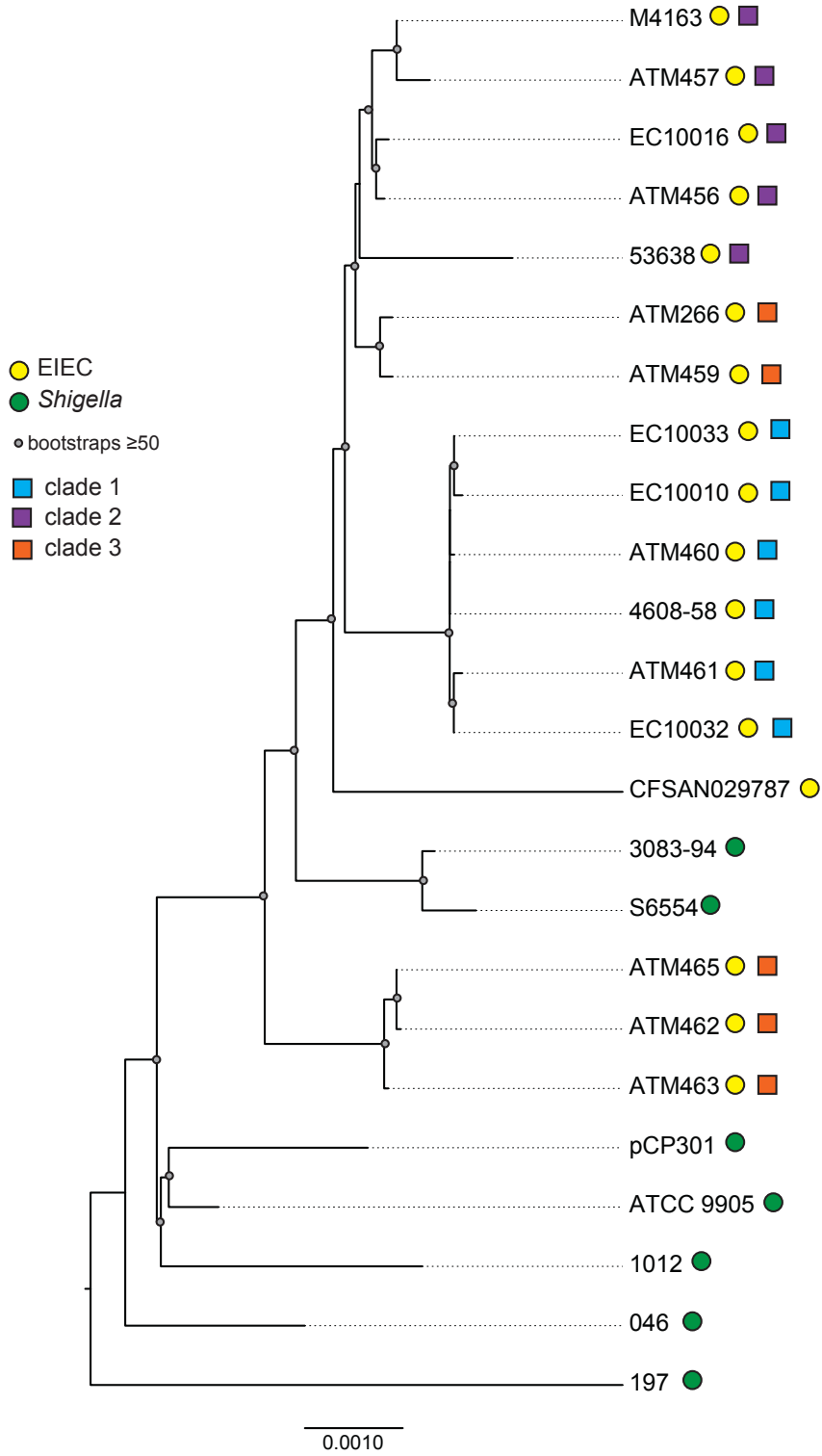


1 Supplemental Figure



2
3

4 **Figure S1. Virulence plasmid phylogeny.** The *Shigella* virulence plasmid, pCP301
5 (NC_004851.1), was aligned to the EIEC and *Shigella* genomes that contained genes
6 that exhibited similarity to those on the plasmid. The alignment was performed using
7 Mugsy (1) and the regions from each genome that aligned to pCP301 were
8 concatenated to generate a 24 kb sequence for each genome. These regions were
9 compared by constructing a maximum-likelihood phylogeny using RAxML v7.2.8 (2) that
10 was visualized using FigTree v1.4.2 (<http://tree.bio.ed.ac.uk/software/figtree/>). The
11 EIEC genomes are indicated by a yellow circle and *Shigella* genomes are indicated by a
12 green circle. The EIEC lineages are indicated by blue (lineage 1), purple (lineage 2),
13 and orange (lineage 3) squares. The scale bar indicates the distance of 0.001
14 nucleotide substitutions per site. Bootstrap values ≥ 50 are designated by a grey circle.

15

16

17

18

- 19 1. **Angiuoli SV, Salzberg SL.** 2011. Mugsy: fast multiple alignment of closely
20 related whole genomes. *Bioinformatics* **27**:334-342.
- 21 2. **Stamatakis A.** 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic
22 analyses with thousands of taxa and mixed models. *Bioinformatics* **22**:2688-2690.

23

Table S1. Genomes included in the phylogenomic analysis

Isolate ID	Serotype	Molecular Serotype	Pathovar	Phylogroup	Accession No.
HS	O9	O9:H4	commensal	A	NC_009800.1
H10407	O78:H11:K80	O78:H11	ETEC	A	FN649414.1
BL21	NA	O7:H-	lab adapted	A	NC_012947.1
K-12, BW2952	NA	O16:H48	lab adapted	A	NC_012759.1
ATCC 8739	NA	O146:H20	lab adapted	A	NC_010468.1
<i>S. dysenteriae</i> 1012	NA	SD2:H14	Shigella	B1	AAMJ000000000.1
<i>S. dysenteriae</i> S6554	NA	SB1?:H45	Shigella	B1	AMJS000000000.1
<i>S. flexneri</i> 2a 2457T	NA	SF2a:H14	Shigella	B1	NC_004741.1
<i>S. flexneri</i> 2a 301	NA	SF2a:H14	Shigella	B1	NC_004337.2
<i>S. flexneri</i> CCH060	NA	SF6:H45	Shigella	B1	AKMW000000000.1
<i>S. boydii</i> 3083-94	NA	SB18:H45	Shigella	B1	NC_010658.1
<i>S. boydii</i> ATCC 9905	NA	SB7:H14	Shigella	B1	AERN000000000.1
<i>S. sonnei</i> 046	NA	SS:H16	Shigella	B1	NC_007384.1
SE11	NA	O173:H28	commensal	B1	NC_011415.1
55989	O104:H4	O104:H4	EAEC	B1	NC_011748.1
11368	O26:H11	O26:H11	EHEC/STEC	B1	NC_013361.1
11128	O111:H-	O111:H8	EHEC/STEC	B1	NC_013364.1
E110019	O111:H9	O111:H9	AEEC	B1	AAJW000000000.2
B171	O111:NM	O111:H2	EPEC2	B1	AAJX000000000.2
E24377A	O139:H28	O139v:H28	ETEC	B1	NC_009801.1
B7A	O148:H28	O148:H28	ETEC	B1	AAJT000000000.2
IAI1	O8	O8/SB17:H19	fecal isolate	B1	NC_011741.1
TY-2482	O104:H4	O104:H4	O104 outbreak	B1	AFOG000000000.1
ED1a	O81	O81:H27	commensal	B2	CU928162.2
E2348/69	O127:H6	O127:H6	EPEC1	B2	NC_011601.1
536	O6:K15:H31	O6:H31	UPEC/ExPEC	B2	NC_008253.1
S88	O45:K1	O45v:H7	UPEC/ExPEC	B2	NC_011742.1
UTI89	NA	O18:H7	UPEC/ExPEC	B2	NC_007946.1
CFT073	O6:K2:H1	O6:H1	UPEC/ExPEC	B2	AE014075.1
042	O44:H18	O44:H18	EAEC	D	FN554766.1
UMN026	O17:K52:H18	O17:H18	UPEC/ExPEC	D	NC_011751.1
<i>S. dysenteriae</i> 197	NA	SD1:H18	Shigella	E	NC_007606.1
EDL933	O157:H7	O157:H7	EHEC/STEC	E	NC_002655.2
Sakai	O157:H7	O157:H7	EHEC/STEC	E	NC_002695.1
CB9615	O55:H7	O55:H7	AEEC	E	NC_013941.1
IA139	O7:K1	O7:H45	UPEC/ExPEC	F	NC_011750.1
SMS_3_5	O19:H34	O153:H34	other	F	CP000970.1

NA indicates that data was not applicable or could not be determined.

Table S2. Predicted protein functions of EIEC genes

LS-BSR Gene Cluster	Gene	Predicted Protein Function
<u>EIEC vs. other <i>E. coli</i> and <i>Shigella</i></u>		
Cluster_178041	<i>pyk</i>	pyruvate kinase
Cluster_178088		putative membrane protein
Cluster_178120		carboxymuconolactone decarboxylase family protein
Cluster_179012		hypothetical protein
Cluster_181725		hypothetical protein
Cluster_183530		dihydrouridine synthase family protein
Cluster_186711	<i>torT</i>	TMAO reductase system periplasmic protein TorT
<u>EIEC vs. other <i>E. coli</i> (not including <i>Shigella</i>)</u>		
Cluster_16689		4Fe-4S dicluster domain protein
Cluster_171880		hypothetical protein
Cluster_171882		conjugal transfer/type IV secretion DotA/TraY family protein
Cluster_171887		hypothetical protein
Cluster_167621		phage-related minor tail family protein
Cluster_168446		enterobacterial exodeoxyribonuclease VIII family protein
Cluster_168480		phage major capsid protein, HK97 family
Cluster_168972		hypothetical protein
Cluster_168973		hypothetical protein
Cluster_169741		bacteriophage CI repressor helix-turn-helix domain protein
Cluster_169827		metallo-beta-lactamase superfamily protein
Cluster_170182		hypothetical protein
Cluster_173349		phage gp6-like head-tail connector family protein
Cluster_175432		hypothetical protein
Cluster_175571	<i>rpoZ</i>	DNA-directed RNA polymerase, omega subunit
Cluster_168990		hypothetical protein
Cluster_168992		hypothetical protein
Cluster_168993		hypothetical protein
Cluster_168995		hypothetical protein
Cluster_168997		hypothetical protein
Cluster_169044		glycosyl transferase 2 family protein
Cluster_169048		glycosyl transferase 2 family protein
Cluster_169049		CDP-Glycerol:Poly(glycerophosphate) glycerophosphotransferase family protein
Cluster_169935		oxidoreductase family, C-terminal alpha/beta domain protein
Cluster_169939		D5 N terminal like family protein
Cluster_169940		conserved hypothetical protein
Cluster_170184		integrating conjugative element , PFL_4695 family domain protein
Cluster_170187		hypothetical protein
Cluster_170188		response regulator
Cluster_170194		conserved hypothetical family protein
Cluster_170195	<i>ssb</i>	single-stranded DNA-binding family protein
Cluster_171910		type IV leader peptidase family protein
Cluster_171911		transglycosylase SLT domain protein
Cluster_178041	<i>pyk</i>	pyruvate kinase
Cluster_178069		kinase/pyrophosphorylase family protein
Cluster_178088		putative membrane protein

Cluster_178120		carboxymuconolactone decarboxylase family protein
Cluster_169052	<i>galE</i>	UDP-glucose 4-epimerase GalE
Cluster_170093		hypothetical protein
Cluster_170303		arylsulfotransferase family protein
Cluster_170304		DSBA-like thioredoxin domain protein
Cluster_170515		conserved hypothetical protein
Cluster_170570		hypothetical protein
Cluster_17062		hypothetical protein
Cluster_170665		hypothetical protein
Cluster_1707		flxA-like family protein
Cluster_170798		fimbrial family protein
Cluster_170859		haemagglutinin family protein
Cluster_171074		SIS domain protein
Cluster_1711		methyltransferase domain protein
Cluster_171226	<i>btuB</i>	tonB-dependent vitamin B12 receptor
Cluster_171408		methyltransferase domain protein
Cluster_171424		transposase IS66 family protein
Cluster_171433		hypothetical protein
Cluster_171446		oligogalacturonate-specific porin family protein
Cluster_171449		FCD domain protein
Cluster_171915		antitoxin Phd_YefM, type II toxin-antitoxin system family protein
Cluster_171916		addiction module toxin, RelE/StbE family protein
Cluster_171917		type IV pilus biogenesis family protein
Cluster_171918		pilin accessory family protein
Cluster_171919		type IVB pilus formation outer membrane protein, R64 PilN family
Cluster_171920		pilM family protein
Cluster_171923		putative membrane protein
Cluster_172064	<i>sfsA</i>	sugar fermentation stimulation protein
Cluster_17208		fimbrial family protein
Cluster_17209		fimbrial family protein
Cluster_17212		fimbrial family protein
Cluster_17213	<i>folK</i>	2-amino-4-hydroxy-6-hydroxymethyldihydropteridine diphosphokinase
Cluster_172174		phosphoribosyl transferase domain protein
Cluster_172190		hypothetical protein
Cluster_172200		hypothetical protein
Cluster_172206		Ash family protein
Cluster_175612		type III secretion apparatus protein, YscI/HrpB, C-terminal domain
Cluster_175616		type III secretion low calcium response chaperone LcrH/SycD
Cluster_176582		phage gp6-like head-tail connector family protein
Cluster_179011		terminase small subunit
Cluster_179012		hypothetical protein
Cluster_170112		description family protein
Cluster_170113		conserved hypothetical protein
Cluster_1723		hypothetical protein
Cluster_172355		plasmid stabilisation system family protein
Cluster_178247		YecM family protein
Cluster_176594		phage baseplate assembly V family protein
Cluster_179014		hypothetical protein
Cluster_173627		hypothetical protein

Cluster_178285		RNA polymerase sigma factor, sigma-70 family protein
Cluster_181210		bacterial mobilization family protein
Cluster_181725		hypothetical protein
Cluster_172738		PTS system sugar-specific permease component family protein
Cluster_181221		hypothetical protein
Cluster_183530		dihydrouridine synthase family protein
Cluster_183668		hypothetical protein
Cluster_183746		efflux pump membrane family protein
Cluster_183764		phosphotransferase system, EIIc family protein
Cluster_18658		phage integrase family protein
Cluster_186711	<i>torT</i>	TMAO reductase system periplasmic protein TorT

EIEC and *Shigella* vs. other *E. coli*

Cluster_16689		4Fe-4S dicluster domain protein
Cluster_167621		phage-related minor tail family protein
Cluster_168446		enterobacterial exodeoxyribonuclease VIII family protein
Cluster_168480		phage major capsid protein, HK97 family
Cluster_168972		hypothetical protein
Cluster_168973		hypothetical protein
Cluster_168990		hypothetical protein
Cluster_168992		hypothetical protein
Cluster_168993		hypothetical protein
Cluster_168995		hypothetical protein
Cluster_168997		hypothetical protein
Cluster_169044		glycosyl transferase 2 family protein
Cluster_169048		glycosyl transferase 2 family protein
Cluster_169049		CDP-Glycerol:Poly(glycerophosphate) glycerophosphotransferase family protein
Cluster_169052	<i>galE</i>	UDP-glucose 4-epimerase GalE
Cluster_169741		bacteriophage CI repressor helix-turn-helix domain protein
Cluster_169827		metallo-beta-lactamase superfamily protein
Cluster_169935		oxidoreductase family, C-terminal alpha/beta domain protein
Cluster_169939		D5 N terminal like family protein
Cluster_169940		conserved hypothetical protein
Cluster_170093		hypothetical protein
Cluster_170112		description family protein
Cluster_170113		conserved hypothetical protein
Cluster_170182		hypothetical protein
Cluster_170184		integrating conjugative element , PFL_4695 family domain protein
Cluster_170187		hypothetical protein
Cluster_170188		response regulator
Cluster_170194		conserved hypothetical family protein
Cluster_170195	<i>ssb</i>	single-stranded DNA-binding family protein
Cluster_170303		arylsulfotransferase family protein
Cluster_170304		DSBA-like thioredoxin domain protein
Cluster_170515		conserved hypothetical protein
Cluster_170570		hypothetical protein
Cluster_17062		hypothetical protein
Cluster_170665		hypothetical protein
Cluster_1707		flxA-like family protein

Cluster_170798		fimbrial family protein
Cluster_170859		haemagglutinin family protein
Cluster_171074		SIS domain protein
Cluster_1711		methyltransferase domain protein
Cluster_171226	<i>btuB</i>	tonB-dependent vitamin B12 receptor
Cluster_171408		methyltransferase domain protein
Cluster_171424		transposase IS66 family protein
Cluster_171433		hypothetical protein
Cluster_171446		oligogalacturonate-specific porin family protein
Cluster_171449		FCD domain protein
Cluster_171880		hypothetical protein
Cluster_171882		conjugal transfer/type IV secretion DotA/TraY family protein
Cluster_171887		hypothetical protein
Cluster_171910		type IV leader peptidase family protein
Cluster_171911		transglycosylase SLT domain protein
Cluster_171915		antitoxin Phd_YefM, type II toxin-antitoxin system family protein
Cluster_171916		addiction module toxin, RelE/StbE family protein
Cluster_171917		type IV pilus biogenesis family protein
Cluster_171918		pilin accessory family protein
Cluster_171919		type IVB pilus formation outer membrane protein, R64 PilN family
Cluster_171920		pilM family protein
Cluster_171923		putative membrane protein
Cluster_172064	<i>sfsA</i>	sugar fermentation stimulation protein
Cluster_17208		fimbrial family protein
Cluster_17209		fimbrial family protein
Cluster_17212		fimbrial family protein
Cluster_17213	<i>folK</i>	2-amino-4-hydroxy-6-hydroxymethyldihydropteridine diphosphokinase
Cluster_172174		phosphoribosyl transferase domain protein
Cluster_172190		hypothetical protein
Cluster_172200		hypothetical protein
Cluster_172206		ash family protein
Cluster_1723		hypothetical protein
Cluster_172355		plasmid stabilisation system family protein
Cluster_172738		PTS system sugar-specific permease component family protein
Cluster_173349		phage gp6-like head-tail connector family protein
Cluster_173627		hypothetical protein
Cluster_175432		hypothetical protein
Cluster_175571	<i>rpoZ</i>	DNA-directed RNA polymerase, omega subunit
Cluster_175612		type III secretion apparatus protein, YscI/HrpB, C-terminal domain
Cluster_175616		type III secretion low calcium response chaperone LcrH/SycD
Cluster_176582		phage gp6-like head-tail connector family protein
Cluster_176594		phage baseplate assembly V family protein
Cluster_178041	<i>pyk</i>	pyruvate kinase
Cluster_178069		kinase/pyrophosphorylase family protein
Cluster_178088		putative membrane protein
Cluster_178120		carboxymuconolactone decarboxylase family protein
Cluster_178247		YecM family protein
Cluster_178285		RNA polymerase sigma factor, sigma-70 family protein
Cluster_179011		terminase small subunit

Cluster_179012
Cluster_179014

hypothetical protein
hypothetical protein

Table S3. Predicted protein functions of EIEC clade-specific genes

LS-BSR Gene Cluster	Gene	Predicted Protein Function
Clade 1		
Cluster_171822		ABC transporter family protein
Cluster_171824		transposase family protein
Cluster_171826		colicin pore forming domain protein
Cluster_171827		colicin E1 (microcin) immunity family protein
Cluster_173138		UreE urease accessory , C-terminal domain protein
Cluster_174773		hypothetical protein
Cluster_174373		fimbrial family protein
Cluster_174374		PapC C-terminal domain protein
Cluster_174375		fimbrial family protein
Cluster_174408		ribosome inactivating family protein
Cluster_174512		bacterial transcriptional regulator family protein
Cluster_174580		hypothetical protein
Cluster_174582		gene 9 domain protein
Cluster_17470	<i>ribH</i>	6,7-dimethyl-8-ribityllumazine synthase
Cluster_174762		hypothetical protein
Cluster_176485		methyltransferase domain protein
Cluster_177135		hypothetical protein
Cluster_177152		UbiE/COQ5 methyltransferase family protein
Cluster_17865		glutathione S-transferase, C-terminal domain protein
Cluster_178767		tetratricopeptide repeat family protein
Cluster_178830		fimbrial family protein
Cluster_178850		coA-transferase III family protein
Cluster_178852	<i>oxc</i>	oxalyl-CoA decarboxylase
Cluster_17886		hypothetical protein
Cluster_178585		FGGY-pentulose kinase family protein
Cluster_178533		glycosyl transferases group 1 family protein
Cluster_178576		HAD hydrolase, IA, variant 1 family protein
Cluster_179214		hypothetical protein
Cluster_179207		hypothetical protein
Cluster_180728		hypothetical protein
Cluster_180729		competence CoiA-like family protein
Cluster_180732		hypothetical protein
Cluster_180737		helix-turn-helix domain protein
Cluster_180738		hypothetical protein
Cluster_180727		hypothetical protein
Cluster_180726		conserved hypothetical protein
Cluster_180680		bacterial DNA-binding family protein
Cluster_181364		RelB antitoxin family protein
Cluster_181369		conserved hypothetical protein
Cluster_181372		hypothetical protein
Cluster_181375		hypothetical protein
Cluster_181376		resolvase, N terminal domain protein
Cluster_181377		conserved hypothetical protein
Cluster_181379		hypothetical protein
Cluster_181380		outer membrane efflux family protein
Cluster_181359		DNA methylase family protein
Cluster_181363		putative membrane protein

Cluster_181355		transposase IS116/IS110/IS902 family protein
Cluster_181347		PAS fold family protein
Cluster_183234		SIS domain protein
Cluster_183235	<i>pyk</i>	pyruvate kinase
Cluster_183237		opacity-associated A LysM-like domain protein
Cluster_18297		biofilm development YmgB/AriR family protein
Cluster_182952		helix-turn-helix domain protein
Cluster_184798		putative membrane protein
Cluster_184808		periplasmic binding protein-like domain protein
Cluster_184859		InaA N-terminal domain protein
Cluster_184884		RND transporter, hydrophobe/amphiphile efflux-1 family protein
Cluster_184888		cytochrome c family protein
Cluster_184908		binding-protein-dependent transport system inner membrane component family protein
Cluster_185221		hypothetical protein
Cluster_185250		formate dehydrogenase family accessory protein FdhD
Cluster_185308		glycyl-radical enzyme activating family protein
Cluster_18561		binding-protein-dependent transport system inner membrane component family protein
Cluster_185625		HNH endonuclease family protein
Cluster_185662		conserved hypothetical family protein
Cluster_18465	<i>pgmB</i>	beta-phosphoglucomutase
Cluster_184656	<i>fmt</i>	methionyl-tRNA formyltransferase
Cluster_184619		ribonuclease, Rne/Rng family domain protein
Cluster_184624		diguanylate cyclase domain protein
Cluster_187913	<i>ubiG</i>	3-demethylubiquinone-9 3-O-methyltransferase
Cluster_187956		NADH dehydrogenase family protein
Cluster_18798		DeoC/LacD aldolase family protein
Cluster_188215		caudovirales tail fiber assembly family protein
Cluster_188216		hypothetical protein
Cluster_187732		polysaccharide pyruvyl transferase family protein
Cluster_191020		polysaccharide biosynthesis family protein
Cluster_191023		hypothetical protein
Cluster_191024		hypothetical protein
Cluster_191027		glycosyl transferase 2 family protein
Cluster_191028		mannose-1-phosphate guanylyltransferase/mannose-6-phosphate isomerase
Cluster_191061		hypothetical protein
Cluster_191062		hypothetical protein
Cluster_191085		hypothetical protein
Cluster_191382		KR domain protein
Cluster_191530		conserved hypothetical protein
Cluster_190944		WGR domain protein
Cluster_190926		outer membrane beta-barrel domain protein
Cluster_193389		DctM-like transporters family protein
Cluster_193410		PapG carbohydrate binding domain protein
Cluster_193450	<i>neuC</i>	UDP-N-acetyl-D-glucosamine 2-epimerase, UDP-hydrolysing
Cluster_193451		GDSL-like Lipase/Acylhydrolase family protein
Cluster_193754		branched-chain amino acid transport system/ permease component family protein
Cluster_193756		periplasmic binding and sugar binding domain of Lacl family protein
Cluster_193912		putative membrane protein
Cluster_194039		hypothetical protein
Cluster_194075		conserved hypothetical protein
Cluster_194123		haemagglutinin family protein

Cluster_193182		KDPG and KHG aldolase family protein
Cluster_197538		hypothetical protein
Cluster_197540		mannose-1-phosphate guanylyltransferase/mannose-6-phosphate isomerase
Cluster_197305		BRO family, N-terminal domain protein
Cluster_197316		hypothetical protein
Cluster_203828		conserved hypothetical protein
Cluster_203846		type II secretion system (T2SS), K family protein
Cluster_203867		hypothetical protein
Cluster_203873		hypothetical protein
Cluster_203875		hypothetical protein
Cluster_203905		YcfA-like family protein
Cluster_203927		HMGL-like family protein
Cluster_203944		conserved hypothetical protein
Cluster_204129		prokaryotic DksA/TraR C4-type zinc finger family protein
Cluster_204133		putative peptidoglycan domain protein
Cluster_203574		Major Facilitator Superfamily protein
Cluster_203598		hypothetical protein
Cluster_20672	<i>hisF</i>	imidazoleglycerol phosphate synthase, cyclase subunit
Cluster_206720		acyl-CoA reductase family protein
Cluster_20675	<i>hisB</i>	histidinol-phosphatase
Cluster_206757		hypothetical protein
Cluster_20677	<i>hisD</i>	histidinol dehydrogenase
Cluster_20678	<i>hisG</i>	ATP phosphoribosyltransferase
Cluster_206784		hypothetical protein
Cluster_206797		hypothetical protein
Cluster_207165		Gram-negative porin family protein
Cluster_207229		cobinamide kinase / cobinamide phosphate guanylyltransferase family protein
Cluster_207285	<i>gnd</i>	6-phosphogluconate dehydrogenase
Cluster_207289		N-terminal domain of galactosyltransferase family protein
Cluster_207292		3-beta hydroxysteroid dehydrogenase/isomerase family protein
Cluster_207293	<i>rfbC</i>	dTDP-4-dehydrorhamnose 3,5-epimerase
Cluster_207294	<i>rfbB</i>	dTDP-glucose 4,6-dehydratase
Cluster_207343		phage tail tape measure protein, TP901 family, core region
Cluster_207358		P2 phage tail completion R family protein
Cluster_207396		nucleoside transporter family protein
Cluster_20746		prokaryotic cytochrome b561 family protein
Cluster_207650		hypothetical protein
Cluster_207652		conserved hypothetical protein
Cluster_207676		NinF family protein
Cluster_207677		conserved hypothetical protein
Cluster_207681		hypothetical protein
Cluster_207694		hypothetical protein
Cluster_207699		hypothetical protein
Cluster_207748	<i>ptsP</i>	phosphoenolpyruvate-protein phosphotransferase
Cluster_207820	<i>xseA</i>	exodeoxyribonuclease VII, large subunit
Cluster_207827	<i>yfgL</i>	outer membrane assembly lipoprotein YfgL
Cluster_207845	<i>iscS</i>	cysteine desulfurase IscS
Cluster_207922	<i>ffh</i>	signal recognition particle protein
Cluster_207935		phage tail tape measure protein, TP901 family, core region
Cluster_207939		caudovirales tail fibre assembly family protein
Cluster_207945		AIPR family protein

Cluster_207946 phage virion morphogenesis protein
Cluster_207948 phage lysis regulatory, LysB family protein
Cluster_207959 hypothetical protein
Cluster_207961 hypothetical protein
Cluster_207966 conserved hypothetical protein
Cluster_207970 hypothetical protein

Clade 2

Cluster_171845 hypothetical protein
Cluster_171850 hypothetical protein
Cluster_174308 mannose-1-phosphate guanylyltransferase/mannose-6-phosphate isomerase
Cluster_176267 IstB-like ATP binding family protein
Cluster_176265 ToIA C-terminal family protein
Cluster_193388 tripartite ATP-independent periplasmic transporters, DctQ component family protein
Cluster_193436 *iutA* ferric aerobactin receptor
Cluster_193448 alpha-2,8-polysialyltransferase family protein
Cluster_197534 putative membrane protein
Cluster_197545 glycosyl transferases group 1 family protein
Cluster_203862 hypothetical protein
Cluster_20700 hypothetical protein

Clade 3

Cluster_181381 hypothetical protein
Cluster_183353 hypothetical protein
Cluster_188221 hypothetical protein
Cluster_197682 LysR substrate binding domain protein
Cluster_197733 conserved hypothetical protein
Cluster_197619 hypothetical protein
Cluster_197635 response regulator
Cluster_204204 bacterial flagellin C-terminal helical region family protein
Cluster_208196 PGAP1-like family protein
Cluster_208198 hypothetical protein
Cluster_208008 hypothetical protein
Cluster_212088 hypothetical protein
Cluster_212334 conserved hypothetical protein
