

TABLE S1 Association among the virulence genes which were reported on various UPEC pathogenicity islands (PAIs)

Virulence factors/description	Virulence genes	Association ^a between UPEC virulence genes (n ^b)																									
		Adhesins							Capsule	Iron uptake systems						Toxins	Haemolysins and hemagglutinins			Miscellaneous							
		<i>papA</i>	<i>papC</i>	<i>papG</i>	<i>pixA</i>	<i>sfaA</i>	<i>sfaD</i>	<i>sfaHII</i>	<i>F17(a-d)-A</i>	<i>kpsM(II)</i>	<i>iroN</i>	<i>fyuA</i>	<i>irp(1)</i>	<i>irp(2)</i>	<i>iucD</i>	<i>iutA</i>	<i>chuA</i>	<i>cnf(1)</i>	<i>sat</i>	<i>hlyA</i>	<i>vat</i>	<i>hraI</i>	<i>malX</i>	<i>usp</i>			
(233)	(139)	(119)	(7)	(12)	(42)	(24)	(4)	(150)	(66)	(125)	(127)	(174)	(55)	(75)	(130)	(46)	(67)	(85)	(90)	(84)	(142)	(100)					
Adhesins																											
P-fimbriae	<i>papA</i>	– ^c																									
	<i>papC</i>	0.40	–																								
	<i>papG</i>	0.35	0.775^d	–																							
Pilus P	<i>pixA</i>		0.17	0.15	–																						
S-fimbriae	<i>sfaA</i>	0.12	0.19	0.26	N ^e	–																					
	<i>sfaD</i>	0.14	0.22	0.28	0.26	0.26	–																				
	<i>sfaHII</i>		0.21	0.25	0.28	0.57	0.56	–																			
F17-fimbriae	<i>F17(a-d)-A</i>				N	0.42	0.18		–																		
Capsule																											
Group II	<i>kpsM(II)</i>	0.34	0.633	0.59	0.16	0.21	0.36	0.30	–																		
Iron uptake systems																											
<i>E. coli</i> siderophore	<i>iroN</i>	0.14	0.27	0.28	0.29	0.35	0.67	0.47	0.41	–																	
Yersiniabactin	<i>fyuA</i>	0.30	0.30	0.32					0.28	0.27	0.20	–															
	<i>irp(1)</i>	0.26	0.34	0.40					0.29	0.32	0.25	0.84	–														
	<i>irp(2)</i>	0.21	0.45	0.46			0.15	0.34	0.21	0.47	0.30	0.70	0.72	–													
Aerobactin	<i>iucD</i>		0.24	0.30	N	N	0.12	N	0.29	0.16	0.35	0.33	0.34	–													
	<i>iutA</i>	0.11	0.37	0.34	N	N			N	0.36	0.14	0.21	0.25	0.30	0.73	–											
heme receptor	<i>chuA</i>	0.27	0.574	0.59			0.14	0.43	0.25	0.69	0.35	0.44	0.55	0.58	0.37	0.37	–										
Toxins																											
Cytotoxic necrotizing factor type 1	<i>cnf(1)</i>	0.15	0.31	0.39	0.24	0.29	0.71	0.46	0.43	0.52	0.26	0.32	0.38	0.15	0.07	0.44	–										
Secreted autotransporter toxin	<i>sat</i>	0.20	0.37	0.42	N	N	0.14	N	N	0.33			0.27	0.32	0.34	0.46	0.57	0.41	0.16								
Haemolysins and hemagglutinins																											
α-hemolysin	<i>hlyA</i>		0.35	0.42	0.25	0.33	0.50	0.47	0.19	0.42	0.53			0.17	0.27			0.30	0.66	–							
Haemoglobin protease	<i>vat</i>	0.28	0.51	0.56	0.19	0.17	0.54	0.25			0.61	0.49	0.40	0.49	0.52			0.12	0.67	0.56	0.21	0.40	–				
Heat resistant agglutinin 1	<i>hraI</i>	0.15	0.41	0.4397	0.25	0.33	0.57	0.48			0.42	0.59	0.17	0.25	0.31			0.33	0.63			0.72	0.45	–			
Miscellaneous																											
Maltose and glucose-specific IIABC component	<i>malX</i>	0.20	0.54	0.46	0.17	0.22	0.42	0.32			0.66	0.46	0.17	0.23	0.36			0.494	0.46			0.46	0.67	0.50	–		
Uropathogenic specific protein	<i>usp</i>	0.25	0.527	0.53	0.18	0.15	0.48	0.27			0.61	0.42	0.36	0.41	0.50			0.12	0.62	0.49	0.20	0.41	0.82	0.49	0.72	–	

^a Pair-wise (a total of 276 pairs of genes) correlation between VGs was determined using the Phi-coefficient (ϕ) of association. Significance level: $P > 0.05$. Only those associations which were positive and significant are presented here.

^b Value in the parentheses represents the number of isolates possessing the indicated gene in a total of 323 screen-positive isolates. The *cswA* (CS12-fimbriae) gene was excluded in the analysis since not a single isolate contained the gene.

^c –: Coefficient of association was not computed among the same pairs of genes.

^d The gene pairs (a total of 41 pairs) with strong association ($\phi > 0.50$) were bold faced.

^e N: Negative association and statistically not significant (12 pairs of genes).

TABLE S2 Change in the prevalence of UPEC virulence gene in non-disinfected and UV or PAA disinfected effluent samples

Isolates or UPEC virulence factors/genes (VFs/VGs)	No. of isolates possessing the UPEC virulence gene									
	UV					PAA				
	Non-Disinfected	Disinfected	% change (+/-) ^a	Ln-odds ratio	Significance of difference with respect to average ^b	Non-Disinfected	Disinfected	% change (+/-) ^a	Ln-odds ratio	Significance of difference with respect to average ^b
Isolates										
Total isolates	370	370	NA	NA	NA ^c	423	434	NA	NA	NA
Screen +ve	83	66	-20.5	-0.287	-	99	75	-26.2	-0.380	+
UPEC	36	17	-52.8	-0.806	+	39	22	-45.0	-0.643	+
Non-UPEC	47	49	+4.26	0.048	-	60	53	-13.9	-0.172	-
UPEC VFs/VGs										
Adherence										
<i>papA</i> ^e	60	44	-7.78	-0.266	-	73	56	+1.26	0.049	-
<i>papC</i>	48	26	-31.9	-0.747	-	43	22	-32.5	-0.615	-
<i>papG</i>	37	19	-35.4	-0.688	-	40	23	-24.1	-0.427	-
<i>pixA</i>	4	1	-68.6	-1.19	-	1	1	+32.0	0.281	-
Average	37	23	-24.0	-0.454	NA	39	26	-14.2	-0.243	NA
Capsule										
<i>kpsM(II)</i>	52	27	-34.7	-0.885	-	45	26	-23.7	-0.451	-
<i>kpsM(III)</i>	1	1	+25.8	0.232	-	2	2	+32.0	0.284	-
Average	26.5	14	-33.6	-0.555	NA	23.5	14	-21.4	-0.305	NA
Iron acquisition systems										
With PAI IV536										
<i>fyuA</i>	17	19	+40.6	0.451	+	45	44	+29.1	0.533	+
<i>irp(1)</i>	19	20	+32.4	0.382	+	45	43	+26.1	0.478	+
<i>irp(2)</i>	46	25	-31.7	-0.712	-	54	49	+20.0	0.451	+
Average (without <i>irp(2)</i>)	18.0	19.5	+36.2	0.415	NA	45.0	43.5	+27.6	0.505	NA
Without PAI IV536										
<i>iroN</i>	25	14	-29.6	-0.471	-	17	10	-22.4	-0.298	-
<i>iucD</i>	16	9	-29.3	-0.414	-	19	11	-23.6	-0.323	-
<i>iutA</i>	25	15	-24.5	-0.382	-	25	10	-47.2	-0.787	-
<i>sitA</i>	60	38	-20.4	-0.653	-	55	30	-28.0	-0.629	-
<i>sitD</i>	35	25	-10.2	-0.179	-	51	27	-30.1	-0.636	-
Average	32.2	20.2	-21.1	-0.363	NA	33.4	17.6	-30.4	-0.507	NA
Toxins										
<i>hlyA</i>	33	11	-58.1	-1.19	-	23	18	+3.3	0.043	-
<i>vat</i>	27	12	-44.1	-0.775	-	33	18	-28.0	-0.460	-
<i>cnf(1)</i>	13	6	-42.0	-0.619	-	13	14	+42.2	0.418	-
<i>cnf(2)</i>	0	0	NC	NC	NC ^d	2	0	-100	NC	NC
<i>sat</i>	14	12	+7.8	0.091	-	20	21	+38.6	0.429	-
Average	17.4	8.2	-40.7	-0.626	NA	18.2	14.2	+3.0	0.036	NA
Miscellaneous										
<i>b(1432)</i>	14	6	-46.1	-0.708	-	13	6	-39.1	-0.553	-
<i>deoK</i>	30	15	-37.1	-0.655	-	32	22	-9.3	-0.140	-
<i>gimB(orf1)</i>	19	7	-53.7	-0.917	-	12	12	+32.0	0.323	-
<i>malX</i>	53	23	-45.4	-1.19	+	41	25	-19.5	-0.346	-
<i>pic</i>	14	6	-46.1	-0.708	-	10	5	-34.0	-0.453	-
<i>usp</i>	27	15	-30.1	-0.494	-	38	20	-30.5	-0.538	-
<i>chuA</i>	40	25	-21.4	-0.422	-	42	23	-27.7	-0.510	-
Average	28.1	13.9	-38.1	-0.658	NA	26.9	16.1	-20.7	-0.306	NA

^a % change: +, Increase; -, decrease^b Statistical significance for the change of frequency of isolates/gene: $P < 0.05$, +; $P > 0.05$, -^c NA: Not applicable^d NC: Not computed due to the low frequency^e bolded virulence genes (VGs) are present on UPEC pathogenicity islands