In 2003, she became Professor and Chair of the Department of Veterinary Microbiology and Preventive Medicine in the College of Veterinary Medicine at Iowa State University and now serves as this college's Associate Dean of Research and Graduate Studies. The enduring focus of her research has been on the virulence mechanisms that extraintestinal pathogenic *Escherichia coli* (ExPEC) uses to cause disease in animal and human hosts, the role of plasmids in the pathogenesis of these diseases, and the relationship between ExPEC diseases in animals and those of human hosts.

