

Table 5. Determinants present in UPEC but absent in *E. coli* K-12 that may contribute to virulence/fitness

UPEC-specific DNA regions, encoded functions	No. of ORFs	Size, bp	Position (<i>E. coli</i> 536 chromosome)		ORF designation (<i>E. coli</i> 536)		ORF designation (<i>E. coli</i> CFT073)	
			Start	End	Start	End	Start	End
PAI VI (<i>asnW</i> -associated polyketide biosynthesis determinant), putative iron transport system, cobalamine biosynthesis	76	92719	2018257	2110976	ECP01965	ECP02038	c2449	c2518
PAI IV (yersiniabactin biosynthesis determinant), putative invasin, pseudogenes	42	56446	1953009	2009455	ECP01913	ECP01955	c2392	c2440
<i>metV</i> island, IHAP-like gene cluster	29	40918	2945739	2986657	ECP02804	ECP02832	c3385	c3409
Fimbrial proteins	24	26111	306644	332755	ECP00287	ECP00310	c1233	c1260
Hemolysin-coregulated protein (Hpc), hypothetical proteins, sugar utilization	23	23182	4199891	4223073	ECP04024	ECP04046	c4753	c4779
Several oxidoreductases, putative permease, putative TetR/AcrR-family transcriptional regulator	11	11033	4655637	4666670	ECP04448	ECP04459	c5295	c5304
Fimbrial proteins	11	8800	4745324	4754124	ECP04531	ECP04541	<i>papX</i>	<i>papA</i>
Fimbrial proteins	10	11757	3670723	3682480	ECP03513	ECP03522	c4204	c4214
Protein involved in the transport of hemolysin A	10	10077	4759984	4770061	ECP04553	ECP04561	<i>hlyD</i>	c3565
O-antigen synthesis/modification	9	10696	2143893	2154589	ECP02076	ECP02084	<i>manB</i>	c2566
2-oxoglutarate utilization system	9	10893	4483619	4494512	ECP04275	ECP04282	c5032	c5041

UPEC-specific DNA regions, encoded functions	No. of ORFs	Size, bp	Position (<i>E. coli</i> 536 chromosome)		ORF designation (<i>E. coli</i> 536)		ORF designation (<i>E. coli</i> CFT073)	
			Start	End	Start	End	Start	End
Putative membrane proteins, putative transcriptional regulator, putative pyridoxine phosphate biosynthetic protein, putative alcohol dehydrogenase	8	8973	732243	741216	ECP00692	ECP00699	c0757	c0765
Putative galacticol/D-tagatose utilization	8	7876	3513256	3521132	ECP03346	ECP03353	c4013	c4021
PTS-dependent fructose utilization	8	7140	3934771	3941911	ECP03753	ECP03760	c4481	c4488
PAI III autotransporter serine protease	7	7445	362627	370072	ECP00343	ECP00349	c0391	c0398
Dehydrogenases, putative allantoin degradation, C4-dicarboxylate transport system	7	9354	3267686	3277040	ECP03103	ECP03109	c3750	<i>ygiK</i>
Sugar utilization	7	7724	4269990	4277714	ECP04087	ECP04093	<i>yihV</i>	c4831
Colicin cluster	6	3462	123276	126738	ECP00113	ECP00118	c0133	c0137
DNA integration/recombination/inversion protein (frame shift), part of the phage tail assembly protein	5	2010	620744	622754	ECP00591	ECP00596	c0646	c0651
Rtn protein (fragment), autotransporter fragmented	5	1443	1242901	1244344	ECP01209	ECP01212	c1610	c1615
EmrE protein	5	7570	1913321	1920891	ECP01866	ECP01874		<i>emrE</i>
Phosphoglycerate transport system transcriptional regulatory protein PgtA, regulatory proteins	5	5278	3147160	3152438	ECP02979	ECP02984	<i>pgtA</i>	c5199

UPEC-specific DNA regions, encoded functions	No. of ORFs	Size, bp	Position (<i>E. coli</i> 536 chromosome)		ORF designation (<i>E. coli</i> 536)		ORF designation (<i>E. coli</i> CFT073)	
			Start	End	Start	End	Start	End
DNA uptake/degradation	5	7184	4331768	4338952	ECP04150	ECP04154	c4894	c4898
Dipeptide uptake system	5	5067	4524502	4529569	ECP04314	ECP04318	<i>yddO</i>	c5081
Arginine degradation	5	5250	4709504	4714754	ECP04496	ECP04500	c5346	c5350
Putative autotransporter (fragment)	4	1227	1246543	1245316	ECP01215	ECP01218	c1618	c1620
Putative membrane and cytosolic proteins	4	5189	2206258	2211447	ECP02124	ECP02127	c2607	c2610
Putative sugar transport	4	2986	3152545	3155531	ECP02985	ECP02990	c0333	c0330
Polysialic acid capsule expression proteins	4	3928	3182730	3186658	ECP03024	ECP03027	c3686	c3689
Putative lipase	4	2311	4280810	4283121	ECP04098	ECP04101	c4836	c4839
C4-dicarboxylate transport system, putative Na (+)/H (+) antiporter	4	5113	4887891	4893004	ECP04684	ECP04687	c5431	c5435
Hypothetical beta-lactamase-related protein, hypothetical flavodoxin	4	3582	2860506	2864088	ECP02721	ECP02724	c3304	c3302
Putative transcriptional regulator, LysR family (frame shift), putative regulator of murein hydrolase (LrgB, LrgA family)	3	1627	4507169	4508796	ECP04296	ECP04298	c5058	c5061
Putative secreted sulfatase, putative cytoplasmic protein	3	3659	16234	19893	ECP00017	ECP00019	c0021	c0023
TonB-dependent receptor, putative cobalamin synthesis protein	3	4234	334467	338701	ECP00314	ECP00316	c1265	c1267
Hypothetical transcriptional regulator, putative aldo/keto reductase family	3	2791	384524	387315	ECP00364	ECP00366	c0412	c0414

UPEC-specific DNA regions, encoded functions	No. of ORFs	Size, bp	Position (<i>E. coli</i> 536 chromosome)		ORF designation (<i>E. coli</i> 536)		ORF designation (<i>E. coli</i> CFT073)	
			Start	End	Start	End	Start	End
Proteins								
Nitrogen assimilation Regulatory protein Nac, transcriptional regulator Cbl, hypothetical protein YeeO	3	2820	2015441	2018261	ECP01962	ECP01964	<i>nac</i>	<i>yeeO</i>
Arylsulfate sulfotransferase, putative disulfide isomerase,	3	1916	3301277	3303193	ECP03134	ECP03136	<i>asst</i>	c3787
Hypothetical proteins	3	2611	4436030	4438641	ECP04229	ECP04231	c4977	c4978
Hypothetical protein	2	1274	72461	73735	ECP00066	ECP00067	c0079	c0080
Hypothetical protein, putative peptide chain release factor	2	1759	284663	286422	ECP00264	ECP00265	c0381	<i>prfH</i>
Putative membrane protein	2	1797	582926	584723	ECP00553	ECP00554	c0608	<i>ybbJ</i>
Hypothetical protein	2	976	2328950	2329926	ECP02234	ECP02235	c2728	c2729
LPS biosynthesis proteins	2	2279	3903467	3905746	ECP03722	ECP03723	<i>waaV</i>	<i>waaW</i>
Starvation sensing protein, transporter	2	2621	4365100	4367721	ECP04174	ECP04175	c4920	<i>yjiZ</i>
Hypothetical protein, putative MP	2	1086	4478381	4479467	ECP04272	ECP04273	c5029	c5030
Hypothetical proteins, DNA binding helix- turn helix protein	2	582	4925545	4926127	ECP04775	ECP04776	c5476	c5477
Putative autotransporter/adhesin	1	2294	404507	402213	ECP00379	-	c0426	-
Putative transcriptional regulator LysR type	1	185	695097	694912	ECP00659	-	c0719	-
Putative membrane protein	1	1874	714130	716004	ECP00679	-	c0741	-

UPEC-specific DNA regions, encoded functions	No. of ORFs	Size, bp	Position (<i>E. coli</i> 536 chromosome)		ORF designation (<i>E. coli</i> 536)		ORF designation (<i>E. coli</i> CFT073)	
			Start	End	Start	End	Start	End
Putative virulence protein GipA (Peyer's patch-specific virulence factor GipA)	1	1127	822759	823886	ECP00787	-	<i>gipA</i>	-
Hemolysin E (fragment)	1	278	1250668	1250946	ECP01225	-	c1630	-
Hypothetical protein	1	275	1281829	1282104	ECP01253	-	c1660	-
Putative autotransporter/adhesin	1	7373	1444970	1452343	ECP01410	-		
Hypothetical protein	1	1397	1534478	1533081	ECP01489	-	c1919	-
6-Phospho-beta-glucosidase	1	896	1568394	1567498	ECP01519	-	c1955	-
Hypothetical protein	1	887	1571107	1570220	ECP01521	-	c1960	-
Ankyrin-like regulatory protein	1	287	1721316	1721603	ECP01671	-	c2118	-
Putative TPR repeat protein	1	1943	1764368	1762425	ECP01711	-	c2163	-
Chaperone protein HscC	1	548	1766084	1765536	ECP01713	-	<i>ybeW</i>	-
Hypothetical protein	1	578	1802382	1801804	ECP01749	-	c2206	-
Outer membrane porin	1	1079	1913838	1914917	ECP01867	-	c2348	-
Outer membrane porin	1	1055	1938975	1940030	ECP01898	-	c2383	-
Putative outer membrane protein (RatA homolog)	1	212	2645565	2645353	ECP02515	-	c3029	-
Putative outer membrane protein (SinI homolog)	1	1469	2647100	2645631	ECP02516	-	c3030	-
Putative intimin/ invasins (SinH homolog)	1	1175	2648396	2647221	ECP02517	-	c3031	-
Putative molybdenum-pterin-binding protein	1	197	2837717	2837914	ECP02700	-	c3292	-

UPEC-specific DNA regions, encoded functions	No. of ORFs	Size, bp	Position (<i>E. coli</i> 536 chromosome)		ORF designation (<i>E. coli</i> 536)		ORF designation (<i>E. coli</i> CFT073)	
			Start	End	Start	End	Start	End
Putative protease	1	905	3095561	3096466	ECP02928	-	c3517	-
YggF (Fructose-1,6-bisphosphatase), partial	1	977	3092035	3091850	ECP02924	-	c3513	-
Hypothetical protein	1	422	3757738	3758160	ECP03592	-	c4303	-
Hypothetical protein	1	461	3856488	3856949	ECP03680	-	c4398	-
Hypothetical protein	1	1076	4476725	4477801	ECP04270	-	c5027	-
Hypothetical protein	1	338	4860237	4860575	ECP04662	-	c5406	-
Putative membrane protein	1	959	4038594	4039553	ECP03876	-	c4594	-