Supporting Information for:

Characterization of a large antibiotic resistance plasmid found in enteropathogenic *Escherichia coli* strain B171 and its relatedness to plasmids of diverse *E. coli*

Tracy H. Hazen[1,2], Jane Michalski[1], Sushma Nagaraj[1], Iruka N. Okeke[3], and David A. Rasko[1,2]*

[1]Institute for Genome Sciences, [2]Department of Microbiology and Immunology, University of Maryland School of Medicine, Baltimore, MD, 21201; [3] Department of Biology, Haverford College, Haverford, PA, USA and Faculty of Pharmacy, University of Ibadan, Ibadan, Nigeria

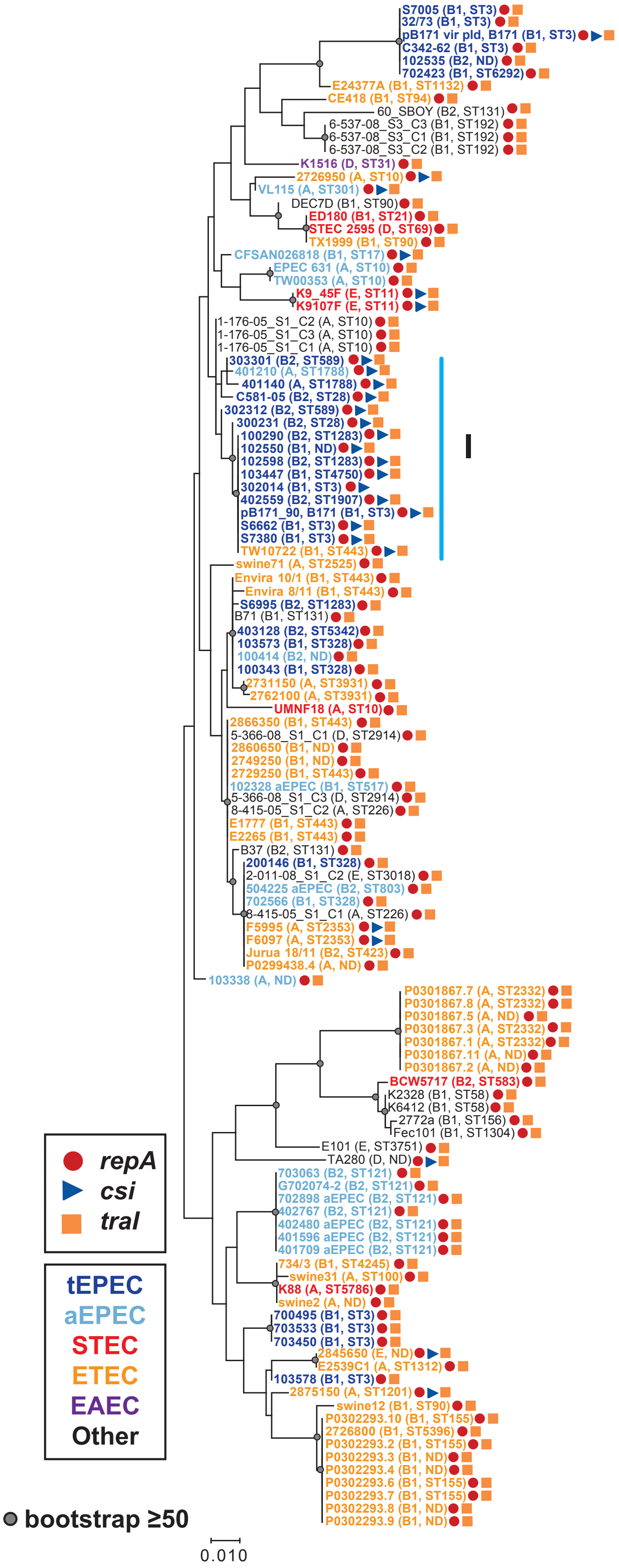
S7005 (B1, ST3)
32/73 (B1, ST3)
pB171 vir pld, B171 (B1, ST3)
C342-62 (B1, ST3)
102535 (B2, ND)
702423 (B1, ST6292)
E24377A (B1, ST1132)
CE418 (B1, ST94)
60_SBOY (B2, ST131)
6-537-08_S3_C3 (B1, ST192)
6-537-08_S3_C1 (B1, ST192)
6-537-08_S3_C2 (B1, ST192)
K1516 (D, ST31)
2726950 (A, ST10)
VL115 (A, ST301)
DEC7D (B1, ST90)
ED180 (B1, ST21)
STEC 2595 (D, ST69)
TX1999 (B1, ST90)
CFSAN026818 (B1, ST17)
EPEC 631 (A, ST10)
TW00353 (A, ST10)
K9_45F (E, ST11)
K9107F (E, ST11)
1-176-05_S1_C2 (A, ST10)
1-176-05_S1_C3 (A, ST10)
1-176-05_S1_C1 (A, ST10)
303301 (B2, ST589)
401210 (A, ST1788)
401140 (A, ST1788)
C581-05 (B2, ST28)
302312 (B2, ST589)
300231 (B2, ST28)
100290 (B2, ST1283)
102550 (B1, ND)
102598 (B2, ST1283)
103447 (B1, ST4750)
302014 (B1, ST3)
402559 (B2, ST1907)
pB171_90, B171 (B1, ST3)
S6662 (B1, ST3)
S7380 (B1, ST3)
TW10722 (B1, ST443)
I
swine71 (A, ST2525)
Envira 10/1 (B1, ST443)
Envira 8/11 (B1, ST443)
S6995 (B2, ST1283)
B71 (B1, ST131)
403128 (B2, ST5342)
103573 (B1, ST328)
100414 (B2, ND)
100343 (B1, ST328)
2731150 (A, ST3931)
2762100 (A, ST3931)
UMNF18 (A, ST10)
2866350 (B1, ST443)
5-366-08_S1_C1 (D, ST2914)
2860650 (B1, ND)
2749250 (B1, ND)
2729250 (B1, ST443)
102328 aEPEC (B1, ST517)
5-366-08_S1_C3 (D, ST2914)
8-415-05_S1_C2 (A, ST226)
E1777 (B1, ST443)
E2265 (B1, ST443)
B37 (B2, ST131)
200146 (B1, ST328)
2-011-08_S1_C2 (E, ST3018)
504225 aEPEC (B2, ST803)
702566 (B1, ST328)
8-415-05_S1_C1 (A, ST226)
F5995 (A, ST2353)
F6097 (A, ST2353)
Jurua 18/11 (B2, ST423)
P0299438.4 (A, ND)
103338 (A, ND)
P0301867.7 (A, ST2332)
P0301867.8 (A, ST2332)
P0301867.5 (A, ND)
P0301867.3 (A, ST2332)
P0301867.1 (A, ST2332)
P0301867.11 (A, ND)
P0301867.2 (A, ND)
BCW5717 (B2, ST583)
K2328 (B1, ST58)
K6412 (B1, ST58)
2772a (B1, ST156)
Fec101 (B1, ST1304)
E101 (E, ST3751)
TA280 (D, ND)
703063 (B2, ST121)
G702074-2 (B2, ST121)
702898 aEPEC (B2, ST121)
402767 (B2, ST121)
402480 aEPEC (B2, ST121)
401596 aEPEC (B2, ST121)
401709 aEPEC (B2, ST121)
734/3 (B1, ST4245)
swine31 (A, ST100)
K88 (A, ST5786)
swine2 (A, ND)
700495 (B1, ST3)
703533 (B1, ST3)
703450 (B1, ST3)
2845650 (E, ND)
E2539C1 (A, ST1312)
103578 (B1, ST3)
2875150 (A, ST1201)
swine12 (B1, ST90)
P0302293.10 (B1, ST155)
2726800 (B1, ST5396)
P0302293.2 (B1, ST155)
P0302293.3 (B1, ND)
P0302293.4 (B1, ND)
P0302293.6 (B1, ST155)
P0302293.7 (B1, ST155)
P0302293.8 (B1, ND)
P0302293.9 (B1, ND)
● repA
▶ csi
■ traI
tEPEC
aEPEC
STEC
ETEC
EAEC
Other
● bootstrap ≥50
0.010

Figure S1. Phylogenetic analysis of *repA*. The *repA* nucleotide sequences from the resistance plasmid of EPEC strain B171 (pB171_90), also from the B171 virulence plasmid (pB171, GenBank: NC_002142.1)(1), and from the other *csi* and/or *traI*-containing *E. coli* genomes (see Table S1) were aligned using ClustalW of MEGA7 (2). The alignment was used to construct a maximum-likelihood phylogeny with the Kimura 2-parameter model and 1,000 bootstraps also using MEGA7 (2). The scale bar represents the approximate distance of 0.01 nucleotide substitutions per site. Bootstrap values ≥50 are indicated by a gray circle. The predicted *E. coli* pathotypes are indicated by the color of their genome label (see legend), and the phylogroup and MLST sequence type (ST) of each genome is denoted in parentheses. A blue vertical line indicates the formation of a clade with the same composition as clade I in the *csi* and *traI* phylogenies, while no similar group of genomes corresponding to clade II was identified.

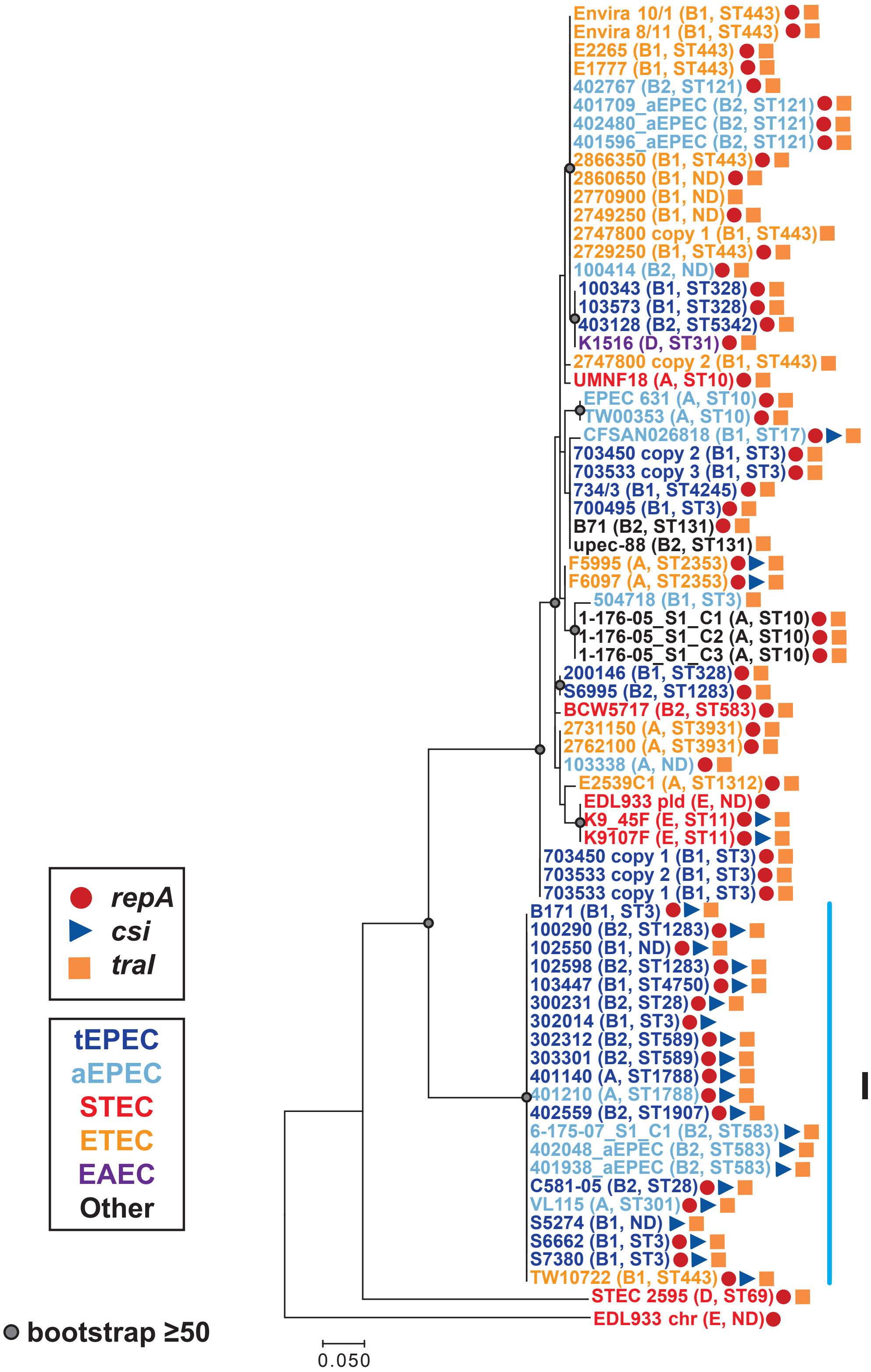
Envira 10/1 (B1, ST443)
Envira 8/11 (B1, ST443)
E2265 (B1, ST443)
E1777 (B1, ST443)
402767 (B2, ST121)
401709_aEPEC (B2, ST121)
402480_aEPEC (B2, ST121)
401596_aEPEC (B2, ST121)
2866350 (B1, ST443)
2860650 (B1, ND)
2770900 (B1, ND)
2749250 (B1, ND)
2747800 copy 1 (B1, ST443)
2729250 (B1, ST443)
100414 (B2, ND)
100343 (B1, ST328)
103573 (B1, ST328)
403128 (B2, ST5342)
K1516 (D, ST31)
2747800 copy 2 (B1, ST443)
UMNF18 (A, ST10)
EPEC 631 (A, ST10)
TW00353 (A, ST10)
CFSAN026818 (B1, ST17)
703450 copy 2 (B1, ST3)
703533 copy 3 (B1, ST3)
734/3 (B1, ST4245)
700495 (B1, ST3)
B71 (B2, ST131)
upec-88 (B2, ST131)
F5995 (A, ST2353)
F6097 (A, ST2353)
504718 (B1, ST3)
1-176-05_S1_C1 (A, ST10)
1-176-05_S1_C2 (A, ST10)
1-176-05_S1_C3 (A, ST10)
200146 (B1, ST328)
S6995 (B2, ST1283)
BCW5717 (B2, ST583)
2731150 (A, ST3931)
2762100 (A, ST3931)
103338 (A, ND)
E2539C1 (A, ST1312)
EDL933 pld (E, ND)
K9_45F (E, ST11)
K9107F (E, ST11)
703450 copy 1 (B1, ST3)
703533 copy 2 (B1, ST3)
703533 copy 1 (B1, ST3)
B171 (B1, ST3)
100290 (B2, ST1283)
102550 (B1, ND)
102598 (B2, ST1283)
103447 (B1, ST4750)
300231 (B2, ST28)
302014 (B1, ST3)
302312 (B2, ST589)
303301 (B2, ST589)
401140 (A, ST1788)
401210 (A, ST1788)
402559 (B2, ST1907)
6-175-07_S1_C1 (B2, ST583)
402048_aEPEC (B2, ST583)
401938_aEPEC (B2, ST583)
C581-05 (B2, ST28)
VL115 (A, ST301)
S5274 (B1, ND)
S6662 (B1, ST3)
S7380 (B1, ST3)
TW10722 (B1, ST443)
STEC 2595 (D, ST69)
EDL933 chr (E, ND)
I
● repA
▶ csi
■ traI
tEPEC
aEPEC
STEC
ETEC
EAEC
Other
● bootstrap ≥50
0.050

Figure S2. Phylogenetic analysis of *hha*. The *hha* nucleotide sequences from the resistance plasmid of EPEC strain B171 (pB171_90), the *csi* and/or *traI*-containing *E. coli* genomes (Table S1), and the *hha* genes located on the chromosome and plasmid of *E. coli* O157:H7 strain EDL933 were aligned using ClustalW of MEGA7 (2). The alignment was used to construct a maximum-likelihood phylogeny with the Kimura 2-parameter model and 1,000 bootstraps also using MEGA7 (2). The scale bar represents the approximate distance of 0.01 nucleotide substitutions per site. Bootstrap values ≥50 are indicated by a gray circle. The predicted *E. coli* pathotypes are indicated by the color of their genome label (see legend), and the phylogroup and MLST sequence type (ST) of each genome is denoted in parentheses. A blue vertical line indicates the formation of a clade with the same composition as clade I in the *csi* and *traI* phylogenies, while no similar group of genomes corresponding to clade II was identified.

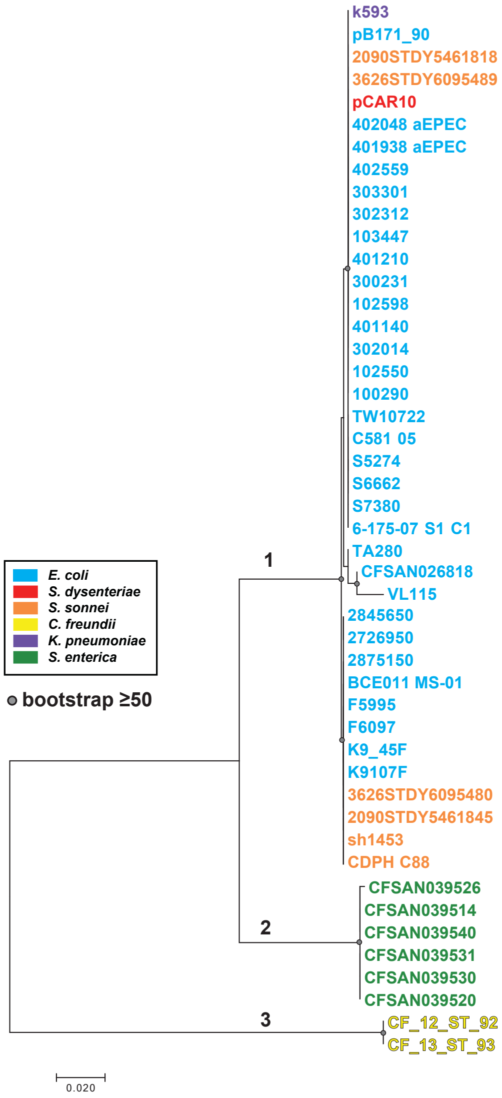

Figure S3. Phylogenetic analysis of *csi*. The *csi* nucleotide sequences of pB171_90, *E. coli,* and other human enteric pathogens were aligned using ClustalW. The alignment was used to construct a maximum-likelihood phylogeny with the Kimura 2-parameter model and 1,000 bootstraps using MEGA7 (2). The scale bar represents the approximate distance of 0.02 nucleotide substitutions per site. Bootstrap values ≥50 are indicated by a shaded circle. The *E. coli* genomes are indicated in black, while the other enterics are indicated by the color of their genome label (see legend). The numbers 1-3 designate three distinct clades of *csi* sequences.

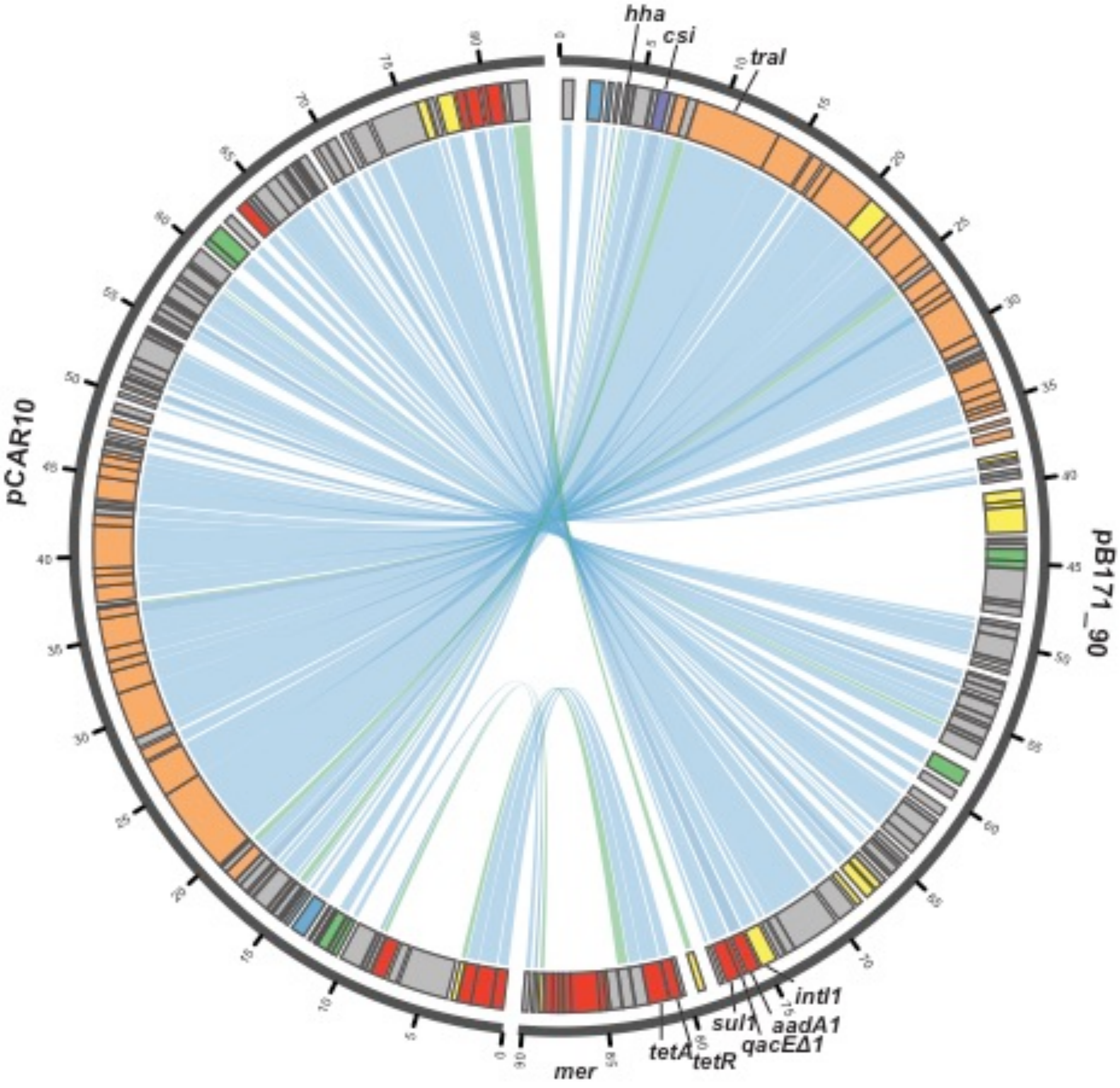

Figure S4. Comparison of plasmids pB171_90 and pCAR10. Similarity among the protein-coding genes of pB171_90 and the *S. dysenteriae* plasmid pCAR10 (GenBank: KT754161.1) is indicated by links representing BSR values obtained by comparing pB171_90 genes to pCAR10 genes using BLASTN (3). Blue links indicate genes (79 genes) with significant similarity (BSR ≥0.7) and green links indicate genes (seven genes) with divergent similarity (BSR <0.7 and ≥0.4). pB171_90 genes with no links (22 genes) were absent (BSR <0.4) from pCAR10. The outer track of the plot contains the predicted protein-coding genes of each plasmid with their protein functions indicated by colors as follows: replication (blue), virulence (purple), antibiotic resistance (red), conjugative transfer (orange), plasmid stability (green), transposition (yellow), and unknown (grey).

REFERENCES

1. **Tobe T, Hayashi T, Han CG, Schoolnik GK, Ohtsubo E, Sasakawa C.** 1999. Complete DNA sequence and structural analysis of the enteropathogenic *Escherichia coli* adherence factor plasmid. Infect Immun **67:**5455-5462.
2. **Kumar S, Stecher G, Tamura K.** 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol **33:**1870-1874.
3. **Sahl JW, Caporaso JG, Rasko DA, Keim P.** 2014. The large-scale blast score ratio (LS-BSR) pipeline: a method to rapidly compare genetic content between bacterial genomes. PeerJ **2:**e332.

Table S1. Genomes analyzed in this study

Strain ID	Pathovar[a]	Molecular Serotype[b]	MLST ST[c]	Phylogroup	EPEC Lineage	Pathovar-Specific Features[d]	Select Genes of pB171_90[e]							Accession No.
							repA	*hha*	*csi*	*traI*	*tet*	*sul1*	*mer*	
100290	tEPEC	O127:H6	1283	B2	EPEC1	LEE, BFP	+	+	+	+	-	+	0	JHQV00000000.1
100343	tEPEC	ONT:H7	328 (278)	B1	EPEC7	LEE, BFP	+	+	-	+	-	+	0	JHQW00000000.1
102535	tEPEC	O76:H51	ND	B2	EPEC8	LEE, BFP	+	-	-	+	-	-	0	JHST00000000.1
102550	tEPEC	ND	ND	B1	None	LEE, BFP	+	+	+	+	+	+	7	JHQY00000000.1
102598	tEPEC	O127:H6	1283	B2	EPEC1	LEE, BFP	+	+	+	+	-	+	0	JHRW00000000.1
103447	tEPEC	ONT:H2	4750	B1	None	LEE, BFP	+	+	+	+	+	+	7	JHSV00000000.1
103573	tEPEC	ONT:H7	328 (278)	B1	EPEC7	LEE, BFP	+	+	-	+	-	+	0	JHRA00000000.1
103578	tEPEC	O111:H2	3 (20)	B1	EPEC2	LEE, BFP	+	-	-	+	-	+	7	JHRY00000000.1
200146	tEPEC	O88:H25	328 (278)	B1	EPEC7	LEE, BFP	+	+	-	+	+	-	0	JHRZ00000000.1
300231	tEPEC	O119:H6	28 (28)	B2	EPEC4	LEE, BFP	+	+	+	+	-	-	0	JHSC00000000.1
302014	tEPEC	O111:H2	3 (20)	B1	EPEC2	LEE, BFP	+	+	+	-	+	+	0	JHRE00000000.1
302312	tEPEC	O51:H49	589	B2	None	LEE, BFP	+	+	+	+	+	+	0	JHTA00000000.1
303301	tEPEC	O51:H49	589	B2	None	LEE, BFP	+	+	+	+	+	+	0	JHTC00000000.1
401140	tEPEC	ONT:H27	1788	A	EPEC5	LEE, BFP	+	+	+	+	+	+	7	JHRM00000000.1
402559	tEPEC	ONT:H34	1907	B2	None	LEE, BFP	+	+	+	+	+	+	0	JHTH00000000.1
403128	tEPEC	O76:H51	5342	B2	EPEC8	LEE, BFP	+	+	-	+	-	-	0	CXZV00000000.1
700495	tEPEC	O126:H2	3 (20)	B1	EPEC2	LEE, BFP	+	+	-	+	-	+	0	CYBA00000000.1
702423	tEPEC	O126:H2	6292	B1	EPEC2	LEE, BFP	+	-	-	+	-	-	0	JHSO00000000.1
703450	tEPEC	O126:H2	3 (20)	B1	EPEC2	LEE, BFP	+	+	-	+	-	+	0	JHTM00000000.1
703533	tEPEC	ONT:H2	3 (20)	B1	EPEC2	LEE, BFP	+	+	-	+	-	+	0	JHRS00000000.1
32/73	tEPEC	O126:H2	3 (20)	B1	EPEC2	LEE, BFP	+	-	-	+	-	-	0	LAFA00000000.1
B171	tEPEC	O111:H2	3 (20)	B1	EPEC2	LEE, BFP	+	+	+	+	+	+	7	AAJX00000000.2
C342-62	tEPEC	O126:H2	3 (20)	B1	EPEC2	LEE, BFP	+	-	-	+	-	-	0	AKNI00000000.1
C581-05	tEPEC	O119:H6	28 (28)	B2	EPEC4	LEE, BFP	+	+	+	+	-	+	0	AIBE00000000.1
S5274	tEPEC	O111:H2	ND	B1	EPEC2	LEE, BFP	-	+	+	+	+	+	0	JICB00000000.1
S6662	tEPEC	ONT:H2	3 (20)	B1	EPEC2	LEE, BFP	+	+	+	+	+	+	0	JICC00000000.1
S6995	tEPEC	O127:H6	1283	B2	EPEC1	LEE, BFP	+	+	-	+	-	-	0	JICH00000000.1
S7005	tEPEC	ONT:H2	3 (20)	B1	EPEC2	LEE, BFP	+	-	-	+	-	+	0	JICG00000000.1
S7380	tEPEC	O111:H2	3 (20)	B1	EPEC2	LEE, BFP	+	+	+	+	+	+	0	JICD00000000.1
100414	aEPEC	O33:H6	ND	B2	EPEC4	LEE	+	+	-	+	-	-	0	JHQX00000000.1
103338	aEPEC	O118/O151:H5	ND	A	EPEC5	LEE	+	+	-	+	-	-	0	JHSU00000000.1
401210	aEPEC	ND	1788	A	EPEC5	LEE	+	+	+	+	+	+	7	JHSJ00000000.1
402767	aEPEC	O33:HNT	121 (28)	B2	EPEC4	LEE	+	+	-	+	+	+	7	CYET00000000.1
504718	aEPEC	O128ac:H2	3 (20)	B1	EPEC2	LEE	-	+	-	+	-	-	0	CYDK00000000.1
702566	aEPEC	O88:H25	328 (278)	B1	EPEC7	LEE	+	-	-	+	+	-	0	CYDT00000000.1
703063	aEPEC	O33:HNT	121 (28)	B2	EPEC4	LEE	+	-	-	+	-	+	6	CYEU00000000.1
102328_aEPEC	aEPEC	O171:H19	517	B1	None	LEE	+	-	-	+	-	-	0	CYBD00000000.1
401596_aEPEC	aEPEC	O33:HNT	121 (28)	B2	EPEC4	LEE	+	+	-	+	+	+	7	CYBO00000000.1
401709_aEPEC	aEPEC	O33:HNT	121 (28)	B2	EPEC4	LEE	+	+	-	+	+	+	7	CYCR00000000.1
401938_aEPEC	aEPEC	O63:H6	583 (122)	B2	None	LEE	-	+	+	+	+	+	7	CYBT00000000.1
402048_aEPEC	aEPEC	O63:H6	583 (122)	B2	None	LEE	-	+	+	+	+	+	7	CXZI00000000.1
402480_aEPEC	aEPEC	O33:HNT	121 (28)	B2	EPEC4	LEE	+	+	-	+	+	+	7	CYBU00000000.1
504225_aEPEC	aEPEC	O85:H31	803	B2	None	LEE	+	-	-	+	-	-	0	CXZC00000000.1
6-175-07_S1_C1	aEPEC	O63:H6	583 (122)	B2	None	LEE	-	+	+	+	-	-	0	JOQR00000000.1
702898_aEPEC	aEPEC	O33:HNT	121 (28)	B2	EPEC4	LEE	+	-	-	+	-	+	6	CYBW00000000.1
CFSAN026818	aEPEC	O103:H2	17 (20)	B1	EPEC2	LEE	+	+	+	+	-	-	0	LDCR00000000.1
EPEC 631	aEPEC	O157:H16	10 (10)	A	EPEC10	LEE	+	+	-	+	-	-	0	LZDW00000000.1
G702074-2	aEPEC	O33:HNT	121 (28)	B2	EPEC4	LEE	+	-	-	+	-	+	6	CYGA00000000.1
TW00353	aEPEC	O157:H16	10 (10)	A	EPEC10	LEE	+	+	-	+	-	-	0	AMUM00000000.1
VL115	aEPEC	O76:H2	301 (165)	A	None	LEE	+	+	+	+	-	-	0	MIWW00000000.1
K1516	EAEC	O15:H18	31 (31)	D	NA	*aggR, aatA, pic*	+	+	-	+	+	+	7	JHJE00000000.1
2726800	ETEC	O98:H32	5396	B1	NA	LT, ST	+	-	-	+	-	-	0	AQFE00000000.1
2726950	ETEC	O128ac:H12	10 (10)	A	NA	ST	+	-	+	+	-	-	0	AQED00000000.1
2729250	ETEC	O115:H5	443 (205)	B1	NA	LT, ST	+	+	-	+	-	-	0	AQEC00000000.1
2731150	ETEC	O25:H16	3931	A	NA	LT	+	+	-	+	-	-	0	AQEB00000000.1
2747800	ETEC	O115:H5	443 (205)	B1	NA	LT, ST	-	+	-	+	-	-	0	AQEA00000000.1
2749250	ETEC	O115:H5	ND	B1	NA	LT, ST	+	+	-	+	-	-	0	AQDZ00000000.1
2762100	ETEC	O25:H16	3931	A	NA	LT	+	+	-	+	-	-	0	AQDX00000000.1
2770900	ETEC	O115:H5	ND	B1	NA	LT, ST	-	+	-	+	-	-	0	AQDW00000000.1
2845650	ETEC	O169/O183:H41	ND	E	NA	ST	+	-	+	+	-	-	0	AQDR00000000.1
2860650	ETEC	O115:H5	ND	B1	NA	LT, ST	+	+	-	+	-	-	0	APXG00000000.1
2866350	ETEC	O115:H5	443 (205)	B1	NA	LT, ST	+	+	-	+	-	-	0	APXJ00000000.1
2875150	ETEC	O25:H42	1201	A	NA	ST	+	-	+	+	-	-	0	APXK00000000.1
BCE011_MS-01	ETEC	O128ac:H12	10 (10)	A	NA	ST	-	-	+	+	-	-	0	APXM00000000.1
CE418	ETEC	O148:H28	94 (448)	B1	NA	ST	+	-	-	+	-	+	1	ATEW00000000.1
E1777	ETEC	O115:H5	443 (205)	B1	NA	LT, ST	+	+	-	+	-	-	0	JTHI00000000.1
E2265	ETEC	O115:H5	443 (205)	B1	NA	LT, ST	+	+	-	+	-	-	0	JUBB00000000.1
E24377A	ETEC	ONT:H28	1132	B1	NA	LT, ST	+	-	-	+	-	-	0	NC_009801.1
E2539C1	ETEC	O25:H16	1312	A	NA	LT	+	+	-	+	-	-	0	JHJV00000000.1
Envira 10/1	ETEC	O167:H5	443 (205)	B1	NA	LT, ST	+	+	-	+	+	-	0	AQFD00000000.1
Envira 8/11	ETEC	O167:H5	443 (205)	B1	NA	LT, ST	+	+	-	+	+	-	0	AQFC00000000.1
F5995	ETEC	O6:H16	2353	A	NA	LT, ST	+	+	+	+	-	-	0	JPXL00000000.1
F6097	ETEC	O6:H16	2353	A	NA	LT, ST	+	+	+	+	-	-	0	JPXJ00000000.1
F6700	ETEC	O148:H28	1988	B1	NA	LT, ST	-	-	-	+	+	-	0	JYIA00000000.1
Jurua 18/11	ETEC	O8:H9	423 (23)	B1	NA	LT	+	-	-	+	-	-	0	AQFB00000000.1
734/3	ETEC	ONT:H19	4245	B1	NA	LT, ST	+	+	-	+	-	+	7	JPQX00000000.1
P0299438.10	ETEC	ONT:H2	100 (165)	A	NA	LT, ST	-	-	-	+	-	-	0	APYA00000000.1
P0299438.4	ETEC	ONT:H2	ND	A	NA	LT, ST	+	-	-	+	-	-	0	APYD00000000.1
P0301867.1	ETEC	O71:H45	2332	A	NA	ST	+	-	-	+	-	-	0	AQEQ00000000.1
P0301867.11	ETEC	O71:H45	ND	A	NA	ST	+	-	-	+	-	-	0	APYT00000000.1
P0301867.13	ETEC	O71:H45	2332	A	NA	ST	-	-	-	+	-	-	0	APZT00000000.1

P0301867.2	ETEC	O71:H45	ND	A	NA	ST	+	-	-	+	-	-	0	APZU00000000.1
P0301867.3	ETEC	O71:H45	2332	A	NA	ST	+	-	-	+	-	-	0	APYU00000000.1
P0301867.4	ETEC	O71:H45	2332	A	NA	ST	-	-	-	+	-	-	0	APYV00000000.1
P0301867.5	ETEC	O71:H45	ND	A	NA	ST	+	-	-	+	-	-	0	APYW00000000.1
P0301867.7	ETEC	O71:H45	2332	A	NA	ST	+	-	-	+	-	-	0	APYX00000000.1
P0301867.8	ETEC	ONT:H45	2332	A	NA	ST	+	-	-	+	-	-	0	APYY00000000.1
P0302293.10	ETEC	ONT:H18	155 (155)	B1	NA	ST	+	-	-	+	-	-	0	AQAA00000000.1
P0302293.2	ETEC	ONT:H18	155 (155)	B1	NA	LT, ST	+	-	-	+	-	-	0	AQEP00000000.1
P0302293.3	ETEC	ONT:H18	ND	B1	NA	LT, ST	+	-	-	+	-	-	0	AQAB00000000.1
P0302293.4	ETEC	ONT:H18	ND	B1	NA	LT, ST	+	-	-	+	-	-	0	AQAC00000000.1
P0302293.6	ETEC	ONT:H18	155 (155)	B1	NA	LT, ST	+	-	-	+	-	-	0	AQAD00000000.1
P0302293.7	ETEC	ONT:H18	155 (155)	B1	NA	LT, ST	+	-	-	+	-	-	0	APZW00000000.1
P0302293.8	ETEC	ONT:H18	ND	B1	NA	LT, ST	+	-	-	+	-	-	0	AQAE00000000.1
P0302293.9	ETEC	ONT:H18	ND	B1	NA	LT, ST	+	-	-	+	-	-	0	AQAF00000000.1
swine12	ETEC	O7:H19	90 (23)	B1	NA	LT	+	-	-	+	+	-	0	LVMO00000000.1
swine2	ETEC	O149:H10	ND	A	NA	LT	+	-	-	+	+	-	0	LVME00000000.1
swine24	ETEC	O149:H10	ND	A	NA	LT	-	-	-	+	-	-	0	LVNA00000000.1
swine31	ETEC	O149:H10	100 (165)	A	NA	LT	+	-	-	+	+	-	7	LVNH00000000.1
swine71	ETEC	O3:H27	2525	A	NA	ST	+	-	-	+	+	-	7	LVOU00000000.1
TW10722	ETEC	O115:H5	443 (205)	B1	NA	ST	+	+	+	+	+	+	3	AELB00000000.1
TX1999	ETEC	O8:H19	90 (23)	B1	NA	LT	+	-	-	+	-	-	5	AFEB00000000.1
BCW5717	STEC	O63:H6	583 (122)	B2	NA	LEE, BFP, STX	+	+	-	+	-	-	0	FBVB00000000.1
ED180	STEC	O26:H11	21 (29)	B1	NA	LEE, STX	+	-	-	+	+	-	0	FAVO00000000.1
K9_45F	STEC	O157:H7	11 (11)	E	NA	LEE, STX	+	+	+	+	-	-	0	LHAP00000000.1
K9107F	STEC	O157:H7	11 (11)	E	NA	LEE, STX	+	+	+	+	-	-	0	LHAQ00000000.1
K88	STEC	O141ab:H4	5786	A	NA	STX, LT	+	-	-	+	-	-	0	LBBN00000000.1
STEC 2595	STEC	O17/O77:H18	69 (69)	D	NA	STX	+	+	-	+	-	-	0	LOIJ00000000.1
UMNF18	STEC	O147:H4	10 (10)	A	NA	ST, STX	+	+	-	+	-	-	7	AGTD00000000.1
upec-39	ExPEC/UPEC	O2:H1	73 (73)	B2	NA	*papG*	-	-	-	+	-	-	0	JSHV00000000.1
60_SBOY	none	O16:H5	131	B2	NA	none	+	-	-	+	+	+	0	JVAO00000000.1
B37	none	O16:H5	131	B2	NA	none	+	-	-	+	-	-	7	LRWD00000000.1
B71	none	O25:H4	131	B2	NA	none	+	+	-	+	-	-	0	LRWM00000000.1
upec-88	none	O16:H5	131	B2	NA	none	-	+	-	+	-	-	0	JSGB00000000.1
5-366-08_S1_C3	none	O166:H15	2914	D	NA	none	+	-	-	+	-	+	0	JONE00000000.1
1-176-05_S1_C1	none	ONT:H10	10 (10)	A	NA	none	+	+	-	+	+	+	7	JHDH00000000.1
1-176-05_S1_C2	none	ONT:H10	10 (10)	A	NA	none	+	+	-	+	+	+	7	JJMA00000000.1
1-176-05_S1_C3	none	ONT:H10	10 (10)	A	NA	none	+	+	-	+	+	+	7	JJLY00000000.1
2-011-08_S1_C2	none	ONT:H41	3018	E	NA	none	+	-	-	+	-	-	0	JJMB00000000.1
2772a	none	ND	156 (156)	B1	NA	none	+	-	-	+	-	-	0	LCVG00000000.1
5-366-08_S1_C1	none	O166:H15	2914	D	NA	none	+	-	-	+	-	+	0	JOQU00000000.1
6-537-08_S3_C1	none	O174:H7	192	B1	NA	none	+	-	-	+	-	-	0	JORP00000000.1
6-537-08_S3_C2	none	O174:H7	192	B1	NA	none	+	-	-	+	-	-	0	JORX00000000.1
6-537-08_S3_C3	none	O174:H7	192	B1	NA	none	+	-	-	+	-	-	0	JORQ00000000.1
8-415-05_S1_C1	none	O40:H4	226 (226)	A	NA	none	+	-	-	+	+	-	7	JORZ00000000.1
8-415-05_S1_C2	none	O40:H4	226 (226)	A	NA	none	+	-	-	+	+	-	7	JOSQ00000000.1
CVM N36834PS	none	O100:H19	1611	B1	NA	none	-	-	-	+	-	-	0	JUCF00000000.1
DEC7D	none	ONT:H43	90 (23)	B1	NA	none	+	-	-	+	-	-	0	AIGD00000000.1
E101	none	O187:H9	3751	E	NA	none	+	-	-	+	-	-	0	AEXG00000000.1
Fec101	none	O182:H16	1304	B1	NA	none	+	-	-	+	-	-	0	MDYZ00000000.1
K2328	none	O88:H25	58 (155)	B1	NA	none	+	-	-	+	+	-	7	MIWN00000000.1
K6412	none	O88:H25	58 (155)	B1	NA	none	+	-	-	+	+	-	7	MIWP00000000.1
TA280	none	ONT:H34	ND	D	NA	none	+	-	+	+	-	-	0	ADBA00000000.1
upec-130	none	O93:H8	ND	B1	NA	none	-	-	-	+	-	-	0	JSNV00000000.1

Reference Genomes in the Phylogenomic Analysis

Strain ID	Pathovar[a]		MLST ST[b]	Phylogroup	EPEC Lineage	Pathovar-Specific Features	*repA*	*hha*	*csi*	*traI*	*tet*	*sul1*	*mer*	Accession No.
100329	tEPEC	ONT:H33	34 (10)	A	EPEC10	LEE, BFP	+	+	-	-	+	+	7	JHRT00000000.1
300214	tEPEC	O55:H51	5342	B2	EPEC8	LEE, BFP	+	+	-	-	+	+	7	JHSB00000000.1
302053	tEPEC	O33:H34	2346	B2	EPEC9	LEE, BFP	+	+	-	-	+	+	0	JHRG00000000.1
302150	tEPEC	ONT:H6	28 (28)	B2	EPEC4	LEE, BFP	+	+	-	-	-	-	0	JHSD00000000.1
303145	tEPEC	O33:H34	2346	B2	EPEC9	LEE, BFP	+	-	-	-	+	+	2	JHSG00000000.1
401091	tEPEC	ONT:H7	ND	B1	EPEC7	LEE, BFP	+	-	-	-	+	+	7	JHSI00000000.1
401588	tEPEC	O55:H51	5342	B2	EPEC8	LEE, BFP	+	+	-	-	+	+	7	JHSK00000000.1
402290	tEPEC	ONT:H7	ND	B1	EPEC7	LEE, BFP	+	-	-	-	+	+	7	JHRQ00000000.1
702324	tEPEC	O142:H34	2346	B2	EPEC9	LEE, BFP	+	+	-	-	+	+	7	JHRR00000000.1
C260-92	tEPEC	ONT:H40	10 (10)	A	EPEC10	LEE, BFP	-	-	-	-	-	-	0	AIAO00000000.1
E2348/69	tEPEC	O127:HNT	15	B2	EPEC1	LEE, BFP	+	-	-	-	-	-	0	NC_011601.1
2362-75	tEPEC	O55:H6	19	B2	EPEC1	LEE, BFP	+	+	-	-	-	+	7	ADUL00000000.1
E110019	aEPEC	ONT:H9	381 (590)	B1	None	LEE	+	+	-	-	-	-	0	AAJW00000000.2
CB9615	aEPEC	O55:H7	335 (11)	E	None	LEE	-	-	-	-	-	-	0	NC_013941.1
HS	commensal	O9:H4	46 (46)	A	NA	none	-	-	-	-	-	-	0	NC_009800.1
042	EAEC	O44:H18	414	D	NA	*aggR, aatA, pic, pet*	+	-	-	-	+	+	7	FN554766.1
55989	EAEC	O104:H4	678	B1	NA	*pic*	-	-	-	-	-	-	0	NC_011748.1
TY-2482	STEC/EAEC	O104:H4	678	B1	NA	*pic*, STX	-	-	-	-	+	+	7	AFOG00000000.1
EDL933	STEC/EHEC	O157:H7	ND	E	NA	LEE, STX	+	+	-	-	-	-	0	NC_002655.2
Sakai	STEC/EHEC	O157:H7	11 (11)	E	NA	LEE, STX	+	+	-	-	-	-	0	NC_002695.1
11368	STEC/EHEC	O26:H11	21 (29)	B1	NA	LEE, STX	+	-	-	-	-	-	0	NC_013361.1
11128	STEC/EHEC	O111:H8	16 (29)	B1	NA	LEE, STX	+	+	-	-	-	-	7	NC_013364.1
2846750	ETEC	O6:H16	4 (10)	A	NA	LT, ST	-	+	-	-	+	+	7	AQGG00000000.1
2851500	ETEC	O114:H49	5305	B1	NA	LT	-	+	-	-	-	-	0	AQDN00000000.1
2854350	ETEC	O98:H32	5396	B1	NA	LT, ST	+	-	-	-	-	-	0	APZL00000000.1
2866450	ETEC	O114:H49	5305	B1	NA	LT	+	+	-	-	-	-	0	AQDI00000000.1
B7A	ETEC	O148:H28	94 (448)	B1	NA	LT, ST	+	-	-	-	+	+	0	AAJT00000000.2
BCE002_MS-12	ETEC	O153/O178:HNT	5587	B1	NA	LT	+	-	-	-	+	-	0	AQDA00000000.1
BCE019_MS-13	ETEC	O6:H16	2353	A	NA	LT, ST	-	+	-	-	-	-	0	AQCZ00000000.1

BCE034_MS-14	ETEC	O78:HNT	173	B1	NA	ST	+	-	-	-	-	-	0	AQCY00000000.1
H10407	ETEC	O78:H11	48 (10)	A	NA	LT, ST	+	+	-	-	-	-	0	FN649414.1
MP021566.1	ETEC	O6:H16	2353	A	NA	LT, ST	+	+	-	-	-	-	0	AQEU00000000.1
ThroopD	ETEC	O63:H9	5392	B1	NA	LT, ST	+	+	-	-	-	-	0	AQEJ00000000.1
TW10598	ETEC	O6:H16	4 (10)	A	NA	LT, ST	+	+	-	-	-	-	0	AELA00000000.2
TW10828	ETEC	O114:H49	173	B1	NA	LT	-	+	-	-	-	-	0	AELC00000000.1
TW11681	ETEC	ONT:H45	2332	A	NA	ST	+	+	-	-	+	-	7	AELD00000000.1
TW14425	ETEC	O78:H9	23 (23)	B1	NA	ST	+	+	-	-	-	-	0	AELE00000000.1
UMNK88	ETEC	O149:H10	100 (165)	A	NA	LT	+	-	-	-	+	+	6	NC_017641.1
SE11	fecal isolate	O173:H28	156 (156)	B1	NA	none	-	+	-	-	+/-	-	0	NC_011415.1
IAI1	fecal isolate	O8:H19	1128	B1	NA	none	-	-	-	-	-	-	0	NC_011741.1
BL21	lab adapted	O7:H48	93 (168)	A	NA	none	-	-	-	-	-	-	0	NC_012947.1
BW2952	lab adapted	O16:H48	10 (10)	A	NA	none	-	-	-	-	-	-	0	NC_012759.1
ATCC 8739	lab adapted	O146:H20	ND	A	NA	none	-	-	-	-	-	-	0	NC_010468.1
SMS_3_5	other	ONT:H34	354 (354)	F	NA	none	+	+	-	-	+	-	0	CP000970.1
53638	EIEC	O124:H30	6	A	NA	*ospB*	+	+	-	-	-	-	0	AAKB00000000.2
S88	ExPEC	O45:H7	95 (95)	B2	NA	*papG*	+	-	-	-	-	-	0	NC_011742.1
536	ExPEC/UPEC	O6:H31	127	B2	NA	*papG*	-	-	-	-	-	-	0	NC_008253.1
UTI89	ExPEC/UPEC	O18:H7	95 (95)	B2	NA	*papG*	+	+	-	-	-	-	0	NC_007946.1
CFT073	ExPEC/UPEC	O6:H1	73 (73)	B2	NA	*papG*	-	-	-	-	-	-	0	AE014075.1
IAI39	ExPEC/UPEC	O7:H45	62	F	NA	*papG*	-	-	-	-	-	-	0	NC_011750.1
UMN026	ExPEC/UPEC	O17/O77:H18	597 (69)	D	NA	*papG*	+	+	-	-	-	+	0	NC_011751.1
S. flexneri 2a 2457T	*Shigella*	ND	245 (245)	B1	NA	NA	-	-	-	-	-	-	0	NC_004741.1
S. boydii 3083-94	*Shigella*	ND	1129	B1	NA	NA	+	+	-	-	-	-	0	NC_010658.1
S. sonnei 046	*Shigella*	ND	152 (152)	B1	NA	NA	+	+	-	-	+	-	0	NC_007384.1
S. dysenteriae 197	*Shigella*	ND	146	E	NA	NA	+	-	-	-	-	-	0	NC_007606.1

[a] The pathovar/pathotype of each genome analyzed was determined by *in silico* detection of virulence genes, or based on previous strain descriptions where available.

[b] The molecular serotype was determined using SerotypeFinder v. 1.1 (https://cge.cbs.dtu.dk/services/SerotypeFinder/) (Joensen *et al.* 2015).

[c] The multilocus sequence typing (MLST) sequence type (ST) of each genome was determined by *in silico* detection of the loci in each genome and comparison with the pubmlst.org database. The number in parentheses indicates the ST clonal complex.

[d] The pathotype-specific virulence genes identified by *in silico* analysis of the genome assemblies.

[e] Genes of pB171_90 that had a BSR ≥0.7 were considered present (+) in each of the genomes, while genes with a BSR <0.7 were absent (-). The *tet* genes are denoted as present (+) when both *tetA* and *tetR* were detected, or partially present (+/-) when one of the genes was present and the other was absent, or absent (-) when both genes were missing from the genome. The number of genes in the mercuric resistance cassette (n=7) that were detected with a BSR ≥0.7 in each genome is indicated.

Table S2. Summary of sequence differences identified in the EPEC virulence plasmid sequences

	Nucleotide Sequence Differences				Protein Sequence Differences					
Location in pB171 Sequence[a]	pB171 (AB024946.1)[b]	pB171_69[c]	pB171 ORF[d]	Base Change of pB171_69 compared to pB171	Location in pB171 Protein Sequence[a]	pB171 (AB024946.1)[b]	pB171_69[c]	Protein Accession No. of pB171[d]	Predicted Protein Function[d]	Description
449	G	C	ORF1	SNV	54	V	L	BAA84836.1	replication protein	Non-synonymous
451	C	G	ORF1	SNV	54	V	L	"	"	Non-synonymous
535	C	G	ORF1	SNV	82	N	K	"	"	Non-synonymous
592	G	T	ORF1	SNV	101	V	V	"	"	Synonymous
2631	T	-	none	Deletion						Non-coding
14,652	A	-	none	Deletion						Non-coding
16,194	T	-	none	Deletion						Non-coding
17,487	G	A	none	SNV						Non-coding
20,870	G	-	ORF24	Deletion	38	G	frameshift	BAA84859.1	bundle-forming pilus protein, BfpT	Frameshift
24,488	-	C	none	Insertion						Non-coding
25,615	T	C	ORF30	SNV	34	L	L	BAA84865.1	transposase	Synonymous
25,658	A	G	ORF30	SNV	48	E	G	"	"	Non-synonymous
26,296	T	C	ORF31	SNV	138	L	L	BAA84866.1	transposase	Synonymous
26,870	G	T	ORF31	SNV	330	V	L	"	"	Non-synonymous
27,080	A	T	ORF31	SNV	400	T	S	"	"	Non-synonymous
27,122	A	C	ORF31	SNV	414	S	R	"	"	Non-synonymous
27,138	T	C	ORF31	SNV	419	V	A	"	"	Non-synonymous
27,158	G	C	ORF31	SNV	426	V	L	"	"	Non-synonymous
27,172	C	T	ORF31	SNV	430	D	D	"	"	Synonymous
27,175	G	C	ORF31	SNV	431	G	G	"	"	Synonymous
65,084	C	T	ORF76	SNV	77	R	W	BAA84911.1	putative DNA binding protein	Non-synonymous
66,376	A	G	ORF77	SNV	296	Q	Q	BAA84912.1	helix-turn-helix family protein	Synonymous
66,395	G	C	ORF77	SNV	303	A	P	"	"	Non-synonymous
66,450	G	A	ORF77	SNV	321	G	D	"	"	Non-synonymous
66,453	G	A	ORF77	SNV	322	S	K	"	"	Non-synonymous
66,454	C	A	ORF77	SNV	322	S	K	"	"	Non-synonymous
68,200	-	T	ORF80	Insertion	6	T	frameshift	BAA84915.1	hypothetical protein	Frameshift
68,287	T	-	none	Deletion						Non-coding

[a] This is the relative location of the sequence differences of pB171_69 compared to the previously-sequence virulence plasmid pB171 (AB024946.1).

[b] The previously-sequenced B171 virulence plasmid, pB171 (GenBank: AB024946.1).

[c] The newly-sequenced B171 virulence plasmid, pB171_69 (Genbank: CP021211).

[d] The designated ORFs and annotation of the previously-sequenced plasmid pB171 described in Tobe *et al.* (1999).

[e] Identification of single nucleotide variant (SNV), insertion, or deletion in pB171_69 compared to pB171.