

Table S1. *E. coli* virulence genes¹.

VFclass	Virulence factors	Genes2	Strain/Pathotype	GenBank Accession/Replicon	Locus ID	Genomic subset ³	EDL933	CFSAN004176	CFSAN004177	RM13514	RM13516	10942	112648	122715	95- 3192 ⁴	2015C- 0315	RM8843- C1	RM8988- C1	RM8995- C1	RM9154- C1	RM9467- C1	RM9872- C1	RM9873- C1	RM10425- C1	RM11626- C1	RM12275- C1	RM12367- C1	RM12522- C8		
Adherence	AAF/II fimbriae	<i>aafA</i>	<i>E. coli</i> O44:H18 str. 042/EAEC	NC_017627/Plasmid pAA	EC042_pAA048	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
		<i>aafB</i>			EC042_pAA030	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
		<i>aafC</i>			EC042_pAA031	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
		<i>aafD</i>			EC042_pAA046	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
	AAF/III fimbriae	<i>aggA</i>	<i>E. coli</i> O104:H4 str. 2011C-3493 /StxEAEC	NC_018666 /Plasmid pAA-EA11	O3K_26197	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
		<i>aggB</i>			O3K_26202	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
		<i>aggC</i>			O3K_26207	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
		<i>aggD</i>			O3K_26212	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
		<i>aggR</i>			O3K_26097	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
	Afimbrial adhesin AFA-I	<i>afaA</i>	<i>E. coli</i> str. VR50 /UPEC	NZ_CP011134/Chromosome	ECVR50_3238	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
		<i>afaB</i>			ECVR50_3237	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
		<i>afaC</i>			ECVR50_3236	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
		<i>afaD</i>			ECVR50_3235	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
		<i>afaE</i>			ECVR50_3233	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
		<i>draP</i>			ECVR50_3234	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	CFA/I fimbriae	<i>cfaA</i>	<i>E. coli</i> O103:H2 str. 12009/EHEC	NC_013353/Chromosome	ECO103_3842	A	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
		<i>cfaB</i>			ECO103_3841	A	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
		<i>cfaC</i>			ECO103_3840	A	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
		<i>cfaD/cfaE</i>			ECO103_3839	A	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	Curli fimbriae	<i>csgD</i>	<i>E. coli</i> O157:H7 str. EDL933/EHEC	NC_002655/Chromosome	Z1673	C	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
		<i>csgE</i>			Z1672	C	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
		<i>csgF</i>			Z1671	C	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
		<i>csgG</i>			Z1670	C	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
		<i>csgA</i>			Z1676	C	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
		<i>csgB</i>			Z1675	C	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
		<i>csgC</i>			Z1677	C	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
	Dispersin	<i>aap</i>	<i>E. coli</i> O44:H18 str. 042/EAEC	NC_017627/Plasmid pAA	EC042_pAA055	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
	<i>E. coli</i> common pilus (ECP)	<i>ecpA</i>	<i>E. coli</i> O104:H4 str. 2011C-3493 /StxEAEC	NC_018658/Chromosome	O3K_20025	C	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
		<i>ecpB</i>			O3K_20030	C	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
		<i>ecpC</i>			O3K_20035	C	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
		<i>ecpD</i>			O3K_20040	C	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
		<i>ecpE</i>			O3K_20045	C	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	
		<i>ecpR</i>			O3K_20020	C	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
	<i>E.coli</i> laminin-binding fimbriae (ELF)	<i>elfA</i>	<i>E.coli</i> O26:H11 str. 11368/EHEC	NC_013361/Chromosome	ECO26_1065	C	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
		<i>elfC</i>			ECO26_1067	C	0	1	0	1	1	1	1	1	0	0	1	0	1	1	1	1	1	1	1	1	0	0	1	
		<i>elfD</i>			ECO26_1066	C	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
		<i>elfG</i>			ECO26_1068	C	0	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	
	Adhesin FdeC	<i>eaeH</i>	<i>E. coli</i> O157:H7 str. EDL933/EHEC	NC_002655/Chromosome	Z0375	A	1	1	1	1	0	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	0	1	1	
	EtpA	<i>etpA</i>	<i>E. coli</i> O78:H11:K80 str. H10407/ETEC	NC_017724/Plasmid p948	ETEC_p948_0110	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
	F1C fimbriae	<i>focA</i>	<i>E. coli</i> str. CFT073/UPEC	NC_004431/Chromosome	c1239	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
		<i>focC</i>			c1241	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
		<i>focD</i>			c1242	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
		<i>focF</i>			c1243	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
		<i>focG</i>			c1244	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
		<i>focH</i>			c1245	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
		<i>focI</i>			c1240	NA	0	0	0	0	0	0	0	0	0															

[illegible]

		NA			EC042_4542	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
		NA			EC042_4543	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		NA			EC042_4544	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		NA			EC042_4545	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		NA			EC042_4546	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		NA			EC042_4547	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		NA			EC042_4548	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		NA			EC042_4549	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		NA			EC042_4550	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Toxin	Alpha-hemolysin	hlyA	<i>E. coli</i> O157:H7 str. EDL933/EHEC	NC_007414/Plasmid pO157	L7048	A	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	
		hlyB			L7049	A	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	
		hlyC			L7047	A	1	1	1	1	1	1	1	0	0	1	1	1	1	1	0	1	1	1	1	1	1	1
		hlyD			L7050	A	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1
		Colicin-like Usp			hcp	<i>E. coli</i> str. UM146/AIEC	NC_017632/Chromosome	UM146_23345	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Cytolethal distending toxin	cdtA	<i>E. coli</i> O18:K1:H7 str. IHE3034/NMEC	NC_017628/Chromosome	ECOK1_2089	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
		cdtB			ECOK1_2088	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
		cdtC			ECOK1_2087	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	Cytotoxic necrotizing factor 1	cnfI	<i>E. coli</i> str. UM146/AIEC	NC_017632/Chromosome	UM146_21825	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
	Enterotoxin 1	pic	<i>E. coli</i> O44:H18 str. 042/EAEC	NC_017626/Chromosome	EC042_4593	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
	Enterotoxin SenB/TieB	senB	<i>E. coli</i> str. UM146/AIEC	NC_017630/Chromosome	UM146_24581	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
	Heat-labile enterotoxin	eltA	<i>E. coli</i> O78:H11:K80 str. H10407/ETEC	NC_017722/Plasmid p666	ETEC_p666_0660	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
		eltB			ETEC_p666_0650	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	Hemolysin/cytolysin A	hlyE/clyA	<i>E. coli</i> O157:H7 str. EDL933/EHEC	NC_002655/Chromosome	Z1944	A	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
	Shiga-like toxin	stx1A	<i>E. coli</i> O157:H7 str. EDL933/EHEC	NC_002655/Chromosome	Z3344	A	1	0	0	0	0	1	1	0	0	1	1	1	1	0	0	0	0	0	1	1	0	1
		stx1B			Z3343	A	1	0	0	0	0	1	1	0	0	1	1	1	0	0	0	0	0	1	1	0	1	
		stx2A	<i>E. coli</i> O157:H7 str. EDL933/EHEC	NC_002655/Chromosome	Z1464	A	1	1	1	1	1	1	2	1	1	1	0	0	0	1	1	1	1	0	0	2	0	
stx2B		Z1465			A	1	2	1	1	1	1	2	1	1	1	0	0	0	1	1	1	1	1	0	0	2	0	

¹The number ≥1 refers to the wild-type (WT) gene and the number 0 indicates that either the gene is absent or carries a loss-of-function mutation. ²NA, gene name is not available in the published genome annotations; ³NA, not applicable; A, accessory gene; C, core gene; ⁴Genome of strain 95-3192 in GenBank does not contain a pEHEC plasmid;