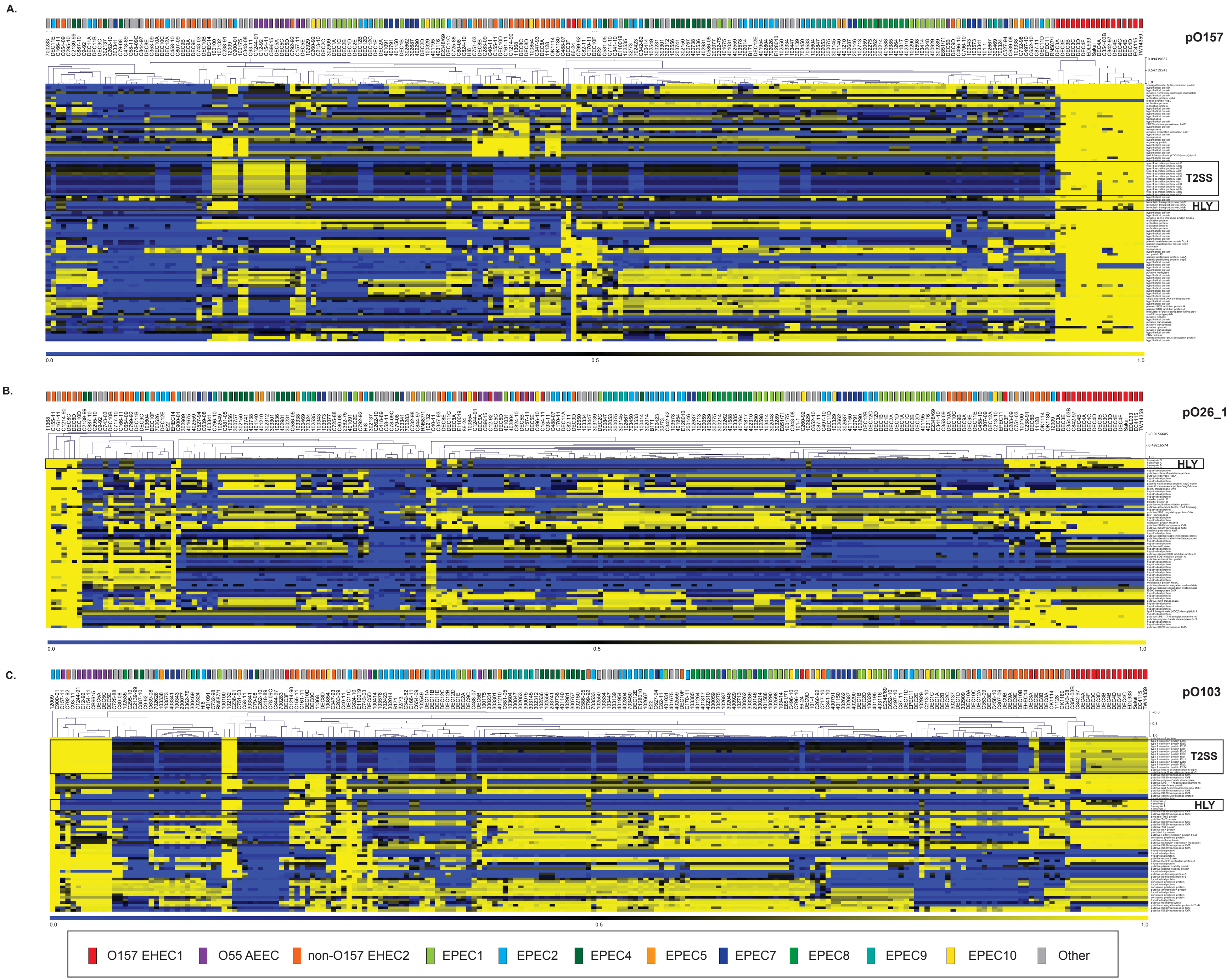
A.
pO157
T2SS
HLY
0.0
0.5
1.0
B.
pO26_1
HLY
0.0
0.5
1.0
C.
pO103
T2SS
HLY
0.0
0.5
1.0
O157 EHEC1
O55 AEEC
non-O157 EHEC2
EPEC1
EPEC2
EPEC4
EPEC5
EPEC7
EPEC8
EPEC9
EPEC10
Other

Figure S1.

Figure S1. *In silico* identification of EHEC virulence plasmids, encoding the hemolysin, *hlyA*, in the AEEC genomes analyzed in this study. BLAST Score Ratio (BSR) analysis was performed as previously described (1). The predicted amino acid sequences of the plasmid-encoded genes compared to each AEEC genome analyzed in this study using TBLASTN (2). Each box represents the BSR value of a single gene in one genome. Each row contains BSR values for a single a gene, while each column contains values identified in a single genome. Hierarchical cluster analysis was performed using the Pearson correlation coefficient of MeV (3-5). The colors of the heatmap indicate the BSR values corresponding to the presence (yellow), divergence (black), or absence (blue) of each of the plasmid genes in each AEEC genome. The colored rectangle above each genome indicates the phylogenomic lineage the genome belongs to in Figure 1. A box indicates the presence of the known and putative virulence-associated genes: hemolysin (HLY) and type II secretion system (T2SS). **A)** *In silico* identification of protein-encoding genes of the O157 EHEC virulence plasmid, pO157, of the O157:H7 EHEC prototype isolate EDL933, in the AEEC isolate genomes. **B)** *In silico* identification of proteins encoded by plasmid pO26_1 from the LEE+/*stx*+ non-O157 EHEC isolate 11368 (6), in the AEEC isolate genomes. **C)** *In silico* identification of proteins encoded by plasmid pO103 from the LEE+/*stx*+ isolate 12009 (7), in the AEEC isolate genomes. In all three of the comparisons, the virulence-associated genes encoded by each plasmid (HLY and T2SS) were most conserved among isolates within the same pathovar. Meanwhile other conserved plasmid genes encoding putative proteins involved in plasmid replication and maintenance, were identified with similarity in both EPEC and EHEC.

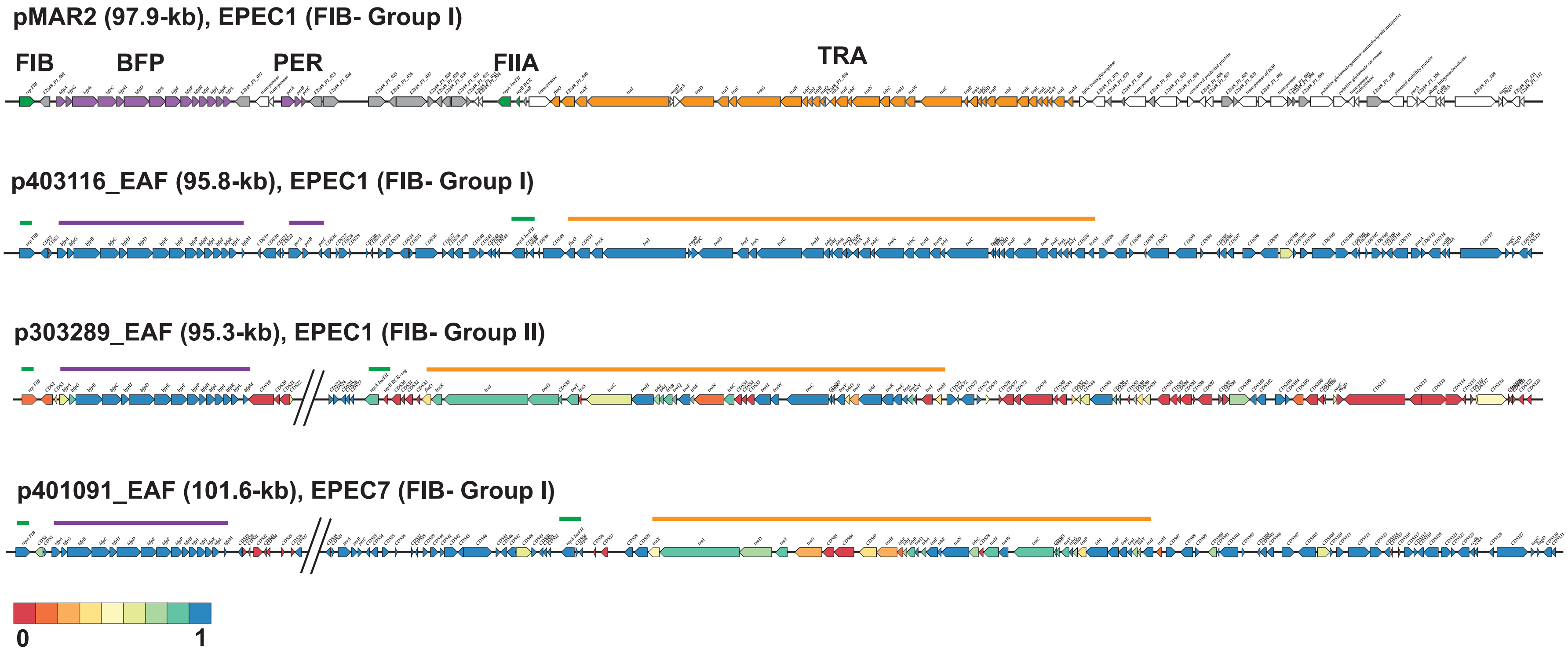
pMAR2 (97.9-kb), EPEC1 (FIB- Group I)
FIB
BFP
PER
FIIA
TRA
p403116_EAF (95.8-kb), EPEC1 (FIB- Group I)
p303289_EAF (95.3-kb), EPEC1 (FIB- Group II)
p401091_EAF (101.6-kb), EPEC7 (FIB- Group I)
0
1

Figure S2

Figure S2. Sequence diversity of the EAF plasmid identified in select EPEC isolates from the EPEC1 and EPEC7 phylogenomic lineages compared to the prototype EAF plasmid, pMAR2, from the EPEC1 isolate E2348/69 (8, 9). The genes identified in each of the EAF plasmids are denoted by arrows, and the direction of the arrows indicate the coding strand. The previously identified protein-encoding genes of pMAR2 are indicated by function as follows: replication (green), bundle-forming pilus (purple), conjugative transfer (orange), and unknown or other functions (white). The pseudogenes are indicated in grey. The colors of each gene of the EAF plasmids from EPEC isolates 403116, 303289, and 401091 indicate a BSR value determined by comparing the nucleotide sequence of the gene to the pMAR2 plasmid sequence using BLASTN. The BSR values indicate a particular gene has significant nucleotide similarity to sequence of pMAR2 (blue) or does not have similarity to pMAR2 (red). The vertical lines indicate the presence of a gap in the plasmid sequence. A green bar indicates the region containing genes encoding proteins involved replication, purple indicates the bundle-forming pilus genes, and orange indicates genes encoding predicted proteins involved in conjugative transfer. The phylogenetic group that the FIB *repA* sequence of each plasmid belongs to is indicated as Group I or Group II. Comparison of the EAF plasmid from two EPEC1 isolates and one EPEC7 isolate characterized in this study to the prototype EPEC1 plasmid pMAR2 demonstrated there can be greater diversity of the EAF plasmid among EPEC isolates within the same phylogenomic lineage compared to the EAF plasmid from EPEC of other phylogenomic lineages (EPEC7).

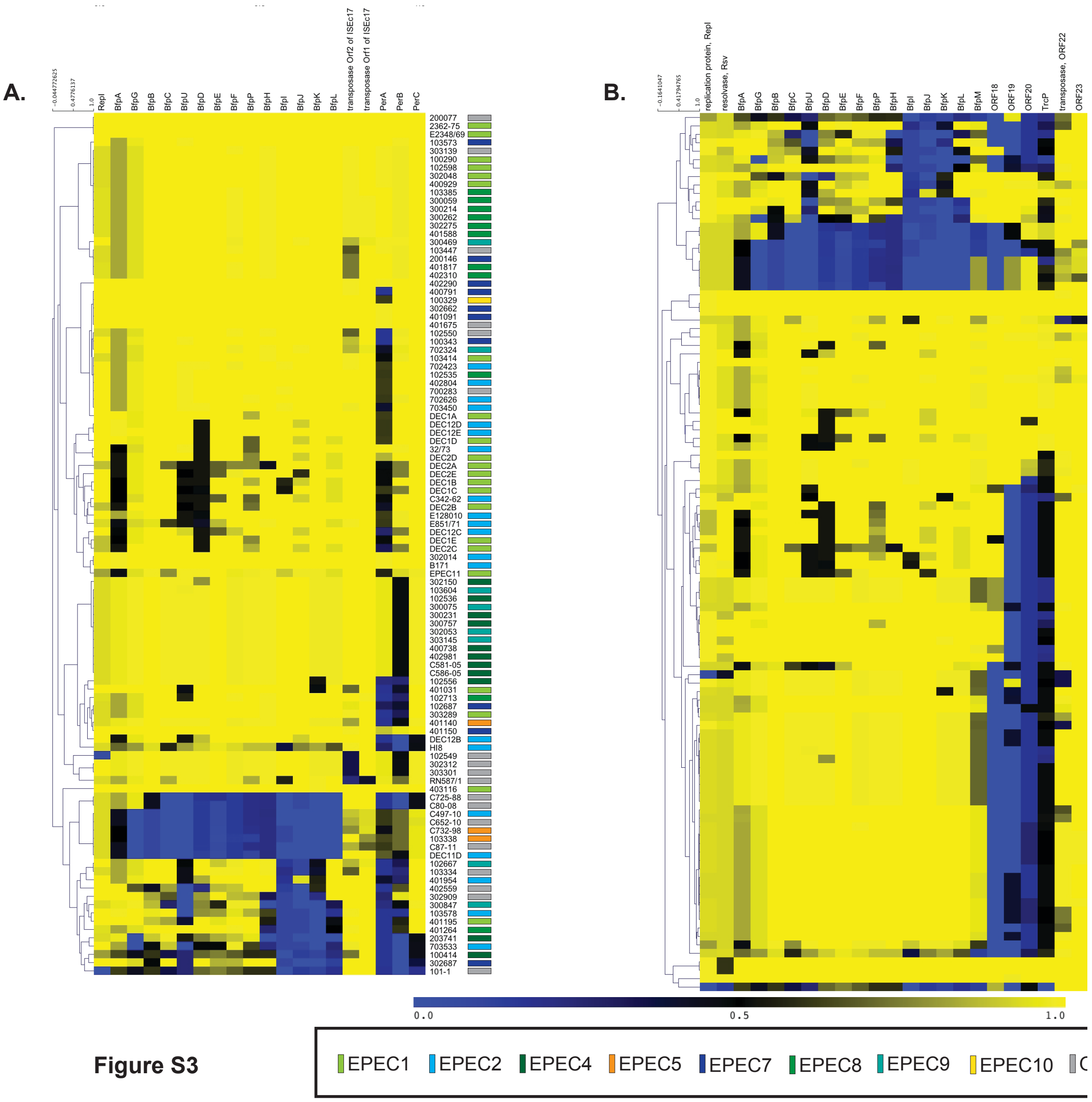
A.
-0.044772625
0.4776137
1.0
RepI
BfpA
BfpG
BfpB
BfpC
BfpU
BfpD
BfpE
BfpF
BfpP
BfpH
BfpI
BfpJ
BfpK
BfpL
transposase Orf2 of ISEc17
transposase Orf1 of ISEc17
PerA
PerB
PerC
200077
2362-75
E2348/69
103573
303139
100290
102598
302048
400929
103385
300059
300214
300262
302275
401588
300469
103447
200146
401817
402310
402290
400791
100329
302662
401091
401675
102550
100343
702324
103414
702423
102535
402804
700283
702626
703450
DEC1A
DEC12D
DEC12E
DEC1D
32/73
DEC2D
DEC2A
DEC2E
DEC1B
DEC1C
C342-62
DEC2B
E128010
E851/71
DEC12C
DEC1E
DEC2C
302014
B171
EPEC11
302150
103604
102536
300075
300231
300757
302053
303145
400738
402981
C581-05
C586-05
102556
401031
102713
102687
303289
401140
401150
DEC12B
HI8
102549
302312
303301
RN587/1
403116
C725-88
C80-08
C497-10
C652-10
C732-98
103338
C87-11
DEC11D
102667
103334
401954
402559
302909
300847
103578
401195
401264
203741
703533
100414
302687
101-1
B.
-0.1641047
0.41794765
1.0
replication protein, RepI
resolvase, Rsv
BfpA
BfpG
BfpB
BfpC
BfpU
BfpD
BfpE
BfpF
BfpP
BfpH
BfpI
BfpJ
BfpK
BfpL
BfpM
ORF18
ORF19
ORF20
TrcP
transposase, ORF22
ORF23
0.0
0.5
1.0
EPEC1
EPEC2
EPEC4
EPEC5
EPEC7
EPEC8
EPEC9
EPEC10

Figure S3

Figure S3. Hierarchical cluster analysis of TBLASTN BSR values demonstrating the similarity of protein-encoding genes of the BFP operons of A) pMAR2, and B) pB171 identified in EPEC genomes analyzed in this study. BLAST Score Ratio (BSR) analysis was performed as previously described (1) using the predicted amino acid sequences encoded by genes of the BFP operons of pMAR2 (A) and pB171 (B) compared to each AEEC genome analyzed in this study using TBLASTN (2). Each box represents the BSR value of a single gene in one genome. Each row contains BSR values for a single gene, while each column contains BSR values identified in a single genome. Hierarchical cluster analysis was performed using the Pearson correlation coefficient of MeV (3-5). The colors of the heatmap indicate the BSR values corresponding to the presence (yellow), divergence (black), or absence (blue) of each of the plasmid genes in each AEEC genome. The colored rectangle next to each EPEC isolate indicates the phylogenomic lineage the genome belongs to in Figure 1. Comparison of the genes in the BFP operon of plasmids pMAR2 and pB171 demonstrated that the BFP genes were highly conserved with the greatest sequence divergence observed in *bfpA* and the *per* genes. The additional genes within the BFP region of pB171 were identified primarily in other EPEC2 isolates, but were also present in a few isolates from other phylogenomic lineages.

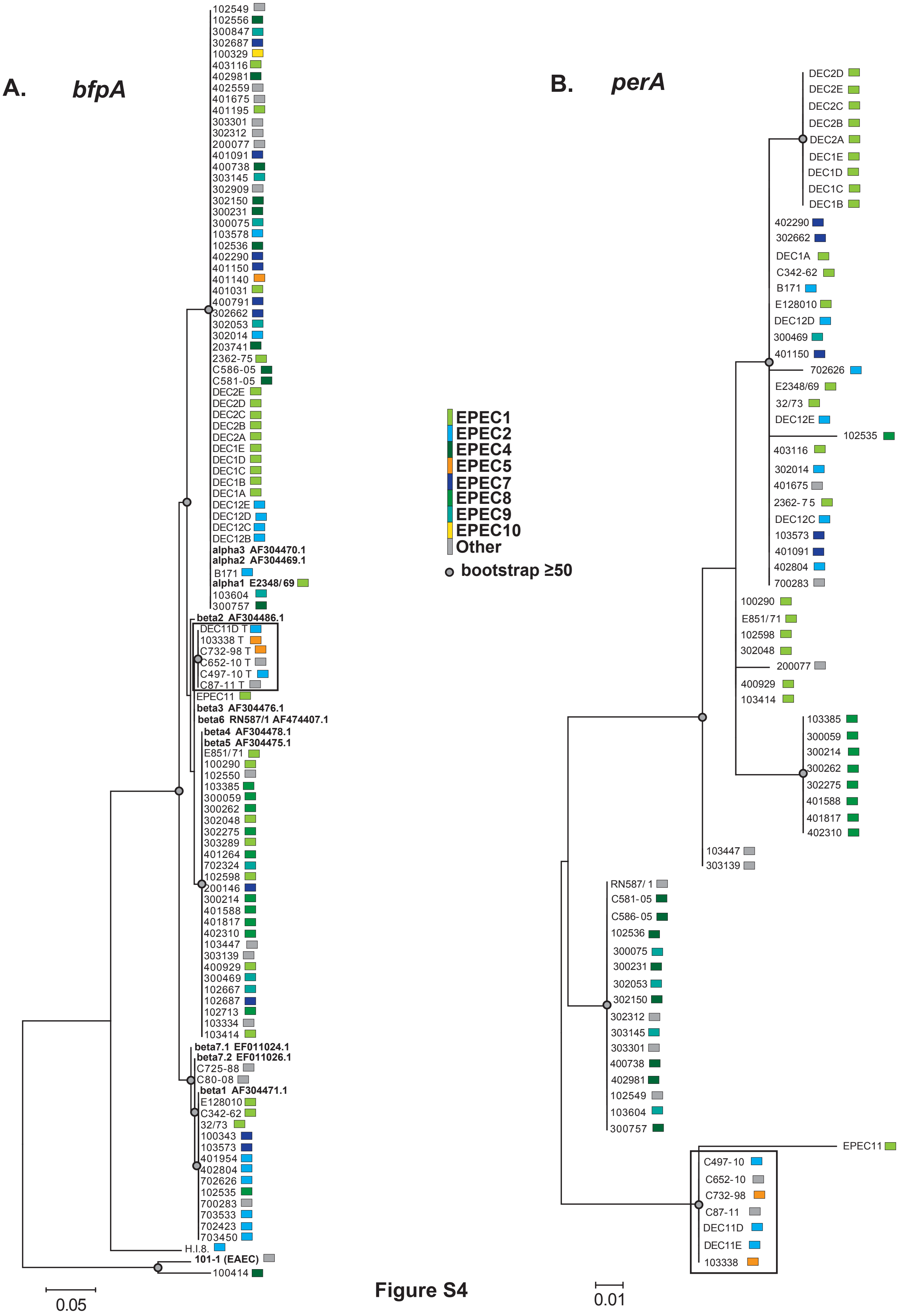
A.
bfpA
102549
102556
300847
302687
100329
403116
402981
402559
401675
401195
303301
302312
200077
401091
400738
303145
302909
302150
300231
300075
103578
102536
402290
401150
401140
401031
400791
302662
302053
302014
203741
2362-75
C586-05
C581-05
DEC2E
DEC2D
DEC2C
DEC2B
DEC2A
DEC1E
DEC1D
DEC1C
DEC1B
DEC1A
DEC12E
DEC12D
DEC12C
DEC12B
alpha3 AF304470.1
alpha2 AF304469.1
B171
alpha1 E2348/69
103604
300757
beta2 AF304486.1
DEC11D T
103338 T
C732-98 T
C652-10 T
C497-10 T
C87-11 T
EPEC11
beta3 AF304476.1
beta6 RN587/1 AF474407.1
beta4 AF304478.1
beta5 AF304475.1
E851/71
100290
102550
103385
300059
300262
302048
302275
303289
401264
702324
102598
200146
300214
401588
401817
402310
103447
303139
400929
300469
102667
102687
102713
103334
103414
beta7.1 EF011024.1
beta7.2 EF011026.1
C725-88
C80-08
beta1 AF304471.1
E128010
C342-62
32/73
100343
103573
401954
402804
702626
102535
700283
703533
702423
703450
H.I.8.
101-1 (EAEC)
100414
0.05
EPEC1
EPEC2
EPEC4
EPEC5
EPEC7
EPEC8
EPEC9
EPEC10
Other
bootstrap ≥50
B.
perA
DEC2D
DEC2E
DEC2C
DEC2B
DEC2A
DEC1E
DEC1D
DEC1C
DEC1B
402290
302662
DEC1A
C342-62
B171
E128010
DEC12D
300469
401150
702626
E2348/69
32/73
DEC12E
102535
403116
302014
401675
2362-75
DEC12C
103573
401091
402804
700283
100290
E851/71
102598
302048
200077
400929
103414
103385
300059
300214
300262
302275
401588
401817
402310
103447
303139
RN587/1
C581-05
C586-05
102536
300075
300231
302053
302150
302312
303145
303301
400738
402981
102549
103604
300757
EPEC11
C497-10
C652-10
C732-98
C87-11
DEC11D
DEC11E
103338
0.01

Figure S4

Figure S4. Phylogenetic analysis of the nucleotide sequences of A) *bfpA,* and B) *perA* that were identified in each of the AEEC isolates included in this study. A maximum likelihood phylogeny was constructed using the Kimura 2-parameter model and 1,000 bootstrap replicates. Bootstrap values ≥50 are designated by a circle. The scale bars represent the distance of the number of nucleotide substitutions per site. The genomes that have a truncated BFP operon are indicated by a "T", and also by a box. The colored rectangle next to each EPEC isolate indicates the phylogenomic lineage the genome belongs to in Figure 1. Phylogenetic analysis of *bfpA* and *perA* demonstrated that for most of the AEEC isolates analyzed, these sequences had less sequence divergence and lineage-specificity than previously observed for other virulence factors such as the T3SS effectors (10).

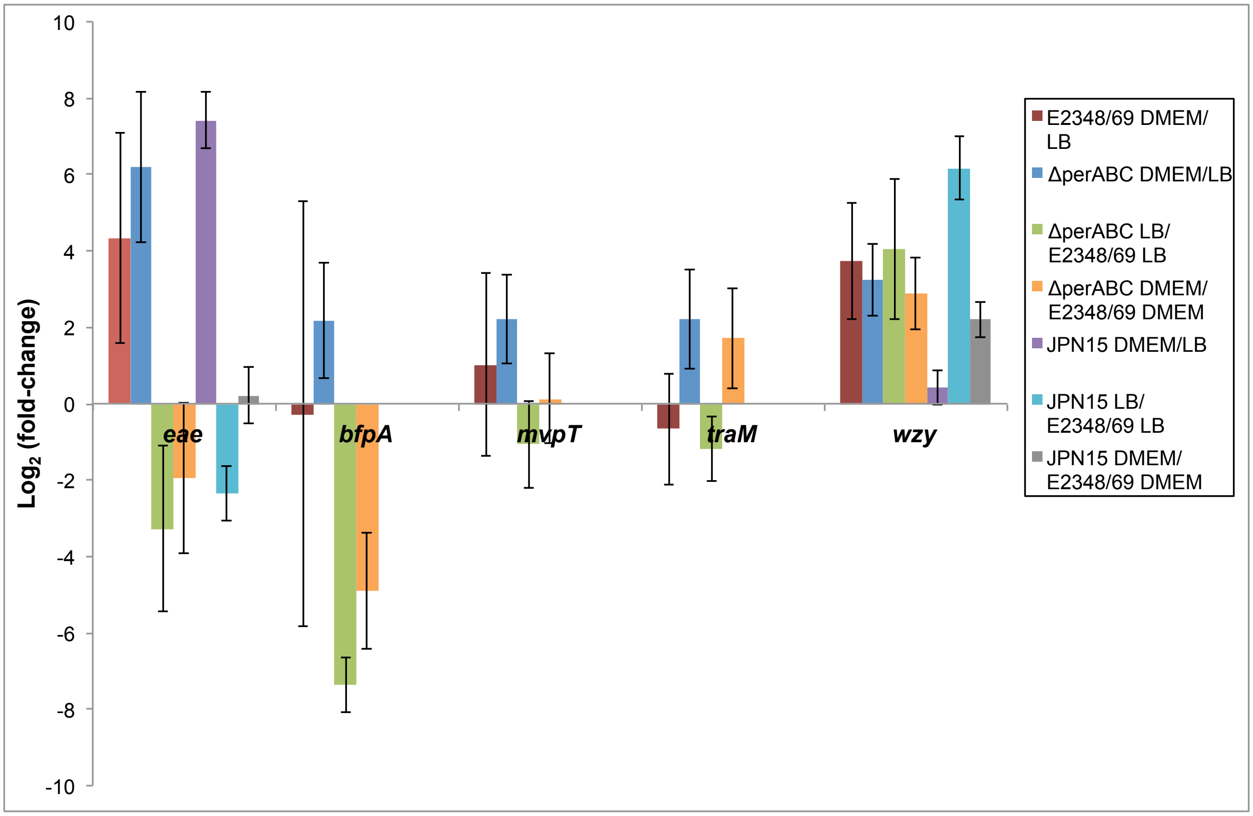

Figure S5. qRT-PCR analysis of select genes that were determined to be differentially-expressed using RNA-Seq analysis. Quantitative reverse transcription PCR (qRT-PCR) analysis of select genes for all RNA samples generated in this study. The values represent the $\log_2$ of the relative fold-difference ($2^{-\Delta\Delta Ct}$) of DMEM samples compared to LB samples for the wild-type E2348/69 and the two plasmid mutants, a Δ*perABC* strain (11), and JPN15 (12, 13). The fold-change of each gene was calculated for two biological replicates and three technical replicates for each condition. The error bars denote the standard deviation of the ΔΔCt values. Each gene was calibrated using the reference gene *rpoA.* The expression trends observed using qRT-PCR were consistent with those determined using RNA-Seq.

Table S1. Genomes analyzed in this study

Isolate ID	Serotype	Location	Inc-types (*In silico*)	Phylogenomic Lineage	Phylogroup	*stx1*	*stx2*	*ehxA*	*bfpA*	*perA*	Accession Number
EDL933	O157:H7	Wisconsin, USA	FIB, FIIA	EHEC1	E	+	+	+	-	-	NC_002655.2
Sakai	O157:H7	Japan	FIB, FIIA	EHEC1	E	+	+	+	-	-	NC_002695.1
EC4115	O157:H7	Maine, USA	FIB, FIIA, FIA	EHEC1	E	-	+	+	-	-	NC_011353.1
TW14359	O157:H7	Michigan, USA	FIB, FIIA, FIA	EHEC1	E	-	+	+	-	-	NC_013008.1
DEC3A	O157:H7	Washington, USA	FIB	EHEC1	E	+	+	+	-	-	AIFE00000000
DEC3B	O157:H7	Washington, USA	FIB, FIIA	EHEC1	E	+	+	+	-	-	AIFF00000000
DEC3C	O157:H7	New Mexico, USA	FIB, FIIA	EHEC1	E	+	+	+	-	-	AIFG00000000
DEC3D	O157:H7	Michigan, USA	FIB, FIIA	EHEC1	E	+	+	+	-	-	AIFH00000000
DEC3E	O157:H7	Canada	FIB, FIIA	EHEC1	E	-	+	+	-	-	AIFI00000000
DEC3F	O157:H7	Germany	FIB, FIIA, FIA	EHEC1	E	-	+	+	-	-	AIFJ00000000
DEC4A	O157:H7	Argentina	FIB, FIIA	EHEC1	E	-	-	+	-	-	AIFK00000000
DEC4B	O157:H7	Denmark	FIB, FIIA, FIA	EHEC1	E	-	+	+	-	-	AIFL00000000
DEC4C	O157:H7	Egypt	FIB, FIIA, FIA	EHEC1	E	-	-	+	-	-	AIFM00000000
DEC4D	O157:H7	Japan	FIB, FIIA, FIA	EHEC1	E	-	+	+	-	-	AIFN00000000
DEC4E	O157:H7	Denmark	FIB, FIIA	EHEC1	E	+	-	+	-	-	AIFO00000000
DEC4F	O157:H7	Michigan, USA	FIB, FIIA	EHEC1	E	+	+	+	-	-	AIFP00000000
86-24	O157:H7	Seattle, USA	FIB, FIIA	EHEC1	E	-	+	-	-	-	AIAK00000000
C842-97	O157:H7	Japan	FIIA	EHEC1	E	-	-	+	-	-	AIBY00000000
C354-03B	O157:H-	Great Britain	FIB, FIIA, FIA	EHEC1	E	-	-	+	-	-	AIAW00000000
CB9615	O55:H7	Germany	FIB	EHEC1	E	-	-	-	-	-	NC_013941.1
DEC5A	O55:H7	New York, USA	FIB	EHEC1	E	-	-	-	-	-	AIFQ00000000
DEC5B	O55:H7	Florida, USA	FIIA	EHEC1	E	-	-	-	-	-	AIFR00000000
DEC5C	O55:H7	New Jersey, USA	FIB	EHEC1	E	-	-	-	-	-	AIFS00000000
DEC5D	O55:H7	Sri Lanka	FIB	EHEC1	E	-	-	-	-	-	AIFT00000000
DEC5E	O55:H7	Iran	FIB, FIIA	EHEC1	E	-	-	-	-	-	AIFU00000000
C12-92	O55:H7	Great Britain	FIB	EHEC1	E	-	-	-	-	-	AIAE00000000
C1244-91	O55:H-	Great Britain	FIB	EHEC1	E	-	-	-	-	-	AIAD00000000
C154-11	O55:H7	Denmark	FIB	EHEC1	E	-	-	-	-	-	AIAF00000000
C157-11	O55:H-	Denmark	FIB, FIIA	EHEC1	E	-	-	-	-	-	AIAH00000000
11128	O111:H-	ND	FIIA, HI1, FIA	EHEC2	B1	+	+	+	-	-	NC_013364.1
11368	O26:H11	ND	FIB, FIIA, K/B/O	EHEC2	B1	+	-	+	-	-	NC_013361.1
OK1114	ND	Oklahoma, USA	FIIA	EHEC2	B1	+	+	+	-	-	AICG00000000
OK1180	ND	Oklahoma, USA	FIIA	EHEC2	B1	+	+	+	-	-	ADUQ00000000
EHEC14	ND	Brazil	FIB, K/B/O	EHEC2	B1	+	-	+	-	-	ADUN00000000
DEC8A	O111a:NM	Maryland, USA	FIB, FIIA, HI2	EHEC2	B1	+	-	+	-	-	AIGF00000000
DEC8B	O111:H8	Idaho, USA	FIIA	EHEC2	B1	+	+	+	-	-	AIGG00000000
DEC8C	O111:NM	South Dakota, USA	FIB, FIIA, K/B/O	EHEC2	B1	+	-	+	-	-	AIGH00000000
DEC8D	O111:H11	Cuba	FIB, FIIA, K/B/O	EHEC2	B1	-	-	+	-	-	AIGI00000000
DEC8E	O111:H8	Denmark	none	EHEC2	B1	+	-	-	-	-	AIGJ00000000
DEC9A	O26:H11	Wisconsin, USA	none	EHEC2	B1	-	-	-	-	-	AIGK00000000
DEC9B	O26:HN	New Hampshire, USA	none	EHEC2	B1	-	-	-	-	-	AIGL00000000
DEC9C	O26:HN	Switzerland	none	EHEC2	B1	-	-	-	-	-	AIGM00000000
DEC9D	O26:H11	Denmark	FIIA	EHEC2	B1	-	-	-	-	-	AIGN00000000
DEC9E	O26:H11	Mexico	none	EHEC2	B1	-	-	-	-	-	AIGO00000000
DEC10A	O26:H11	United Kingdom	none	EHEC2	B1	+	-	-	-	-	AIGP00000000
DEC10B	O26:H11	Australia	FIB, K/B/O	EHEC2	B1	+	-	-	-	-	AIGQ00000000
DEC10C	O26:H11	Delaware, USA	none	EHEC2	B1	+	-	-	-	-	AIGR00000000
DEC10D	O26:H11	France	FIB, FIIA, K/B/O	EHEC2	B1	-	-	+	-	-	AIGS00000000
DEC10F	O15	South Carolina, USA	FIB, FIIA, I1, FIA	EHEC2	B1	-	-	-	-	-	AIGU00000000
C1214-90	O26:H11	India	FIB, FIIA, K/B/O	EHEC2	B1	-	-	+	-	-	AIAC00000000
C792-92	O26:K-:H11	Italy	FIB, FIIA	EHEC2	B1	-	-	-	-	-	AIBR00000000
C799-92	O26:K-:H-	Italy	FIB, FIIA, I1	EHEC2	B1	-	-	-	-	-	AIBT00000000
C347-93	O26:K-:H-	Thailand	FIIA	EHEC2	B1	-	-	-	-	-	AIAU00000000
C654-09	O26:H-	France	FIB, I1, Y	EHEC2	B1	-	-	-	-	-	AIBJ00000000
C717-10	O103:H-	Denmark	FIB, FIIA, I1, Y	EHEC2	B1	-	-	-	-	-	AIBK00000000
C824-10	O128abc:H8	Denmark	FIB, FIIA	EHEC2	B1	-	-	-	-	-	AIBX00000000
C155-11	O26:H11	Denmark	FIB, FIIA, K/B/O	EHEC2	B1	-	-	+	-	-	AIAG00000000
C166-11	O26:H-	Denmark	I1	EHEC2	B1	-	-	-	-	-	AICF00000000
E2348/69	O127:H6	Taunton, England	FIB, FIIA	EPEC1	B2	-	-	-	+	+	NC_011601.1
2362-75	O55:H6	New Mexico, USA	FIB, FIIA	EPEC1	B2	-	-	-	+	+	ADUL01000000
E851/71	O142:H6	Scotland	FIB, FIIA, Y	EPEC1	B2	-	-	-	+	+	ALNX00000000
EPEC11	NK	Brazil	FIB, FIIA, FIA	EPEC1	B2	-	-	-	+	T	AKYH00000000
DEC1A	O55:H6	Pennsylvania, USA	FIB, FIIA	EPEC1	B2	-	-	-	+	+	AIEV00000000
DEC1B	O55:H6	Dutch Guiana	FIB, FIIA	EPEC1	B2	-	-	-	+	+	AIEW00000000
DEC1C	O55:H6	Germany	FIB, FIIA	EPEC1	B2	-	-	-	+	+	AIEX00000000
DEC1D	O55:H6	Egypt	FIB	EPEC1	B2	-	-	-	+	+	AIEY00000000
DEC1E	O55:H6	Mexico	FIB, FIIA	EPEC1	B2	-	-	-	+	T	AIEZ00000000
DEC2A	O55:H6	Congo	FIB, FIIA	EPEC1	B2	-	-	-	+	+	AIFA00000000
DEC2B	O55:NM	Pennsylvania, USA	FIB, FIIA	EPEC1	B2	-	-	-	+	+	AFJB00000000
DEC2C	O55:H6	Pennsylvania, USA	FIB, FIIA, Y, K/B/O	EPEC1	B2	-	-	-	+	+	AIFB00000000
DEC2D	O55:H6	France	FIB, FIIA	EPEC1	B2	-	-	-	+	+	AIFC00000000
DEC2E	O55:H6	Texas, USA	FIB	EPEC1	B2	-	-	-	+	+	AIFD00000000
100290	ND	The Gambia	FIB, FIIA	EPEC1	B2	-	-	-	+	+	JHQV00000000
302048	ND	Mozambique	FIB, FIIA	EPEC1	B2	-	-	-	+	+	JHRF00000000
303289	ND	Mozambique	FIB, FIIA	EPEC1	B2	-	-	-	+	-	JHRJ00000000
401031	ND	Kenya	FIB, FIIA	EPEC1	B2	-	-	-	+	-	JHRL00000000
102598	ND	The Gambia	FIB, FIIA	EPEC1	B2	-	-	-	+	+	JHRW00000000
401195	ND	Kenya	FIB, FIIA, P	EPEC1	B2	-	-	-	+	-	JHTF00000000
403116	ND	Kenya	FIB, FIIA, P	EPEC1	B2	-	-	-	+	+	JHTJ00000000
400929	ND	Kenya	FIB, FIIA, P	EPEC1	B2	-	-	-	+	+	JHTE00000000
B171	O111:NM	Washington D.C., USA	FIB, FIIA, P	EPEC2	B1	-	-	-	+	+	AAJX00000000

32/73	O126:H2	England	FIB, FIIA	EPEC2	B1	-	-	-	+	+	LAFA00000000
C342-62	O126:H2	Egypt	FIB, FIIA	EPEC2	B1	-	-	-	+	+	AKNI00000000
E128010	O114:H2	Bangladesh	FIB, FIIA, FIA	EPEC2	B1	-	-	-	+	+	ADUO00000000
DEC11A	O128a:H2	Montana, USA	FIB, FIIA, FIC	EPEC2	B1	-	-	-	-	-	AIGV00000000
DEC11B	O128a:H3	Missouri, USA	none	EPEC2	B1	-	-	-	-	-	AIGW00000000
DEC11C	O45:H2	Michigan, USA	FIB	EPEC2	B1	+	-	+	-	-	AIGX00000000
DEC11D	O128:H2	United Kingdom	FIB, FIIA, FIA	EPEC2	B1	-	-	-	T	+	AIGY00000000
DEC11E	O128:H2	Brazil	FIIA	EPEC2	B1	-	-	-	-	+	AIGZ00000000
DEC12A	O111:H2	United Kingdom	FIB	EPEC2	B1	-	-	-	-	-	AIHA00000000
DEC12B	O111:H2	Florida, USA	FIB, FIIA, Y	EPEC2	B1	-	-	-	+	-	AIHB00000000
DEC12C	O111:NM	Panama	FIB, FIIA	EPEC2	B1	-	-	-	+	T	AIHC00000000
DEC12D	O111:H2	Peru	FIB, FIIA	EPEC2	B1	-	-	-	+	+	AIHD00000000
DEC12E	O111:HN	Kenya	FIB, FIIA, I1	EPEC2	B1	-	-	-	+	+	AIHE00000000
E22	O103:H2	ND	FIB, FIIA, FIC	EPEC2	B1	-	-	-	-	-	AAJV00000000
H.I.8	O128:H-	ND	FIB, FIIA	EPEC2	B1	+	-	-	+	-	AFDY00000000
12009	O103:H2	ND	FIB	EPEC2	B1	+	+	+	-	-	NC_013353.1
C900-01	O103:H2	Guinea-Bissau	FIB	EPEC2	B1	-	-	+	-	-	AICC00000000
C488-07	O128abc:H-	Denmark	FIB, FIIA, FIA	EPEC2	B1	-	-	-	-	-	AIBA00000000
C341-10	O128ab:H-	Denmark	FIB, FIIA	EPEC2	B1	-	-	-	-	-	AIAS00000000
C497-10	O119:H2	Denmark	FIB, FIIA, FIA	EPEC2	B1	-	-	-	T	+	AIBC00000000
C82-11	O128abc:H2	Denmark	FIB, FIIA, FIC, FIA	EPEC2	B1	-	-	-	-	-	AIBW00000000
C170-11	O128ab:H2	Denmark	FIB, FIIA, FIC, FIA	EPEC2	B1	-	-	-	-	-	AIAJ00000000
302014	ND	Mozambique	FIB, FIIA, P	EPEC2	B1	-	-	-	+	+	JHRE00000000
401954	ND	Kenya	FIB, FIIA, FIA	EPEC2	B1	-	-	-	+	-	JHRP00000000
703533	ND	Pakistan	FIB, FIIA	EPEC2	B1	-	-	-	+	-	JHRS00000000
103578	ND	The Gambia	FIB, FIIA	EPEC2	B1	-	-	-	+	-	JHRY00000000
402804	ND	Kenya	FIB, FIIA, FIA	EPEC2	B1	-	-	-	+	+	JHSN00000000
702423	ND	Pakistan	FIB, FIIA	EPEC2	B1	-	-	-	+	-	JHSO00000000
702626	ND	Pakistan	FIB, FIIA, FIA, K/B/O	EPEC2	B1	-	-	-	+	+	JHSP00000000
703450	ND	Pakistan	FIB, FIIA	EPEC2	B1	-	-	-	+	-	JHTM00000000
C2139-99	O111:H7	Denmark	FIIA, I1	EPEC4	B2	-	-	-	-	-	AIAM00000000
C581-05	O119:H6	Africa	FIB, FIIA	EPEC4	B2	-	-	-	+	+	AIBE00000000
C586-05	O119:H-	Africa	FIB, FIIA	EPEC4	B2	-	-	-	+	+	AIBG00000000
C79-08	O33:H6	Slovenia	none	EPEC4	B2	-	-	-	-	-	AIBQ00000000
C887-10	O113:H6	Denmark	FIIA, I1	EPEC4	B2	-	-	-	-	-	AICB00000000
100414	ND	The Gambia	FIB, FIIA	EPEC4	B2	-	-	-	+	-	JHQX00000000
203741	ND	Mali	FIB, FIIA, Y, P	EPEC4	B2	-	-	-	+	-	JHRB00000000
102536	ND	The Gambia	FIB, FIIA, Y, P	EPEC4	B2	-	-	-	+	+	JHRV00000000
302150	ND	Mozambique	FIB, FIIA	EPEC4	B2	-	-	-	+	+	JHSD00000000
300231	ND	Mozambique	FIB, FIIA	EPEC4	B2	-	-	-	+	+	JHSC00000000
400738	ND	Kenya	FIB, FIIA, P	EPEC4	B2	-	-	-	+	+	JHSH00000000
302137	ND	Mozambique	none	EPEC4	B2	-	-	-	-	-	JHSZ00000000
402981	ND	Kenya	FIB, FIIA, P	EPEC4	B2	-	-	-	+	+	JHTI00000000
401140	ND	Kenya	FIB, FIIA, P	EPEC5	B1	-	-	-	+	-	JHRM00000000
401210	ND	Kenya	FIB, FIIA, P	EPEC5	B1	-	-	-	-	-	JHSJ00000000
103338	ND	The Gambia	FIB, FIIA, FIA	EPEC5	B1	-	-	-	T	+	JHSU00000000
C732-98	O88:H-	Israel	FIB, FIIA, FIA	EPEC5	A	-	-	-	T	+	AIBM00000000
C751-03	O49:H-	Austria	FIIA	EPEC5	A	-	-	+	-	-	AIBO00000000
100343	ND	The Gambia	FIB, FIIA	EPEC7	B1	-	-	-	+	-	JHQW00000000
103573	ND	The Gambia	FIB, FIIA	EPEC7	B1	-	-	-	+	+	JHRA00000000
302662	ND	Mozambique	FIB, FIIA, P	EPEC7	B1	-	-	-	+	+	JHRI00000000
400791	ND	Kenya	FIB, FIIA, P	EPEC7	B1	-	-	-	+	-	JHRK00000000
402290	ND	Kenya	FIB, FIIA, P	EPEC7	B1	-	-	-	+	+	JHRQ00000000
401150	ND	Kenya	FIB, FIIA, P	EPEC7	B1	-	-	-	+	T	JHRN00000000
200146	ND	Mali	FIB, FIIA	EPEC7	B1	-	-	-	+	+	JHRZ00000000
302687	ND	Mozambique	FIB, FIIA	EPEC7	B1	-	-	-	+	-	JHSE00000000
401091	ND	Kenya	FIB, FIIA, P	EPEC7	B1	-	-	-	+	+	JHSI00000000
303341	ND	Mozambique	none	EPEC7	B1	-	-	-	-	-	JHTD00000000
403341	ND	Kenya	FIIA	EPEC7	B1	-	-	-	-	-	JHTK00000000
C527-94	O153:H12	Denmark	FIB, FIIA, FIC	EPEC7	B1	-	-	-	-	-	AIBD00000000
103385	ND	The Gambia	FIB, FIIA	EPEC8	B2	-	-	-	+	+	JHQZ00000000
300059	ND	Mozambique	FIB, FIIA, P	EPEC8	B2	-	-	-	+	+	JHRC00000000
300262	ND	Mozambique	FIB, FIIA, P	EPEC8	B2	-	-	-	+	+	JHRD00000000
302275	ND	Mozambique	FIB, FIIA, P	EPEC8	B2	-	-	-	+	+	JHRH00000000
401264	ND	Kenya	FIB, FIIA, P	EPEC8	B2	-	-	-	+	-	JHRO00000000
300214	ND	Mozambique	FIB, FIIA, P	EPEC8	B2	-	-	-	+	+	JHSB00000000
402310	ND	Kenya	FIB, FIIA	EPEC8	B2	-	-	-	+	+	JHSM00000000
401817	ND	Kenya	FIB, FIIA	EPEC8	B2	-	-	-	+	+	JHSL00000000
401588	ND	Kenya	FIB, FIIA, P	EPEC8	B2	-	-	-	+	+	JHSK00000000
102535	ND	The Gambia	FIB, FIIA	EPEC8	B2	-	-	-	+	+	JHST00000000
302053	ND	Mozambique	FIB, FIIA	EPEC9	B2	-	-	-	+	+	JHRG00000000
702324	ND	Pakistan	FIB, FIIA, FIA, P	EPEC9	B2	-	-	-	+	-	JHRR00000000
303145	ND	Mozambique	FIB, FIIA	EPEC9	B2	-	-	-	+	+	JHSG00000000
300075	ND	Mozambique	FIB, FIIA	EPEC9	B2	-	-	-	+	+	JHSA00000000
300847	ND	Mozambique	FIB, FIIA	EPEC9	B2	-	-	-	+	-	JHSY00000000
300469	ND	Mozambique	FIB, FIIA, FIA	EPEC9	B2	-	-	-	+	-	JHSX00000000
C796-10	O142:H34	Denmark	FIIA	EPEC9	B2	-	-	-	-	-	AIBS00000000
100329	ND	The Gambia	FIB, FIIA	EPEC10	B1	-	-	-	+	-	JHRT00000000
100854	ND	The Gambia	FIIA, P	EPEC10	B1	-	-	-	-	-	JHRU00000000
C260-92	O127:K-:H40	Italy	FIB, FIIA, K/B/O	EPEC10	A	-	-	-	-	-	AIAO00000000
C213-10	O127:H40	Denmark	FIB, FIIA, FIA	EPEC10	A	-	-	-	-	-	AIAL00000000
C496-10	O145:H40	Denmark	FIIA	EPEC10	A	-	-	-	-	-	AIBB00000000
102549	ND	The Gambia	FIIA, P	NC	B2	-	-	-	+	+	JZXE00000000
103334	ND	The Gambia	FIB, FIIA, FIC, P	NC	B1	-	-	-	+	-	JZXF00000000

102687	ND	The Gambia	FIB, FIIA	EPEC7	B1	-	-	-	+	-	JZXG00000000
102556	ND	The Gambia	FIB, FIIA, P, Y	EPEC4	B2	-	-	-	+	-	JZXH00000000
103414	ND	The Gambia	FIB, FIIA, Y	EPEC1	B2	-	-	-	+	T	JZXI00000000
102667	ND	The Gambia	FIB, FIIA, P, FIA	EPEC9	B2	-	-	-	+	-	JZXJ00000000
102713	ND	The Gambia	FIB, FIIA, P	EPEC8	B2	-	-	-	+	-	JZXK00000000
103604	ND	The Gambia	FIB, FIIA, I1	EPEC9	B2	-	-	-	+	+	JZXL00000000
300757	ND	Mozambique	FIB, FIIA	EPEC4	B2	-	-	-	+	+	JZXM00000000
102550	ND	The Gambia	FIB, FIIA, P, FIC	NC	B1	-	-	-	+	-	JHQY00000000
102929	ND	The Gambia	FIB, FIIA	NC	B1	-	-	-	-	-	JHRX00000000
302909	ND	Mozambique	FIB, FIIA	NC	B2	-	-	-	+	-	JHSF00000000
100175	ND	The Gambia	FIB	NC	B1	-	-	-	-	-	JHSR00000000
102132	ND	The Gambia	FIB, I1	NC	E	-	-	+	-	-	JHSS00000000
103447	ND	The Gambia	FIB, FIIA, P, FIC	NC	B1	-	-	-	+	+	JHSV00000000
100100	ND	The Gambia	FIB	NC	E	-	-	+	-	-	JHSQ00000000
200077	ND	Mali	FIB, FIIA	NC	B2	-	-	-	+	+	JHSW00000000
303301	ND	Mozambique	FIB, FIIA, P	NC	B2	-	-	-	+	+	JHTC00000000
303139	ND	Mozambique	FIB, FIIA, FIC	NC	B1	-	-	-	+	+	JHTB00000000
302312	ND	Mozambique	FIB, FIIA, P	NC	B2	-	-	-	+	+	JHTA00000000
402559	ND	Kenya	FIB, FIIA	NC	B2	-	-	-	+	-	JHTH00000000
401675	ND	Kenya	FIB, FIIA	NC	B2	-	-	-	+	+	JHTG00000000
700283	ND	Pakistan	FIB, FIIA, P, FIA	NC	B2	-	-	-	+	+	JHTL00000000
RN587/1	O157:H8	Brazil	FIB, FIIA, FIA	NC	B2	-	-	-	+	+	ADUS00000000
C9-92	O127:H45	Great Britain	I1	NC	B2	-	-	-	-	-	AICE00000000
C844-97	O157:H45	Japan	none	NC	B2	-	-	-	-	-	AIBZ00000000
C639-08	O157:H45	Denmark	FIB, FIIA	NC	B2	-	-	-	-	-	AIBH00000000
E110019	O111:H9	Finland	FIB, FIIA	NC	B1	-	-	-	-	-	AAJW00000000
C725-88	O86:H6	Germany	FIB, FIIA	NC	B1	-	-	-	+	-	AIBL00000000
C418-89	O125ac:H6	Italy	none	NC	B2	-	-	-	-	-	AIAY00000000
C238-91	O103:H25	Germany	FIB	NC	B1	-	-	+	-	-	AIAN00000000
C743-03	O111:H19	Austria	FIB	NC	B1	-	-	-	-	-	AIBN00000000
C80-08	O98:H8	Slovenia	FIB, FIIA	NC	B1	-	-	-	+	-	AIBU00000000
C343-08	O111:H38	Denmark	FIB	NC	B1	-	-	-	-	-	AIAT00000000
C78-09C	O132:H34	Belgium	none	NC	B2	-	-	-	-	-	AIBP00000000
C283-09	O177:H-	Slovenia	FIB, FIIA	NC	B1	-	-	+	-	-	AIAQ00000000
C353-09	O167:H9	Denmark	none	NC	D	-	-	-	-	-	AIAV00000000
C807-09	O3:H8	Denmark	none	NC	B1	-	-	-	-	-	AIBV00000000
C262-10	O114:H49	Denmark	none	NC	B2	-	-	-	-	-	AIAP00000000
C295-10	O55:H34	Denmark	I1	NC	B2	-	-	-	-	-	AIAR00000000
C458-10	O153:H21	Denmark	none	NC	B1	-	-	-	-	-	AIAZ00000000
C652-10	O88:H25	Denmark	FIB, FIIA, FIA	NC	B1	-	-	-	T	+	AIBI00000000
C40-11	O145:H25	Denmark	FIB, FIIA, K/B/O	NC	B1	-	-	+	-	-	AIAX00000000
C58-11	O145:H34	Denmark	Y	NC	B2	-	-	-	-	-	AIBF00000000
C87-11	O111:H8	Denmark	FIB, FIIA, FIA, Y	NC	B1	-	-	-	T	+	AICA00000000
C93-11	O157:H-	Denmark	FIB	NC	B2	-	-	-	-	-	AICD00000000
C161-11	O145:H-	Denmark	FIB, FIIA, K/B/O	NC	E	-	-	+	-	-	AIAI00000000

E. coli and *Shigella* reference genomes included in the phylogenomic analysis

Isolate ID	Pathovar	Phylogroup	Accession Number
53638	EIEC	A	AAKB00000000.2
HS	commensal	A	NC_009800.1
ATCC 8739	lab adapted	A	NC_010468.1
BL21	lab adapted	A	NC_012947.1
BW2952	lab adapted	A	NC_012759.1
H10407	ETEC	A	NC_017633.1
S. flexneri 2A 2457T	*Shigella*	B1	NC_004741.1
S. sonnei 046	*Shigella*	B1	NC_007384.1
S. boydii 3083-94	*Shigella*	B1	NC_010658.1
B7A	ETEC	B1	AAJT00000000.2
E24377A	ETEC	B1	NC_009801.1
IAI1	fecal isolate	B1	NC_011741.1
SE11	fecal isolate	B1	NC_011415.1
CFT073	ExPEC	B2	NC_004431.1
S88	ExPEC	B2	NC_011742.1
UTI89	ExPEC	B2	NC_007946.1
536	ExPEC	B2	NC_008253.1
TY-2482	EAEC	B1	AFOG00000000.1
55989	EAEC	B1	NC_011748.1
042	EAEC	D	FN554766.1
101-1	EAEC	A	AAMK00000000.2
UMN026	ExPEC	D	NC_011751.1
S. dysenteriae Sd197	*Shigella*	E	NC_007606.1
IAI39	ExPEC	F	NC_011750.1
SMS-3-5	other	F	NC_010498.1

Table S2. Primers used in this study

Gene	Predicted Protein Function	Protein Accession No.[a]	Primer ID	Sequence 5'-3'	Source
rpoA	RNA polymerase, alpha subunit	NA	Ec-rpoA-F	TGTAGGCAATACGCTCCACA	(14)
			Ec-rpoA-R	GGTTATGTGCCGGCTTCTAC	"
eae	Intimin, attachment protein	YP_002331401.1	eae_361F	TATAGTGCCTTACCTGTCTTAGGTTCGG	(15)
			eae_495R	ATAATTTAGAGCCTTGTCATCGGTCGTG	"
bfpA	major subunit of the bundle-forming pilus	YP_002332157.1	bfpA_118F	GCGTCTGATTCCAATAAGTCGCAGAATG	"
			bfpA_380R	GCTTTATCCAACCTGGTAAGCGTCAGAT	"
perA	transcriptional activator	YP_002332173.1	perA_332F	CTGCGAACCTCAATGAAATGCAATTG	"
			perA_506R	CTGTCTACGATGCTCTTAGATGAAGCAC	"
rbsD	D-ribose pyranase	YP_002331517.1	rbsD_57F	ACATACCGATACGCTGGTGGTGTGT	"
			rbsD_253R	GCAACGTTTCGTGGAGTTGCGGATTA	"
nirC	nitrite transporter	YP_002331085.1	nirC_190F	ACCTTTGGTATCGCCTTAACGCTGG	"
			nirC_357R	AACGAAGACGGAACCGACCAGGTTAC	"
mvpT	post-segregational killing toxin	YP_002332186.1	mvpT_107F	CAGGAACTGCTTCAATGCTTCTTGAATTAGG	This study
			mvpT_286R	CGGATACATGAGGACTGAGAGACTCAATACC	"
traM	conjugal transfer protein	YP_002332219.1	traM_101F	GCCTTTCAGGAACTGCTTCAATGCTTCTTG	"
			traM_287R	CCGGATACATGAGGACTGAGAGACTCAATACC	"
wzy	O-antigen polymerase	YP_002329685.1	wzy_E2348_859F	CCAGCGTCACCAATTAGCTTGGAG	"
			wzy_E2348_1039R	CATCGCAAATATGCAGGCCATAATGAC	"

[a] Primers were selected by comparing with the sequences from all EPEC genomes analyzed; however, the protein accession number is provided for the respective gene from the E2348/69 genome.

Data Sets

Data Set 1. Differentially-expressed genes identified by RNA-Seq analysis of E2348/69 grown in DMEM compared to LB

Data Set 2. Differentially-expressed genes identified by RNA-Seq analysis of E2348/69 compared to *perABC* grown in LB

Data Set 3. Differentially-expressed genes identified by RNA-Seq analysis of E2348/69 compared to *perABC* grown in DMEM

Data Set 4. Differentially-expressed genes identified by RNA-Seq analysis of *perABC* grown in DMEM compared to LB

Data Set 5. Differentially-expressed genes identified by RNA-Seq analysis of JPN15 compared to *perABC* grown in LB

Data Set 6. Differentially-expressed genes identified by RNA-Seq analysis of JPN15 compared to *perABC* grown in DMEM

Data Set 7. Differentially-expressed genes identified by RNA-Seq analysis of E2348/69 compared to JPN15 grown in LB

Data Set 8. Differentially-expressed genes identified by RNA-Seq analysis of E2348/69 compared to JPN15 grown in DMEM

Data Set 9. Differentially-expressed genes identified by RNA-Seq analysis of JPN15 grown in DMEM compared to LB

References

1. **Rasko DA, Myers GS, Ravel J.** 2005. Visualization of comparative genomic analyses by BLAST score ratio. BMC Bioinf **6:**2.
2. **Gertz EM, Yu YK, Agarwala R, Schaffer AA, Altschul SF.** 2006. Composition-based statistics and translated nucleotide searches: improving the TBLASTN module of BLAST. BMC Biol **4:**41.
3. **Eisen MB, Spellman PT, Brown PO, Botstein D.** 1998. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci USA **95:**14863-14868.
4. **Saeed AI, Bhagabati NK, Braisted JC, Liang W, Sharov V, Howe EA, Li J, Thiagarajan M, White JA, Quackenbush J.** 2006. TM4 microarray software suite. Methods Enzymol **411:**134-193.
5. **Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N, Braisted J, Klapa M, Currier T, Thiagarajan M, Sturn A, Snuffin M, Rezantsev A, Popov D, Ryltsov A, Kostukovich E, Borisovsky I, Liu Z, Vinsavich A, Trush V, Quackenbush J.** 2003. TM4: a free, open-source system for microarray data management and analysis. Biotechniques **34:**374-378.
6. **Lacher DW, Steinsland H, Blank TE, Donnenberg MS, Whittam TS.** 2007. Molecular evolution of typical enteropathogenic *Escherichia coli*: clonal analysis by multilocus sequence typing and virulence gene allelic profiling. J Bacteriol **189:**342-350.
7. **Ogura Y, Ooka T, Iguchi A, Toh H, Asadulghani M, Oshima K, Kodama T, Abe H, Nakayama K, Kurokawa K, Tobe T, Hattori M, Hayashi T.** 2009. Comparative genomics reveal the mechanism of the parallel evolution of O157 and non-O157 enterohemorrhagic *Escherichia coli*. Proc Natl Acad Sci USA **106:**17939-17944.
8. **Brinkley C, Burland V, Keller R, Rose DJ, Boutin AT, Klink SA, Blattner FR, Kaper JB.** 2006. Nucleotide sequence analysis of the enteropathogenic *Escherichia coli* adherence factor plasmid pMAR7. Infect. Immun. **74:**5408-5413.
9. **Iguchi A, Thomson NR, Ogura Y, Saunders D, Ooka T, Henderson IR, Harris D, Asadulghani M, Kurokawa K, Dean P, Kenny B, Quail MA, Thurston S, Dougan G, Hayashi T, Parkhill J, Frankel G.** 2009. Complete genome sequence and comparative genome analysis of enteropathogenic *Escherichia coli* O127:H6 strain E2348/69. J Bacteriol **191:**347-354.
10. **Hazen TH, Sahl JW, Fraser CM, Donnenberg MS, Scheutz F, Rasko DA.** 2013. Refining the pathovar paradigm via phylogenomics of the attaching and effacing *Escherichia coli*. Proc Natl Acad Sci USA **110:**12810-12815.
11. **Leverton LQ.** 2004. An examination of enteropathogenic *Escherichia coli* virulence gene regulation. University of Maryland, Baltimore.
12. **Jerse AE, Yu J, Tall BD, Kaper JB.** 1990. A genetic locus of enteropathogenic *Escherichia coli* necessary for the production of attaching and effacing lesions on tissue culture cells. Proc Natl Acad Sci USA **87:**7839-7843.
13. **Levine MM, Nataro JP, Karch H, Baldini MM, Kaper JB, Black RE, Clements ML, O'Brien AD.** 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.

14. **Sahl JW, Rasko DA.** 2012. Analysis of Global Transcriptional Profiles of Enterotoxigenic *Escherichia coli* Isolate E24377A. Infect. Immun. **80:**1232-1242.
15. **Hazen TH, Daugherty SC, Shetty A, Mahurkar AA, White O, Kaper JB, Rasko DA.** 2015. RNA-Seq analysis of isolate- and growth phase-specific differences in the global transcriptomes of enteropathogenic *Escherichia coli* prototype isolates. Front Microbiol **6**.