

## Escherichia coli EHEC Germany outbreak preliminary functional annotation using BG7 system

Marina Manrique, Pablo Pareja-Tobes, Eduardo Pareja-Tobes, Eduardo Pareja, Raquel Tobes

Oh no sequences! Group. Era7 Bioinformatics.

Corresponding autor: Raquel Tobes rtobes@era7.com

We have done the annotation of the genome sequenced by BGI (6-2-2011, <http://www.bgisequence.com/eu/index.php?cid=194>) and assembled with MIRA by Nick Loman (6-2-2011 <http://pathogenomics.bham.ac.uk/blog/2011/06/ehec-genome-assembly/>).

Our system BG7 (Bacterial Genome annotation of Era7 Bioinformatics, [https://registration.hinxton.wellcome.ac.uk/display\\_info.asp?id=227](https://registration.hinxton.wellcome.ac.uk/display_info.asp?id=227), [http://www.slideshare.net/marina\\_manrique/bg7-a-new-system-for-bacterial-genome-annotation-designed-for-ngs-data](http://www.slideshare.net/marina_manrique/bg7-a-new-system-for-bacterial-genome-annotation-designed-for-ngs-data)) predicts ORFs and annotates them based on fragments of similarity with Uniprot proteins.

In contrast to other annotation pipelines where finding ORFs is the first step followed by the annotation one, BG7 system first searches for protein similarity and then defines the ORF searching for start and stop signals. It is specifically designed for annotating prokaryotic genomes obtained with NGS data since it handles the principal errors of these technologies: false indels in homopolymer regions and substitutions. Annotation systems based on initial and exact ORF detection often may lose ORFs due to these kinds of sequencing errors that may lead to introduction or lack of stop codons and modification of start signals. BG7 is also designed to work with genomes fragmented in many contigs solving the problem of the detection of incomplete genes at the end of contigs. The system is especially suitable to detect rare genes similar to proteins from taxonomically distant organisms. BG7 takes advantage of cloud computing to perform extensive computing tasks in a reasonable time. The annotation of a 3Mb bacterial genome can be performed in less than 12 hours.

Our prediction of genes of the MIRA assembly of the e.coli EHEC responsible for the last European outbreak sequenced by BGI was based on a set of 137063 proteins composed by:

- All the representative Uniprot proteins corresponding to all Uniref90 clusters for all *Escherichia coli* proteins
- All Uniprot proteins from organisms including in their name the terms "EHEC" or "EAEC"
- All Uniprot proteins from bacteria including in any field the term "toxin"

- All Uniprot proteins from bacteria including in any field the term “hemolysin”
- All the proteins from *Shigella dysenteriae*, *Yersinia pestis* and *Shigella dysenteriae*

## RESULTS

We have predicted **6327** genes, **6156** encoding proteins y **171** corresponding to ribosomal and tRNA.

Only **1326** out of the 6156 protein encoding genes have canonical start and stop codon and haven't frame-shifts neither intragenic stop codons. **2479** protein encoding genes (out of the 6156 predicted) include some frameshift or some intragenic stop codon in their sequences, probably caused by inherent technology errors. However our system is tolerant to errors of massive sequencing technologies and it has been able to detect a rich set of genes even with very preliminary sequencing results.

Probably some of the proteins detected are fragmented and some of them could appear as two different predicted genes if they are in different contigs.

We have analyzed the taxonomic origin of the proteins responsible of the prediction of the detected genes. Table 1, Figure 1 and Figure 2 display the result of this analysis.

**Table 1: Taxonomic origin of proteins responsible of the prediction of the detected genes**

Organism	number of proteins
Escherichia coli O26:H11 (strain 11368 / EHEC)	2810
Escherichia coli (strain 55989 / EAEC)	1166
Escherichia coli O44:H18 (strain 042 / EAEC)	339
Escherichia coli O103:H2 (strain 12009 / EHEC)	296
Escherichia coli	221
Escherichia coli O111:H- (strain 11128 / EHEC)	151
Escherichia coli O157:H7 (strain EC4115 / EHEC)	148
Escherichia coli O157:H7 (strain TW14359 / EHEC)	144
Escherichia coli (strain K12)	51
Salmonella typhi	51
Escherichia coli O1:K1 / APEC	50
Escherichia coli (strain UTI89 / UPEC)	40
Escherichia coli O81 (strain ED1a)	30
Yersinia pestis	29
Escherichia coli O139:H28 (strain E24377A / ETEC)	18
Escherichia coli B354	14
Escherichia coli O55:H7 (strain CB9615 / EPEC)	13

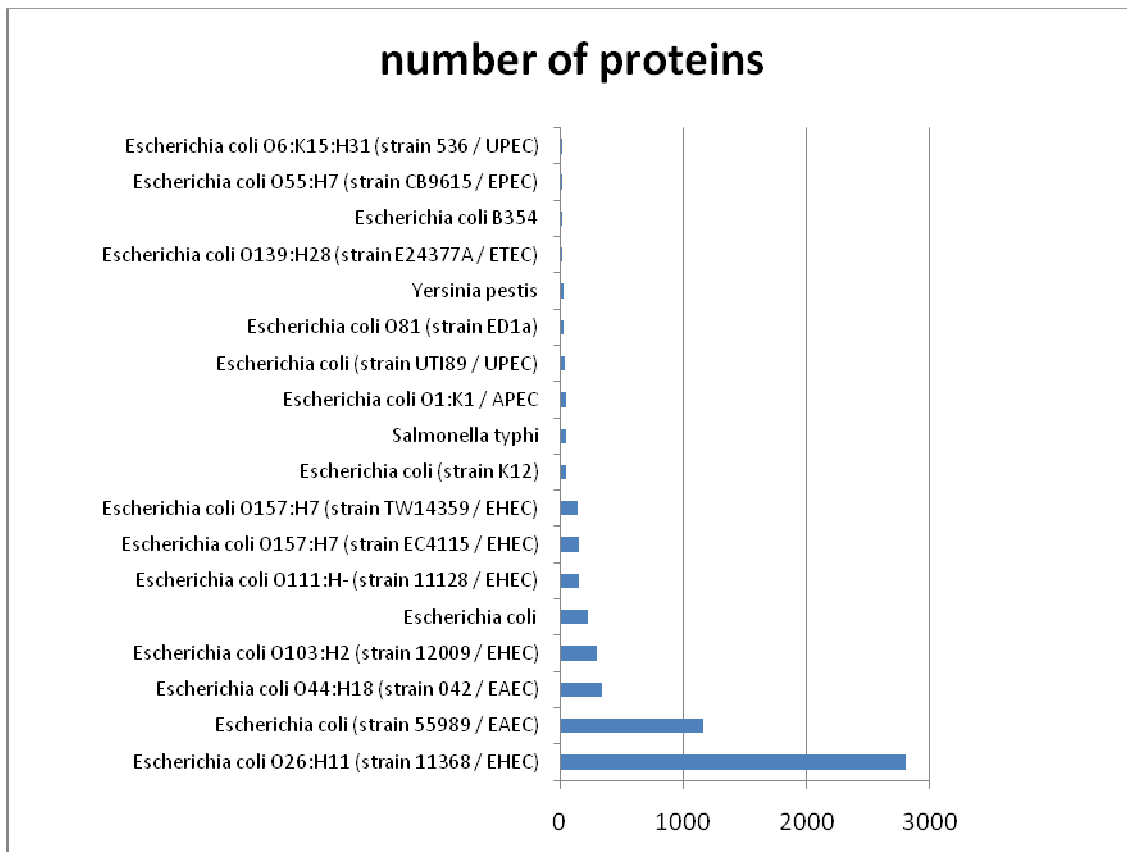
Escherichia coli O6:K15:H31 (strain 536 / UPEC)	13
Escherichia coli B088	12
Escherichia coli MS 119-7	12
Escherichia coli O6	12
Escherichia coli TA007	12
Escherichia coli (strain SE11)	11
Escherichia coli 1827-70	11
Escherichia coli MS 107-1	11
Escherichia coli O127:H6 (strain E2348/69 / EPEC)	11
Escherichia coli EPECa14	10
Escherichia coli M863	10
Escherichia coli MS 124-1	10
Escherichia coli B7A	9
Escherichia coli H120	9
Escherichia coli MS 117-3	9
Escherichia coli MS 198-1	9
Escherichia coli MS 21-1	9
Escherichia coli O157:H7	9
Shigella dysenteriae	9
Shigella flexneri 2a str. 2457T	9
Escherichia coli 3431	8
Escherichia coli LT-68	8
Escherichia coli MS 116-1	8
Escherichia coli MS 182-1	8
Escherichia coli O17:K52:H18 (strain UMN026 / ExPEC)	8
Escherichia coli O45:K1 (strain S88 / ExPEC)	8
Escherichia coli EC4100B	7
Escherichia coli MS 145-7	7
Escherichia coli MS 196-1	7
Escherichia coli MS 84-1	7
Escherichia coli (strain SMS-3-5 / SECEC)	6
Escherichia coli B185	6
Escherichia coli E128010	6
Escherichia coli E482	6
Escherichia coli FVEC1302	6
Escherichia coli MS 16-3	6
Escherichia coli MS 175-1	6
Escherichia coli MS 187-1	6
Escherichia coli MS 69-1	6
Escherichia coli O9:H4 (strain HS)	6
Escherichia sp. 3_2_53FAA	6
Shigella boydii serotype 18 (strain CDC 3083-94 / BS512)	6
Shigella dysenteriae serotype 1 (strain Sd197)	6
Shigella flexneri	6
Escherichia coli E1520	5

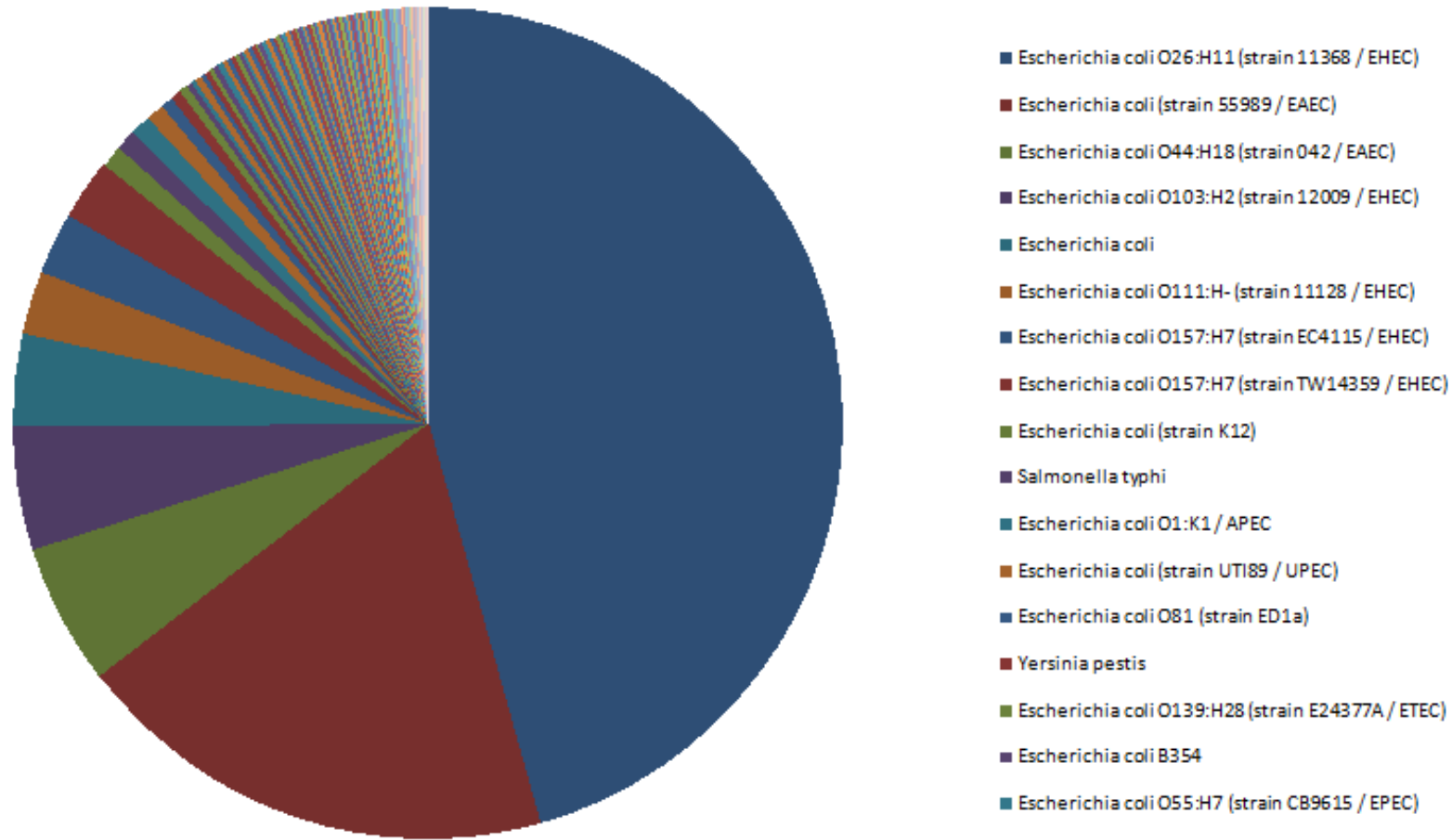
Escherichia coli MS 110-3	5
Escherichia coli MS 115-1	5
Escherichia coli MS 45-1	5
Escherichia coli O157:H7 str. TW14588	5
Escherichia coli O7:K1 (strain IAI39 / ExPEC)	5
Enterobacter sakazakii (strain ATCC BAA-894)	4
Escherichia coli 1357	4
Escherichia coli 53638	4
Escherichia coli B171	4
Escherichia coli E110019	4
Escherichia coli E22	4
Escherichia coli F11	4
Escherichia coli MS 185-1	4
Escherichia coli MS 78-1	4
Escherichia coli MS 85-1	4
Escherichia coli O111:H-	4
Escherichia coli O55:H7 str. USDA 5905	4
Escherichia coli O78:H11 (strain H10407 / ETEC)	4
Escherichia fergusonii (strain ATCC 35469 / DSM 13698 / CDC 0568-73)	4
Citrobacter koseri (strain ATCC BAA-895 / CDC 4225-83 / SGSC4696)	3
Enterobacteria phage lambda (Bacteriophage lambda)	3
Escherichia albertii TW07627	3
Escherichia coli (strain ATCC 55124 / KO11)	3
Escherichia coli (strain B / BL21)	3
Escherichia coli (strain B / REL606)	3
Escherichia coli 1180	3
Escherichia coli 83972	3
Escherichia coli H263	3
Escherichia coli MS 57-2	3
Escherichia coli MS 60-1	3
Escherichia coli RN587/1	3
Escherichia sp. 1_1_43	3
Salmonella typhimurium	3
Shigella boydii serotype 4 (strain Sb227)	3
Shigella flexneri serotype 5b (strain 8401)	3
Bacillus cereus G9241	2
Citrobacter youngae ATCC 29220	2
Enterobacteria phage CUS-3	2
Enterobacteria phage VT2phi_272	2
Escherichia coli (strain K12 / DH10B)	2
Escherichia coli (strain K12 / MC4100 / BW2952)	2
Escherichia coli 2362-75	2
Escherichia coli FVEC1412	2
Escherichia coli H489	2
Escherichia coli MS 146-1	2

Escherichia coli MS 200-1	2
Escherichia coli NC101	2
Escherichia coli O150:H5 (strain SE15)	2
Escherichia coli O157:H- str. H 2687	2
Escherichia coli O157:H7 str. 1125	2
Escherichia coli O157:H7 str. EC869	2
Escherichia coli O55:H7 str. 3256-97	2
Escherichia coli O8 (strain IAI1)	2
Escherichia coli TW10509	2
Salmonella choleraesuis	2
Serratia marcescens	2
Shigella flexneri CDC 796-83	2
Shigella sonnei (strain Ss046)	2
Citrobacter rodentium (strain ICC168) (Citrobacter freundii biotype 4280)	1
Citrobacter sp. 30_2	1
Cronobacter turicensis (strain DSM 18703 / LMG 23827 / z3032)	1
Enterobacteria phage H19B (Bacteriophage H19B)	1
Enterobacteria phage Sf6 (Shigella flexneri bacteriophage VI) (Bacteriophage SfVI)	1
Enterobacteria phage VT2-Sa (Bacteriophage VT2-Sa)	1
Erwinia amylovora (strain CFBP1430)	1
Escherichia coli (strain UM146)	1
Escherichia coli 101-1	1
Escherichia coli H252	1
Escherichia coli MS 153-1	1
Escherichia coli O157:H- str. 493-89	1
Escherichia coli O157:H7 str. 1044	1
Escherichia coli O157:H7 str. EC1212	1
Escherichia coli O157:H7 str. EC4042	1
Escherichia coli O157:H7 str. EC4486	1
Escherichia coli O157:H7 str. G5101	1
Escherichia coli O83:H1 (strain NRG 857C / AIEC)	1
Escherichia coli OR:K5:H- (strain ABU 83972)	1
Haemophilus ducreyi	1
Klebsiella pneumoniae (strain 342)	1
Klebsiella pneumoniae subsp. pneumoniae (strain ATCC 700721 / MGH 78578)	1
Klebsiella pneumoniae subsp. pneumoniae NTUH-K2044	1
Klebsiella variicola (strain At-22)	1
Pantoea sp. (strain At-9b)	1
Proteus mirabilis	1
Salmonella enterica subsp. enterica serovar Typhimurium str. TN061786	1
Salmonella paratyphi A (strain AKU_12601)	1
Shigella boydii ATCC 9905	1
Shigella dysenteriae 1617	1
Shigella dysenteriae CDC 74-1112	1

Shigella sonnei	1
Shigella sonnei 53G	1
Uncultured gamma proteobacterium HF0010_10D20	1
Uncultured Oceanospirillales bacterium HF4000_21D01	1
Vibrio cholerae serotype O1 (strain ATCC 39541 / Ogawa 395 / O395)	1
Wolbachia endosymbiont of Drosophila simulans	1
Yersinia pestis Pestoides A	1

**Figure 1: Taxonomic origin of proteins responsible of the prediction of the detected genes**





**Figure 2: Taxonomic origin of proteins responsible of the prediction of the detected genes**

## Fast manual annotation

Based on the preliminary results of our semi-automated method of annotation we have reviewed manually the annotation tagging the principal genes and functions. This preliminary tagging has been carried out analyzing the annotation for each predicted protein following specific interesting points: toxins, hemolysins, antibiotic resistance, pathogenicity, adhesion, plasmid, phage and other features.

We have selected and clustered genes with specific functions especially important from the human health perspective. The tagged genes are displayed in the following simplified tables. The complete annotation table is available at:

[http://www.era7bioinformatics.com/docs/EHEC\\_E\\_COLI\\_GERMANY\\_OUTBREAK\\_Annotation\\_Era7\\_Bioinformatics\\_v1\\_5\\_6\\_2011.xls](http://www.era7bioinformatics.com/docs/EHEC_E_COLI_GERMANY_OUTBREAK_Annotation_Era7_Bioinformatics_v1_5_6_2011.xls)

## Table 2: Toxins and peptidases

In this table we have selected predicted proteins with annotations related with toxin function and some peptidases that could be related with toxin-like activity.

Contig ID	Gen ID	start	end	S	tags	Similar to	Protein names	Organism
husec41_c1134	36180	829	1278	+	Toxin	Q83Z99	Putative acyltransferase MchD	Escherichia coli
husec41_c1252	46188	1087	1662	+	Toxin	D8CGU2	Toxin-antitoxin system, toxin component, PIN family	Escherichia coli MS 185-1
husec41_c136	40883	561	271	-	Toxin	D7YRT1	Toxin-antitoxin system, antitoxin component, Xre family	Escherichia coli MS 182-1
husec41_c145	66526	1166	1423	+	Toxin	C8THQ6	Predicted antitoxin of YafQ-DinJ toxin-antitoxin system	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c145	54284	1429	1704	+	Toxin	B7LHF6	Toxin of the YafQ-DinJ toxin-antitoxin system	Escherichia coli (strain 55989 / EAEC)



husec41_c1518	5119	1	1132	+	Toxin	Q8VSL2	Serine protease sepA autotransporter (EC 3.4.21.-) [Cleaved into: Serine protease sepA; Serine protease sepA translocator]	Shigella flexneri
husec41_c1554	45805	104	485	+	Toxin	C1NDL5	Secreted autotransporter toxin Sat	Escherichia sp. 1_1_43
husec41_c1932	50684	2	403	+	Toxin	B7LBW0	Serine protease pic (ShMu)	Escherichia coli (strain 55989 / EAEC)
husec41_c2009	41962	488	249	-	Toxin	D8ET41	Toxin-antitoxin system protein	Escherichia coli MS 107-1
husec41_c2132	76041	3712	6777	+	Toxin	B7MC95	Vacuolating autotransporter toxin	Escherichia coli O45:K1 (strain S88 / ExPEC)
husec41_c291	110361	1432	2418	+	Toxin	B7LBY5	Serine protease pet (Plasmid-encoded toxin pet) (EC 3.4.21.72)	Escherichia coli (strain 55989 / EAEC)
husec41_c380	110360	135	2288	+	Toxin	B7LBY5	Serine protease pet (Plasmid-encoded toxin pet) (EC 3.4.21.72)	Escherichia coli (strain 55989 / EAEC)
husec41_c482	84923	4515	4030	-	Toxin	C8TUV8	Membrane protein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c543	42847	2851	3272	+	Toxin	E6ARB5	Toxin-antitoxin system, antitoxin component, Xre family	Escherichia coli MS 16-3
husec41_c567	110363	3	803	+	Toxin	B7LBY5	Serine protease pet (Plasmid-encoded toxin pet) (EC 3.4.21.72)	Escherichia coli (strain 55989 / EAEC)
husec41_c581	45751	1411	1797	+	Toxin	E7J0S6	Putative shET2 enterotoxin	Escherichia coli 1357
husec41_c59	63485	4593	4940	+	Toxin	C8TP46	Toxin ChpB of the ChpB-ChpS toxin-antitoxin system	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c672	46347	2184	1269	-	Toxin	D7X7P6	ShET2 enterotoxin, region	Escherichia coli MS 198-1
husec41_c69	41563	3801	4068	+	Toxin	D8BT94	Toxin-antitoxin system, antitoxin component, HicB family	Escherichia coli MS 196-1
husec41_c74	69806	4058	3807	-	Toxin	C8TTV0	Toxin of the YoeB-YefM toxin-antitoxin system	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c74	68869	4333	4058	-	Toxin	A1ACN0	Antitoxin of the YoeB-YefM toxin-	Escherichia coli O1:K1 / APEC

							antitoxin system	
husec41_c746	89468	3163	3465	+	Toxin	C8TG29	Toxin ChpA	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c782	5124	1124	1528	+	Toxin	Q8VSL2	Serine protease sepA autotransporter (EC 3.4.21.-) [Cleaved into: Serine protease sepA; Serine protease sepA translocator]	Shigella flexneri
husec41_c797	36434	360	1	-	Toxin	D7ZZI1	Toxin-antitoxin system, antitoxin component, AbrB family	Escherichia coli MS 187-1
husec41_rep_c2274	50678	1	835	+	Toxin	B7LBW0	Serine protease pic (ShMu)	Escherichia coli (strain 55989 / EAEC)
husec41_rep_c2292	59806	323	3	-	Toxin	B7LBU0	Toxin of the YeeV-YeeU toxin-antitoxin system	Escherichia coli (strain 55989 / EAEC)
husec41_rep_c2292	25611	641	268	-	Toxin	E3PD62	Putative antitoxin	Escherichia coli O78:H11 (strain H10407 / ETEC)
husec41_rep_c2295	19119	3	1366	+	Toxin	B3HJU4	Serine protease eata (EC 3.4.21.-)	Escherichia coli B7A
husec41_rep_c2297	108631	295	29	-	Toxin	C6UP08	Shiga toxin II subunit B	Escherichia coli O157:H7 (strain TW14359 / EHEC)
husec41_rep_c2297	42022	1350	310	-	Toxin	B3BQ93	Shiga toxin subunit A (EC 3.2.2.22)	Escherichia coli O157:H7 str. EC869
husec41_rep_c2359	5123	121	618	+	Toxin	Q8VSL2	Serine protease sepA autotransporter (EC 3.4.21.-) [Cleaved into: Serine protease sepA; Serine protease sepA translocator]	Shigella flexneri
husec41_rep_c2607	104486	98	346	+	Toxin	B7MN77	Toxin of the YeeV-YeeU toxin-antitoxin system	Escherichia coli O45:K1 (strain S88 / ExPEC)
husec41_rep_c2848	5126	261	3	-	Toxin	Q8VSL2	Serine protease sepA autotransporter (EC 3.4.21.-) [Cleaved into: Serine protease sepA; Serine protease sepA translocator]	Shigella flexneri

husec41_c1549	66741	87	372	+	Toxin?	D3QPV9	Small toxic membrane polypeptide	Escherichia coli O55:H7 (strain CB9615 / EPEC)
husec41_c685	78980	170	463	+	Toxin?	E8YB84	Hok/gef cell toxic protein	Escherichia coli (strain ATCC 55124 / KO11)
husec41_c201	19118	1698	3	-	Toxin?	B3HJU4	Serine protease eata (EC 3.4.21.-)	Escherichia coli B7A
husec41_c1030	92954	1343	1858	+	protease	C8TI62	Predicted intracellular protease	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c108	112662	18221	17895	-	protease	C8TIC9	ClpXP protease specificity-enhancing factor	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c375	93700	5121	4207	-	protease	C8TJ82	Predicted membrane anchored protease	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c375	67160	9116	8491	-	protease	C8TJ87	Multifunctional acyl-CoA thioesterase I and protease I and lysophospholipase L1	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c397	61897	731	3	-	protease	C8TNZ5	Modulator for HflB protease	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c426	102625	1	612	+	protease	B7LC29	Modulator for HflB protease specific for phage lambda cII repressor	Escherichia coli (strain 55989 / EAEC)
husec41_c426	81158	622	1623	+	protease	C8TNZ6	Modulator for HflB protease	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c435	94605	2312	4164	+	protease	C8TST2	Protease IV	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1014	93222	6068	5204	-	peptidase	C8TSZ0	Protease HtpX (EC 3.4.24.-) (Heat shock protein HtpX)	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1033	96990	797	210	-	peptidase	C8UF29	Protease III	Escherichia coli O111:H- (strain 11128 / EHEC)
husec41_c1037	111767	2669	1388	-	peptidase	C8TVL8	Peptidase B (EC 3.4.11.23) (Aminopeptidase B)	Escherichia coli O26:H11 (strain 11368 / EHEC)

husec41_c108	66438	22197	23562	+	peptidase	C8TID5	Serine endoprotease DegQ, periplasmic	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c108	60523	23656	24135	+	peptidase	B7LHU3	Serine endoprotease, periplasmic (EC 3.4.21.-)	Escherichia coli (strain 55989 / EAEC)
husec41_c1290	78186	235	1	-	peptidase	D3GXP8	Putative peptidase	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c1290	63732	1224	229	-	peptidase	C8TI67	Predicted peptidase	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1335	63160	1148	399	-	peptidase	C8TU14	Predicted peptidase	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1397	66644	1941	2	-	peptidase	C8TT72	Protease II	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1409	107804	98	913	+	peptidase	D3GSR7	Protease III (Pitriylisin) (EC 3.4.24.55)	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c1462	102352	531	6	-	peptidase	B5Z0E1	Protease Do (EC 3.4.21.-)	Escherichia coli O157:H7 (strain EC4115 / EHEC)
husec41_c1563	48591	328	1148	+	peptidase	Q8XG80	Putative secreted protein	Salmonella typhi
husec41_c1633	103360	1534	3	-	peptidase	C8TG70	Protease III	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1652	104691	557	3	-	peptidase	D3GTU8	Putative peptidase	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c179	84429	2955	5002	+	peptidase	C8TSZ1	Carboxy-terminal protease for penicillin-binding protein 3	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1807	49156	813	84	-	peptidase	Q8Z3I4	Putative uncharacterized protein STY3458 (Putative uncharacterized protein yhbV)	Salmonella typhi
husec41_c221	75835	1078	197	-	peptidase	B7LDC6	Putative peptidase	Escherichia coli (strain 55989 / EAEC)

husec41_c237	99842	283	817	+	peptidase	C8TJ30	ATP-dependent Clp protease proteolytic subunit (EC 3.4.21.92) (Endopeptidase Clp)	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c256	71982	1	632	+	peptidase	C8TJL3	Rhomboid protease glpG (EC 3.4.21.105) (Intramembrane serine protease)	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c266	29610	105	955	+	peptidase	B7L8Y4	Chaperone protein hchA (Hsp31)	Escherichia coli (strain 55989 / EAEC)
husec41_c270	82448	511	308	-	peptidase	C8TI90	D-alanyl-D-alanine carboxypeptidase	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c292	88489	1728	1041	-	peptidase	C8TMQ5	Peptidase E (EC 3.4.13.21) (Alpha-aspartyl dipeptidase) (Asp-specific dipeptidase) (Dipeptidase E)	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c352	61049	1	1054	+	peptidase	C8TQ29	Peptidase T (EC 3.4.11.4) (Aminotripeptidase) (Tripeptide aminopeptidase)	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c46	14703	6657	5911	-	peptidase	E9XBV5	D-alanyl-D-alanine carboxypeptidase	Escherichia coli H120
husec41_c492	59662	7133	8342	+	peptidase	C8TKF4	D-alanyl-D-alanine carboxypeptidase	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c495	60308	2896	1402	-	peptidase	C8TJW4	Predicted zinc-dependent peptidase	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c5	95828	2380	1237	-	peptidase	C8TTU3	D-alanyl-D-alanine carboxypeptidase	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c521	106852	4249	1905	-	peptidase	B7LFN6	Putative uncharacterized protein	Escherichia coli (strain 55989 / EAEC)
husec41_c522	103530	8385	7260	-	peptidase	C8TVG5	Succinyl-diaminopimelate desuccinylase (SDAP desuccinylase) (EC 3.5.1.18) (N-succinyl-LL-2,6-diaminoheptanedioate amidohydrolase)	Escherichia coli O26:H11 (strain 11368 / EHEC)

husec41_c532	111201	2889	109	-	peptidase	C8TRL4	Predicted peptidase	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c538	80157	2353	3102	+	peptidase	B7LC50	Putative hydrolase	Escherichia coli (strain 55989 / EAEC)
husec41_c558	95229	748	1146	+	peptidase	B7LGT6	Putative uncharacterized protein ycgK	Escherichia coli (strain 55989 / EAEC)
husec41_c590	60522	1	656	+	peptidase	B7LHU3	Serine endoprotease, periplasmic (EC 3.4.21.-)	Escherichia coli (strain 55989 / EAEC)
husec41_c591	111649	2463	3377	+	peptidase	C8TQK4	L,D-carboxypeptidase A	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c645	96639	2369	1334	-	peptidase	C8TQX0	Predicted inner membrane peptidase	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c657	89840	1179	1795	+	peptidase	B5YUD9	Peptidase, U32 family	Escherichia coli O157:H7 (strain EC4115 / EHEC)
husec41_c69	71210	1051	3009	+	peptidase	C8TRG1	Predicted peptidase	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c758	61491	1358	2143	+	peptidase	B5YT43	Peptidase, M48B family (EC 3.4.24.-)	Escherichia coli O157:H7 (strain EC4115 / EHEC)
husec41_c791	82447	1310	3	-	peptidase	C8TI90	D-alanyl-D-alanine carboxypeptidase	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c809	58434	2925	1725	-	peptidase	C8TLV5	D-alanyl-D-alanine carboxypeptidase DacC	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c833	67161	448	975	+	peptidase	C8TK00	ATP-dependent protease subunit HsIV (EC 3.4.25.2) (Heat shock protein HsIV)	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c84	106278	1	951	+	peptidase	B7LGM3	Serine endoprotease (Protease Do), membrane-associated (EC 3.4.21.-)	Escherichia coli (strain 55989 / EAEC)
husec41_c878	104156	2554	4050	+	peptidase	C8TP80	Probable cytosol aminopeptidase (EC 3.4.11.1) (Leucine aminopeptidase)	Escherichia coli O26:H11 (strain 11368 / EHEC)

							(Leucyl aminopeptidase)	
husec41_c88	92547	2555	710	-	peptidase	C8TM40	Predicted carboxypeptidase	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c919	110827	1534	1169	-	peptidase	C6UT42	Predicted peptidase	Escherichia coli O157:H7 (strain TW14359 / EHEC)

**Table 3: Hemolysins and heme metabolism related proteins**

Contig ID	Gen ID	start	end	S	tags	Similar to	Protein names	Organism
husec41_c1375	71299	2390	1337	-	Hemolysin	B5YXK6	Hemolysin E, chromosomal	Escherichia coli O157:H7 (strain EC4115 / EHEC)
husec41_c1786	63671	984	695	-	Hemolysin	Q0TDJ8	Putative hemolysin expression modulating protein	Escherichia coli O6:K15:H31 (strain 536 / UPEC)
husec41_c604	100245	211	864	+	Hemolysin	D3RHI6	Channel protein, hemolysin III family	Klebsiella variicola (strain At-22)
husec41_c1103	111924	1751	1305	-	Heme	C8U6D2	Heme lyase, CcmH subunit	Escherichia coli O103:H2 (strain 12009 / EHEC)
husec41_c1469	111923	672	2	-	Heme	C8U6D2	Heme lyase, CcmH subunit	Escherichia coli O103:H2 (strain 12009 / EHEC)
husec41_c1469	109614	3166	1239	-	Heme	C8TUD4	Heme lyase, CcmF subunit	Escherichia coli O26:H11 (strain 11368 / EHEC)

husec41_c1914	84812	451	36	-	Heme	C8TUD7	Heme exporter protein C	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c194	62795	796	164	-	Heme	C8TTP2	Sulfoxide reductase heme-binding subunit yedZ (Flavocytochrome yedZ)	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c231	55526	779	392	-	Heme	D3GXL9	Sulfoxide reductase heme-binding subunit yedZ (Flavocytochrome yedZ)	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c458	57035	782	2159	+	Heme	C8TIQ7	Siroheme synthase	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c656	88320	6148	5555	-	Heme	C8TMW6	Heme lyase, NrfG subunit	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c656	82983	6528	6148	-	Heme	C8TMW5	Heme lyase, NrfF subunit	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c656	99672	8184	6528	-	Heme	C8TMW4	Heme lyase, NrfE subunit	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c849	90431	5575	6195	+	Heme	B7LAM4	Heme exporter subunit ; ATP-binding component of ABC superfamily	Escherichia coli (strain 55989 / EAEC)
husec41_c849	110508	6195	6854	+	Heme	B7LAM3	Heme exporter subunit ; membrane component of ABC superfamily	Escherichia coli (strain 55989 / EAEC)
husec41_c849	84813	6899	7210	+	Heme	C8TUD7	Heme exporter protein C	Escherichia coli O26:H11 (strain 11368 / EHEC)



**Table 4: Antibiotic resistance.**

In table 4 we have selected genes involved in specific antibiotic resistance and also genes encoding efflux pumps and multidrug resistance proteins that could be involved in some additional antibiotic resistance capabilities of this strain.

Contig ID	Gen ID	start	end	S	tags	Similar to	Protein names	Organism
husec41_c1261	66325	2	1849	+	Aminoglycoside resistance	C8TVG3	<b>Aminoglycoside</b> /multidrug efflux system protein AcrD	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c177	66326	3	2575	+	Aminoglycoside resistance	C8TVG3	<b>Aminoglycoside</b> /multidrug efflux system protein AcrD	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c626	66329	1284	23	-	Aminoglycoside resistance	C8TVG3	<b>Aminoglycoside</b> /multidrug efflux system protein AcrD	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c697	66327	3298	151	-	Aminoglycoside resistance	C8TVG3	<b>Aminoglycoside</b> /multidrug efflux system protein AcrD	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c738	66328	3	3033	+	Aminoglycoside resistance	C8TVG3	<b>Aminoglycoside</b> /multidrug efflux system protein AcrD	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c312	94767	270	3	-	Macrolide resistance	C8TLZ5	Fused <b>macrolide</b> transporter subunits of ABC superfamily: ATP-binding component/membrane component	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c584	94764	3955	2332	-	Macrolide resistance	C8TLZ5	Fused <b>macrolide</b> transporter subunits of ABC superfamily: ATP-binding component/membrane component	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c312	94763	1384	270	-	Macrolide resistance	C8TLZ4	<b>Macrolide</b> transporter subunit, membrane fusion protein component	Escherichia coli O26:H11 (strain 11368 / EHEC)

husec41_c48	30491	2597	1778	-	Penicillin resistance	B7LBI2	<b>Penicillin</b> -insensitive murein endopeptidase (EC 3.4.24.-) (D-alanyl-D-alanine-endopeptidase) (DD-endopeptidase)	Escherichia coli (strain 55989 / EAEC)
husec41_c1681	110509	694	942	+	Polymyxin resistance	B7LAS5	<b>Polymyxin</b> resistance protein B	Escherichia coli (strain 55989 / EAEC)
husec41_c497	12057	676	18	-	Polymyxin resistance	D6I9J9	<b>Polymyxin</b> resistance protein PmrM	Escherichia coli B185
husec41_c2170	101844	1528	671	-	beta-lactam resistance	C8UQP5	TEM-1 <b>beta-lactamase</b>	Escherichia coli O111:H- (strain 11128 / EHEC)
husec41_c642	87098	2345	1148	-	Tetracycline resistance	D3H382	<b>Tetracycline</b> resistance protein	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c559	82209	4864	5178	+	beta-lactam resistance	C8TJ28	Regulator of penicillin binding proteins and <b>beta-lactamase</b> transcription	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1436	94598	2158	931	-	Tetracycline resistance	C8TNP7	Multidrug resistance protein mdtG	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c55	72338	3557	2312	-	Fosfomycin and deoxycholate resistance	C8TU05	Multidrug resistance protein mdtA (Multidrug transporter mdtA)	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1334	58199	4219	3307	-	Novobiocin and deoxycholate resistance	C6V0N5	Multidrug resistance protein MdtB (Multidrug transporter MdtB)	Escherichia coli O157:H7 (strain TW14359 / EHEC)
husec41_c1334	62131	3303	225	-	Novobiocin and deoxycholate resistance	C8TU07	Multidrug resistance protein MdtC (Multidrug transporter MdtC)	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c55	62133	2267	13	-	Novobiocin and deoxycholate resistance	C8TU07	Multidrug resistance protein MdtC (Multidrug transporter MdtC)	Escherichia coli O26:H11 (strain 11368 / EHEC)

husec41_c1063	30470	2629	1456	-	chloramphenicol resistance	B7L855	Multidrug resistance protein mdtL	Escherichia coli (strain 55989 / EAEC)
husec41_c654	59511	3708	4838	+	beta-lactam resistance	C8TNX4	<b>Beta-lactamase</b> /D-alanine carboxypeptidase AmpC	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c229	90034	4731	5885	+	beta-lactam resistance	C8TIW8	<b>Beta-lactamase</b> /D-alanine carboxypeptidase	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1120	93622	903	706	-	Beta-lactamic resistance	B7L5H7	<b>Beta-lactam</b> resistance membrane protein	Escherichia coli (strain 55989 / EAEC)
husec41_c30	34317	1963	1039	-	Beta-lactamic resistance	Q6BBP7	<b>Beta-lactamase</b> (Beta-lactamase CTX-M-3) (CTX-M-3 extended-spectrum beta-lactamase) (Extended-spectrum class A beta-lactamase CTX-M-3)	Escherichia coli
husec41_rep_c2845	16578	194	3	-	Beta-lactamic resistance	B2CD48	<b>Beta-lactamase</b> TEM (Fragment)	Escherichia coli
husec41_c132	57069	2484	3222	+	Bicyclomycin resistance	B7LAK6	<b>Bicyclomycin</b> /multidrug efflux system	Escherichia coli (strain 55989 / EAEC)
husec41_c720	91855	4731	4495	-	Bicyclomycin resistance	C6V2S0	<b>Bicyclomycin</b> /multidrug efflux system	Escherichia coli O157:H7 (strain TW14359 / EHEC)
husec41_c1017	67341	200	1048	+	Polymyxin resistance	C8TUQ2	Bifunctional <b>polymyxin</b> resistance protein ArnA	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c497	67340	4376	3269	-	Polymyxin resistance	C8TUQ2	Bifunctional <b>polymyxin</b> resistance protein ArnA	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c770	30454	1469	2675	+	norfloxacin and enoxacin resistance	B7LFZ9	Multidrug resistance protein mdtH	Escherichia coli (strain 55989 / EAEC)
husec41_c310	79562	1741	560	-	6-mercaptopurine resistance	C8TLE9	Purine ribonucleoside efflux pump nepI	Escherichia coli O26:H11 (strain 11368 / EHEC)

husec41_c212	106418	13194	12706	-	Antibiotic resistance related?	B5Z494	Drug resistance transporter, Bcr/CflA subfamily	Escherichia coli O157:H7 (strain EC4115 / EHEC)
husec41_c642	68106	2452	3099	+	Antibiotic resistance related?	D3H381	Tetracycline repressor	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c18	111861	2412	1280	-	Antibiotic resistance related?	C8TME8	Membrane fusion protein (MFP) component of efflux pump, signal anchor	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c348	59452	1242	2272	+	Antibiotic resistance related?	B7LB31	Membrane fusion protein of efflux pump	Escherichia coli (strain 55989 / EAEC)
husec41_c1372	77754	2	178	+	Antibiotic resistance related?	D3GVT1	Modulator of drug activity B	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c1352	97206	3	942	+	Antibiotic resistance related?	B7LEA2	Multidrug efflux system	Escherichia coli (strain 55989 / EAEC)
husec41_c348	89848	2387	4188	+	Antibiotic resistance related?	B7LB24	Multidrug efflux system component	Escherichia coli (strain 55989 / EAEC)
husec41_c1670	95215	665	1261	+	Antibiotic resistance related?	B7LHX4	Multidrug efflux system protein	Escherichia coli (strain 55989 / EAEC)
husec41_c684	66306	397	1424	+	Antibiotic resistance related?	C8TQ83	Multidrug efflux system protein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c709	75513	1460	1245	-	Antibiotic resistance related?	C8UN45	Multidrug efflux system protein	Escherichia coli O111:H- (strain 11128 / EHEC)

husec41_c697	86040	4513	3324	-	Antibiotic resistance related?	C8TJ55	Multidrug efflux system protein AcrA	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1084	83109	107	1337	+	Antibiotic resistance related?	C8TLV8	Multidrug efflux system protein Cmr	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c75	86744	3472	2170	-	Antibiotic resistance related?	C8TFT3	Multidrug efflux system protein EmrB	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c654	103489	5777	5463	-	Antibiotic resistance related?	C8TNX2	Multidrug efflux system protein SugE	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c214	54051	3736	2598	-	Antibiotic resistance related?	B5YX79	Multidrug resistance protein D	Escherichia coli O157:H7 (strain EC4115 / EHEC)
husec41_c212	88104	10368	8997	-	Antibiotic resistance related?	C8TSE0	Multidrug resistance protein mdtK (Multidrug-efflux transporter)	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c288	18218	391	11	-	Antibiotic resistance related?	Q3I3P0	ORF3-QacEdelta1 fusion protein	Escherichia coli
husec41_c656	90365	2	682	+	Antibiotic resistance related?	B7LB29	Outer membrane factor of efflux pump	Escherichia coli (strain 55989 / EAEC)
husec41_c407	99788	250	840	+	Antibiotic resistance related?	C8TJM6	Predicted antibiotic transporter	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c993	68739	4215	2530	-	Antibiotic resistance related?	C8TTQ9	Predicted multidrug efflux system	Escherichia coli O26:H11 (strain 11368 / EHEC)

husec41_c1456	111119	1023	1	-	Antibiotic resistance related?	C8TV77	Predicted multidrug efflux system protein Y	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c481	111121	953	1525	+	Antibiotic resistance related?	C8TV77	Predicted multidrug efflux system protein Y	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1726	77783	3	666	+	Antibiotic resistance related?	C8ULT4	Predicted outer membrane factor of efflux pump	Escherichia coli O111:H- (strain 11128 / EHEC)
husec41_c1014	109696	3780	5152	+	Antibiotic resistance related?	C8TSY9	Predicted transporter	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c497	28091	3269	2382	-	Antibiotic resistance related?	B7LAS1	Probable 4-deoxy-4-formamido-L-arabinose-phosphoundecaprenol deformylase ArnD (EC 3.5.1.n3)	Escherichia coli (strain 55989 / EAEC)
husec41_c1334	105658	221	3	-	Antibiotic resistance related?	D3GZE1	Putative multidrug resistance protein mdtD	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c464	109529	2	1296	+	Antibiotic resistance related?	C8TU08	Putative multidrug resistance protein mdtD	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1656	54287	399	1	-	Antibiotic resistance related?	B7L8U8	Putative multidrug resistance protein; DLP12 prophage	Escherichia coli (strain 55989 / EAEC)
husec41_c1277	106968	879	1	-	Antibiotic resistance related?	B7L7J5	Putative multidrug transporter fused subunits of ABC superfamily transporter: permease component; ATP-binding component	Escherichia coli (strain 55989 / EAEC)

husec41_c225	85500	6378	7532	+	Antibiotic resistance related?	C8TUQ0	UDP-4-amino-4-deoxy-L-arabinose--oxoglutarate aminotransferase (EC 2.6.1.87) (UDP-(beta-L-threo-pentapyranosyl-4''-ulose diphosphate) aminotransferase) (UDP-4-amino-4-deoxy-L-arabinose aminotransferase)	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1017	75853	3	197	+	Antibiotic resistance related?	C6V2Y6	Undecaprenyl-phosphate 4-deoxy-4-formamido-L-arabinose transferase (EC 2.7.8.30) (Undecaprenyl-phosphate Ara4FN transferase)	Escherichia coli O157:H7 (strain TW14359 / EHEC)
husec41_c225	57126	7539	7742	+	Antibiotic resistance related?	D3H1C1	Undecaprenyl-phosphate 4-deoxy-4-formamido-L-arabinose transferase (EC 2.7.8.30) (Undecaprenyl-phosphate Ara4FN transferase)	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c181	28083	2	670	+	Antibiotic resistance related?	B7LAR9	Undecaprenyl-phosphate 4-deoxy-4-formamido-L-arabinose transferase (EC 2.7.8.30) (Undecaprenyl-phosphate Ara4FN transferase) (Ara4FN transferase)	Escherichia coli (strain 55989 / EAEC)
husec41_c908	66282	3242	1862	-	Antibiotic resistance related?	C8TMT4	DNA-damage-inducible SOS response protein DinF	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c481	103492	3	950	+	Antibiotic resistance related?	C8TV78	EmrKY-TolC multidrug resistance efflux pump protein K, membrane fusion protein component	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c2258	104799	370	138	-	Antibiotic resistance related?	B7LBS1	EmrKY-TolC multidrug resistance efflux pump, membrane fusion protein component	Escherichia coli (strain 55989 / EAEC)

husec41_c2050	83548	121	388	+	Antibiotic resistance related?	C8U6F0	Fused predicted multidrug transport subunits of ABC superfamily: membrane component/ATP-binding component	Escherichia coli O103:H2 (strain 12009 / EHEC)
husec41_c931	95462	1	1347	+	Antibiotic resistance related?	C8TUF0	Fused predicted multidrug transport subunits of ABC superfamily: membrane component/ATP-binding component	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c449	97110	3737	1960	-	Antibiotic resistance related?	C8TJ42	Fused predicted multidrug transporter subunits of ABC superfamily: ATP-binding components	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c449	74859	5504	3733	-	Antibiotic resistance related?	C8TJ41	Fused predicted multidrug transporter subunits of ABC superfamily: ATP-binding components	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c897	107526	69	812	+	Antibiotic resistance related?	C8U8G6	Fused predicted multidrug transporter subunits of ABC superfamily: membrane component/ATP-binding component	Escherichia coli O103:H2 (strain 12009 / EHEC)
husec41_c1382	111871	312	1	-	Antibiotic resistance related?	B7L9Q8	Wzx	Escherichia coli (strain 55989 / EAEC)
husec41_c2032	19066	290	2	-	Antibiotic resistance related?	Q93NP6	Wzx	Escherichia coli
husec41_c2082	19068	3	197	+	Antibiotic resistance related?	Q93NP6	Wzx	Escherichia coli
husec41_c2085	111873	241	2	-	Antibiotic resistance related?	B7L9Q8	Wzx	Escherichia coli (strain 55989 / EAEC)



husec41_c832	75877	3646	3266	-	Antibiotic resistance related?	C8TRP6	DNA-binding transcriptional dual activator MarA of multiple antibiotic resistance	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c832	94641	4100	3669	-	Antibiotic resistance related?	C8TRP5	DNA-binding transcriptional repressor MarR of multiple antibiotic resistance	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c795	88581	4640	6397	+	Antibiotic resistance related?	C8THH6	Fused glycosyl transferase and transpeptidase	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c440	88099	1589	4137	+	Antibiotic resistance related?	C8TJI6	Fused penicillin-binding protein 1a: murein transglycosylase/murein transpeptidase	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1203	111077	1355	197	-	Antibiotic resistance related?	C8TJV0	Multidrug resistance efflux transporter MdtE	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1099	78392	6	794	+	Antibiotic resistance related?	D3GR96	Penicillin-binding protein 1B [includes: penicillin-insensitive transglycosylase; penicillin sensitive transpeptidase] (EC 2.4.1.129) (EC 3.4.-.-)	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c1011	84329	914	298	-	Antibiotic resistance related?	C6UUQ5	p-hydroxybenzoic acid efflux pump subunit AaeA (pHBA efflux pump protein A)	Escherichia coli O157:H7 (strain TW14359 / EHEC)
husec41_c1123	65704	391	14	-	Antibiotic resistance related?	C8TIE2	p-hydroxybenzoic acid efflux pump subunit AaeA (pHBA efflux pump protein A)	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1011	101419	289	2	-	Antibiotic resistance related?	C8TIE1	p-hydroxybenzoic acid efflux pump subunit AaeB (pHBA efflux pump protein B)	Escherichia coli O26:H11 (strain 11368 / EHEC)

husec41_c150	101418	2121	470	-	Antibiotic resistance related?	C8TIE1	p-hydroxybenzoic acid efflux pump subunit AaeB (pHBA efflux pump protein B)	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c228	99138	1021	143	-	Antibiotic resistance related?	C6UU60	Predicted methyl viologen efflux pump	Escherichia coli O157:H7 (strain TW14359 / EHEC)
husec41_c1148	100146	922	2262	+	Antibiotic resistance related?	C8TL54	Predicted multidrug or homocysteine efflux system protein HsrA	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c124	54288	252	1	-	Antibiotic resistance related?	B7L8U8	Putative multidrug resistance protein; DLP12 prophage	Escherichia coli (strain 55989 / EAEC)

**Table 5: Adhesion related, secretion system and pathogenicity and virulence related proteins**

Contig ID	Gen ID	start	end	S	tags	Similar to	Protein names	Organism
husec41_c1007	74820	2	517	+	Adhesion	C8ULG1	Predicted fimbrial-like adhesin protein	Escherichia coli O111:H- (strain 11128 / EHEC)
husec41_c101	53306	1	2252	+	Adhesion	C8TPC9	Adhesin AIDA-I	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1040	93676	642	198	-	Adhesion	C8TNN7	Curlin major subunit CsgA	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c110	111590	1573	2158	+	Adhesion	C8U1Y5	Predicted fimbrillin	Escherichia coli O103:H2 (strain 12009 / EHEC)
husec41_c117	111092	3162	2	-	Adhesion	C8TI10	AidA-I adhesin-like protein	Escherichia coli O26:H11 (strain 11368 / EHEC)

husec41_c123	65701	1318	2408	+	Adhesion	C8TI54	Predicted fimbrial-like adhesin protein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1256	104796	2339	1841	-	Adhesion	B7LH27	Putative adhesin major subunit pilin	Escherichia coli (strain 55989 / EAEC)
husec41_c128	67824	3982	4510	+	Adhesion	C8U8H5	Predicted fimbrial-like adhesin protein	Escherichia coli O103:H2 (strain 12009 / EHEC)
husec41_c1306	71578	1086	19	-	Adhesion	C8TM57	Predicted fimbrial-like adhesin protein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1331	79683	1335	1	-	Adhesion	C8TUN4	Adhesin YfaL	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1427	88182	352	11	-	Adhesion	C8U5A7	Predicted fimbrial-like adhesin protein	Escherichia coli O103:H2 (strain 12009 / EHEC)
husec41_c1427	62747	902	363	-	Adhesion	C8U5A6	Predicted fimbrial-like adhesin protein	Escherichia coli O103:H2 (strain 12009 / EHEC)
husec41_c1503	26686	1064	1736	+	Adhesion	C6UM96	Predicted fimbrial-like adhesin protein	Escherichia coli (strain B / REL606)
husec41_c151	102071	2	1744	+	Adhesion	C8TNL5	Putative AidA-I adhesin-like protein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1510	68137	24	1781	+	Adhesion	C8TQI1	AidA-I adhesin-like protein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1570	61843	1	1035	+	Adhesion	C8TGZ3	Putative Iha adhesin	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1732	107162	721	2	-	Adhesion	C8TZB2	Putative adhesin	Escherichia coli O103:H2 (strain 12009 / EHEC)
husec41_c1801	26687	510	1	-	Adhesion	C6UM96	Predicted fimbrial-like adhesin protein	Escherichia coli (strain B / REL606)

husec41_c1871	74608	3	1188	+	Adhesion	C8TLM3	Putative adhesin	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c2014	85241	951	136	-	Adhesion	D3GX42	Putative adhesin autotransporter	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c2017	71758	645	1426	+	Adhesion	C8TTS6	AidA-I adhesin-like protein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c202	60438	2197	1167	-	Adhesion	C8TU34	Predicted fimbrial-like adhesin protein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c279	78547	51	632	+	Adhesion	C8TI51	Predicted fimbrial-like adhesin protein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c297	61844	1535	2622	+	Adhesion	C8TGZ3	Putative Iha adhesin	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c311	105224	1682	277	-	Adhesion	C6V2W1	Adhesin	Escherichia coli O157:H7 (strain TW14359 / EHEC)
husec41_c353	85943	693	1589	+	Adhesion	B7L4E2	Putative uncharacterized protein	Escherichia coli (strain 55989 / EAEC)
husec41_c571	97235	49	1040	+	Adhesion	C8UG07	Conserved predicted protein	Escherichia coli O111:H- (strain 11128 / EHEC)
husec41_c678	79682	201	1390	+	Adhesion	C8TUN4	Adhesin YfaL	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c749	76923	100	639	+	Adhesion	C8TJB6	Predicted fimbrial-like adhesin protein SfmA	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c822	90032	53	733	+	Adhesion	C8TKM1	Predicted fimbrial-like adhesin protein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c92	93858	1136	162	-	Adhesion	C8TJB9	Predicted fimbrial-like adhesin protein SfmH	Escherichia coli O26:H11 (strain 11368 / EHEC)

husec41_c98	61848	16	564	+	Adhesion	C8TH66	Predicted fimbrial-like adhesin protein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_rep_c2345	102079	422	2	-	Adhesion	C8TNL5	Putative AidA-I adhesin-like protein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_rep_c2430	26688	282	13	-	Adhesion	C6UM96	Predicted fimbrial-like adhesin protein	Escherichia coli (strain B / REL606)
husec41_rep_c2436	53313	2	635	+	Adhesion	C8TPC9	Adhesin AIDA-I	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1650	54314	2971	3963	+	Flagellum	B7LBH5	Flagella control of anti-sigma factor FlgM secretion into the periplasm	Escherichia coli (strain 55989 / EAEC)
husec41_c106	69548	1443	350	-	Flagellum	C8TPF0	Flagellar P-ring protein 2 (Basal body P-ring protein 2)	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c106	66532	2153	1458	-	Flagellum	C8TPE9	Flagellar L-ring protein 2 (Basal body L-ring protein 2)	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c106	64941	2988	2208	-	Flagellum	C8TPE8	Flagellar component FlgG of cell-distal portion of basal-body rod	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c106	69184	3917	3163	-	Flagellum	B7LG11	Flagellar component of cell-proximal portion of basal-body rod	Escherichia coli (strain 55989 / EAEC)
husec41_c106	111682	5144	3940	-	Flagellum	C8TPE6	Flagellar hook protein FlgE	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c106	89749	5865	5172	-	Flagellum	C8TPE5	Flagellar hook assembly protein FlgD	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c106	80984	6281	5880	-	Flagellum	C8TPE4	Flagellar component FlgC of cell-proximal portion of basal-body rod	Escherichia coli O26:H11 (strain 11368 / EHEC)

husec41_c106	100036	6703	6288	-	Flagellum	C8TPE3	Flagellar component FlgB of cell-proximal portion of basal-body rod	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c106	86915	6871	7515	+	Flagellum	C8TPE2	Assembly protein FlgA for flagellar basal-body periplasmic P ring	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c106	76560	7892	8305	+	Flagellum	B7LG04	Export chaperone for FlgK and FlgL	Escherichia coli (strain 55989 / EAEC)
husec41_c1100	54209	1008	57	-	Flagellum	B7LG17	Flagellar hook-filament junction protein	Escherichia coli (strain 55989 / EAEC)
husec41_c1100	109008	2664	1023	-	Flagellum	C8TPF2	Flagellar hook-filament junction protein FlgK	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1100	94938	3324	2733	-	Flagellum	B5YVV1	Flagellar rod assembly protein/muramidase FlgJ	Escherichia coli O157:H7 (strain EC4115 / EHEC)
husec41_c1317	81793	1569	853	-	Flagellum	C8TJW2	EAL domain containing protein involved in flagellar function	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1479	60328	470	859	+	Flagellum	C8TTB9	Conserved predicted protein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1600	34314	527	883	+	Flagellum	Q4W8H8	Flagellar transcriptional regulator FlhD 1	Shigella dysenteriae
husec41_c1656	103938	878	566	-	Flagellum	D3GXI6	Flagellar hook-basal body complex protein FliE 2	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c178	101183	8378	7331	-	Flagellum	B7L8T6	Flagellin	Escherichia coli (strain 55989 / EAEC)
husec41_c178	104604	8626	10040	+	Flagellum	C6UYC8	Flagellar filament capping protein	Escherichia coli O157:H7 (strain TW14359 / EHEC)

husec41_c180	84613	1132	2277	+	Flagellum	C8TTC1	Predicted flagellar export pore protein FlhB	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c180	70181	2273	3936	+	Flagellum	C8TTC0	Predicted flagellar export pore protein FlhA	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c229	72306	8963	6603	-	Flagellum	B7L524	Putative flagellin structural protein; putative exported protein	Escherichia coli (strain 55989 / EAEC)
husec41_c27	73300	1039	38	-	Flagellum	C8TTM1	Flagellar motor switching and energizing component FliM	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c27	104771	1508	1047	-	Flagellum	C8TTM0	Flagellar biosynthesis protein FliL	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c27	106776	2658	1616	-	Flagellum	B7L8V5	Flagellar hook-length control protein	Escherichia coli (strain 55989 / EAEC)
husec41_c27	72296	3181	2741	-	Flagellum	C8TTL8	Flagellar protein FliJ	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c27	76765	4573	3203	-	Flagellum	C8TTL7	Flagellum-specific ATP synthase FliI	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c27	74878	5259	4576	-	Flagellum	B7L8V2	Flagellar biosynthesis protein	Escherichia coli (strain 55989 / EAEC)
husec41_c27	77068	6249	5255	-	Flagellum	C8TTL5	Flagellar motor switching and energizing component protein FliG	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c27	54779	7901	6245	-	Flagellum	C8TTL4	Flagellar basal-body MS-ring and collar protein FliF	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c518	87009	1	366	+	Flagellum	C8TTK6	Flagellar protein FliS, potentiates polymerization	Escherichia coli O26:H11 (strain 11368 / EHEC)

husec41_c518	85404	369	726	+	Flagellum	D3GXH8	Flagellar protein FliT	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c591	69223	1016	1747	+	Flagellum	B7LGV4	Flagellar brake protein YcgR (Cyclic di-GMP binding protein YcgR)	Escherichia coli (strain 55989 / EAEC)
husec41_c677	55280	1590	19	-	Flagellum	C8THR0	Predicted lateral flagellar system protein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c723	58591	28	265	+	Flagellum	C8TTM2	Flagellar motor switching and energizing component FliN	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c723	95166	271	632	+	Flagellum	B7L8V9	Flagellar biosynthesis protein	Escherichia coli (strain 55989 / EAEC)
husec41_c723	94985	635	1368	+	Flagellum	B5YRX7	Flagellar biosynthetic protein FliP	Escherichia coli O157:H7 (strain EC4115 / EHEC)
husec41_c723	105503	1381	1647	+	Flagellum	C8TTM5	Flagellar biosynthesis protein FliQ	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c723	57335	1659	2441	+	Flagellum	B7L8W2	Flagellar biosynthetic protein fliR	Escherichia coli (strain 55989 / EAEC)
husec41_c751	93919	2309	3247	+	Flagellum	C8TUT4	DNA-binding transcriptional repressor LrhA of flagellar, motility and chemotaxis genes	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c807	66266	938	438	-	Flagellum	C8TMI2	Predicted outer membrane lipoprotein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1479	100867	21	467	+	Flagellum	C8UC56	Predicted flagellar export pore protein FliA	Escherichia coli O111:H- (strain 11128 / EHEC)
husec41_c1940	23875	2	489	+	immune sistem interaction	B7N797	Surface presentation of antigens protein	Escherichia coli O17:K52:H18 (strain UMN026 / ExPEC)



husec41_c1047	103919	916	1893	+	invasin	D3GYK0	Putative invasin	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c215	92470	6900	2649	-	invasin	C8THZ4	Putative invasin	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c540	52585	589	2029	+	invasin	Q8Z7G3	Putative invasin	Salmonella typhi
husec41_c955	103920	1672	1091	-	invasin	D3GYK0	Putative invasin	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c1172	106992	1	1838	+	pathogenesis	B7LBK5	Putative uncharacterized protein	Escherichia coli (strain 55989 / EAEC)
husec41_c2051	25619	713	171	-	pathogenesis	C8TGA6	Conserved predicted protein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_rep_c2406	71603	107	397	+	pathogenesis	B7LF41	Putative uncharacterized protein	Escherichia coli (strain 55989 / EAEC)
husec41_c1167	99224	791	77	-	Secretion system	C8TGA7	Type III secretion system lipoprotein EprK	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c146	76579	2079	1546	-	Secretion system	B7LG99	Putative secretion pathway M-type protein, membrane anchored	Escherichia coli (strain 55989 / EAEC)
husec41_c146	102053	3259	2084	-	Secretion system	B7LGA0	Putative secretion pathway protein, L-type protein	Escherichia coli (strain 55989 / EAEC)
husec41_c146	61761	4235	3259	-	Secretion system	B7LGA1	Putative type II secretion protein (GspK-like)	Escherichia coli (strain 55989 / EAEC)
husec41_c146	91476	4837	4241	-	Secretion system	B7LG98	Putative uncharacterized protein	Escherichia coli (strain 55989 / EAEC)
husec41_c146	84087	5205	4837	-	Secretion system	B7LGA2	Putative type II secretion protein (GspI-like)	Escherichia coli (strain 55989 / EAEC)

husec41_c146	76580	5723	5205	-	Secretion system	B7LGA4	Putative uncharacterized protein	Escherichia coli (strain 55989 / EAEC)
husec41_c146	57151	6252	5773	-	Secretion system	B7LGA5	Putative general secretion pathway protein G (EpsG-like)	Escherichia coli (strain 55989 / EAEC)
husec41_c146	106738	7465	6245	-	Secretion system	B7LGA3	General secretion pathway protein F	Escherichia coli (strain 55989 / EAEC)
husec41_c146	105746	8959	7468	-	Secretion system	D3GVL9	Type II secretion system protein E	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c146	71843	10826	8892	-	Secretion system	B7LGA8	General secretion pathway protein D	Escherichia coli (strain 55989 / EAEC)
husec41_c146	72393	12009	11052	-	Secretion system	B7LGA9	Putative secretion pathway protein, C-type protein	Escherichia coli (strain 55989 / EAEC)
husec41_c1468	14001	413	2	-	Secretion system	D8AT29	Type III secretion system protein PrgH-EprH	Escherichia coli MS 116-1
husec41_c1538	69188	492	1	-	Secretion system	B7LF39	Putative Type III secretion EprH protein	Escherichia coli (strain 55989 / EAEC)
husec41_c1598	84444	192	449	+	Secretion system	C8TGB3	Type III secretion protein EpaQ	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1598	90155	454	834	+	Secretion system	B7LF43	Putative Type III secretion protein EpaR	Escherichia coli (strain 55989 / EAEC)
husec41_c16	74266	276	52	-	Secretion system	D3GUQ5	Microcin H47	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c16	67217	513	307	-	Secretion system	D3GUQ6	Microcin H47 immunity protein	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c1639	74493	1007	4	-	Secretion system	C8TGF0	T3SS effector-like protein EspX-homolog	Escherichia coli O26:H11 (strain 11368 / EHEC)

husec41_c1643	9758	652	750	+	Secretion system	E9Z029	DotU family protein type IV/VI secretion system protein	Escherichia coli M863
husec41_c1674	85180	703	1	-	Secretion system	C8U8E4	T3SS effector-like protein EspR-homolog	Escherichia coli O103:H2 (strain 12009 / EHEC)
husec41_c1783	75594	1429	1010	-	Secretion system	D3GUZ5	Putative type VI secretion protein	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c1791	11194	371	1	-	Secretion System	Q3SBC5	EpaS1	Escherichia coli
husec41_c1791	90154	832	383	-	Secretion system	B7LF43	Putative Type III secretion protein EpaR	Escherichia coli (strain 55989 / EAEC)
husec41_c1849	71007	243	1	-	Secretion system	D3GV03	Putative type VI secretion protein	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c1983	71004	856	1	-	Secretion system	D3GV03	Putative type VI secretion protein	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c2008	71005	11	646	+	Secretion system	D3GV03	Putative type VI secretion protein	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c2036	71006	3	428	+	Secretion system	D3GV03	Putative type VI secretion protein	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c2178	103825	354	163	-	Secretion system	D3GUY9	Putative type VI secretion protein	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c323	96371	2	1412	+	Secretion system	D3GUZ0	Putative type VI secretion protein	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c323	82460	1425	1929	+	Secretion system	D3GUZ1	Putative type VI secretion protein	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c333	95705	2539	1340	-	Secretion system	C8THD4	Assembly protein HofC in type IV pilin biogenesis, transmembrane protein	Escherichia coli O26:H11 (strain 11368 / EHEC)

husec41_c343	69947	966	1166	+	Secretion system	C8TKE9	Sec-independent protein translocase protein tatA/E homolog 1	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c539	14843	534	1031	+	Secretion System	Q3SBB5	EpaO	Escherichia coli
husec41_c542	106725	186	764	+	Secretion System	B7LF47	Putative type III secretion system protein EpaP	Escherichia coli (strain 55989 / EAEC)
husec41_c576	94943	3858	3274	-	Secretion system	B5YZD3	Secretion monitor protein	Escherichia coli O157:H7 (strain EC4115 / EHEC)
husec41_c716	86738	7	510	+	Secretion system	C8TIZ8	SecYEG protein translocase auxillary subunit SecD	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c716	105425	524	1557	+	Secretion system	C8TIZ9	SecYEG protein translocase auxillary subunit SecF	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c72	100206	1409	3509	+	Secretion system	D3GUQ1	Probable microcin H47 secretion/processing ATP-binding protein (EC 3.4.22.-)	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c734	86737	1422	2	-	Secretion system	C8TIZ8	SecYEG protein translocase auxillary subunit SecD	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c734	101983	1728	1453	-	Secretion system	C8TIZ7	SecYEG protein translocase auxillary subunit	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c771	111502	199	1558	+	Secretion system	D3GU38	Putative type VI secretion protein	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c813	112612	26	1597	+	Secretion system	C8UKR4	T3SS effector-like protein EspL-homolog	Escherichia coli O111:H- (strain 11128 / EHEC)
husec41_c854	96482	144	382	+	Secretion system	D3GSV7	Type III secretion system protein	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c854	75072	405	691	+	Secretion system	C8UAK5	Type III secretion protein EprJ	Escherichia coli O103:H2 (strain 12009 / EHEC)

husec41_c862	103824	68	379	+	Secretion system	D3GUY9	Putative type VI secretion protein	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c929	75593	1650	3	-	Secretion system	D3GUZ5	Putative type VI secretion protein	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c975	54598	1101	847	-	Secretion system	C8TKY1	Sec-independent protein translocase protein tatA/E homolog 2	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c335	64567	2147	728	-	Secretion system?	C8TVQ3	Predicted transglycosylase	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c553	107942	1760	696	-	Secretion System?	B7L5T3	Putative HlyD family secretion protein	Escherichia coli (strain 55989 / EAEC)
husec41_c902	111517	1527	2473	+	virulence, plasmid	D3H5A9	Virulence protein required for expression/correct membrane localisation of IcsA (VirG)	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c1048	80194	1120	1570	+	fimbria	B7L7K6	Putative major fimbrial subunit FmlA	Escherichia coli (strain 55989 / EAEC)
husec41_c1130	72963	296	18	-	fimbria	B7LAC4	Putative uncharacterized protein ybgQ	Escherichia coli (strain 55989 / EAEC)
husec41_c1130	93863	871	308	-	fimbria	C8TKM4	Putative fimbrial-like protein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c123	77229	71	1304	+	fimbria	C8TI53	Predicted outer membrane protein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1256	104676	3085	2372	-	fimbria	C8THB2	Putative fimbrial protein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1279	80482	1413	964	-	fimbria	C8U6Q7	Putative fimbrial subunit protein	Escherichia coli O103:H2 (strain 12009 / EHEC)

husec41_c128	80195	3	224	+	fimbria	B7L7K6	Putative major fimbrial subunit FmlA	Escherichia coli (strain 55989 / EAEC)
husec41_c128	70269	1367	3965	+	fimbria	C8TQ64	Outer membrane usher protein FimD	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c128	101175	4526	5026	+	fimbria	B7L7K2	Putative fimbrial-like adhesin exported protein	Escherichia coli (strain 55989 / EAEC)
husec41_c128	102666	5088	6004	+	fimbria	B7L7K1	Putative fimbrial-like exported adhesin protein	Escherichia coli (strain 55989 / EAEC)
husec41_c1369	66153	161	1215	+	fimbria	C8TL83	Putative fimbrial protein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1400	93626	588	145	-	fimbria	C8U6Q8	Putative fimbrial subunit protein	Escherichia coli O103:H2 (strain 12009 / EHEC)
husec41_c1400	54315	1084	614	-	fimbria	B7LBJ0	Putative minor fimbrial subunit	Escherichia coli (strain 55989 / EAEC)
husec41_c1503	86233	3	458	+	fimbria	C8THG4	Putative fimbrial protein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1503	94749	601	1080	+	fimbria	C8THG3	Putative fimbrial protein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1527	210	3	356	+	fimbria, plasmid	P46005	Outer membrane usher protein AggC	Escherichia coli
husec41_c1592	66268	103	2506	+	fimbria	C8TMJ0	Putative fimbrial usher protein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1651	97157	260	820	+	fimbria	C8TUY3	Putative fimbrial-like protein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1651	70863	905	1975	+	fimbria	C8TUY2	Outer membrane usher protein	Escherichia coli O26:H11 (strain 11368 / EHEC)

husec41_c1675	70369	65	737	+	fimbria	D3H3A6	Fimbrial outer membrane usher protein	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c1701	209	2	387	+	fimbria, plasmid	P46005	Outer membrane usher protein AggC	Escherichia coli
husec41_c1953	76524	363	710	+	fimbria, plasmid	B7LWW0	14 kDa aggregative adherence fimbriae I protein (Modular protein)	Escherichia coli (strain 55989 / EAEC)
husec41_c202	99238	4697	2216	-	fimbria	C8TU35	Putative outer membrane protein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c202	95193	6013	5474	-	fimbria	B7L9Y3	Putative fimbrial-like adhesin protein	Escherichia coli (strain 55989 / EAEC)
husec41_c2205	205	654	63	-	fimbria, plasmid	P46005	Outer membrane usher protein AggC	Escherichia coli
husec41_c279	77228	1441	2735	+	fimbria	C8TI53	Predicted outer membrane protein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c353	70368	3	677	+	fimbria	D3H3A6	Fimbrial outer membrane usher protein	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c487	108564	2468	2009	-	fimbria	B7L8D9	Putative fimbrial-like adhesin protein	Escherichia coli (strain 55989 / EAEC)
husec41_c498	70862	1	1572	+	fimbria	C8TUY2	Outer membrane usher protein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c507	203	1	656	+	fimbria, plasmid	P46005	Outer membrane usher protein AggC	Escherichia coli
husec41_c514	89862	842	2148	+	fimbria	C8THG6	Probable outer membrane porin protein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c636	62099	1846	1448	-	fimbria	C8TMJ2	Putative fimbrial adhesin protein	Escherichia coli O26:H11 (strain 11368 / EHEC)

husec41_c636	80512	2907	1855	-	fimbria	C8TMJ1	Putative minor fimbrial subunit	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c636	88197	5440	3692	-	fimbria	C8TM56	Putative outer membrane usher protein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c636	85971	6768	6247	-	fimbria	C8TMI8	Putative fimbrial major protein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c679	92542	568	2	-	fimbria	C8THG8	Putative fimbrial-like protein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c775	202	864	1	-	fimbria, plasmid	P46005	Outer membrane usher protein AggC	Escherichia coli
husec41_c777	86234	236	3	-	fimbria	C8THG4	Putative fimbrial protein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c777	96182	820	254	-	fimbria	C8U1K8	Putative fimbrial protein	Escherichia coli O103:H2 (strain 12009 / EHEC)
husec41_c777	89861	2216	840	-	fimbria	C8THG6	Probable outer membrane porin protein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c92	70271	3747	1460	-	fimbria	C8TQ64	Outer membrane usher protein FimD	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c972	51908	1727	1158	-	Fimbria	Q8Z2Q1	Probable fimbrial subunit protein	Salmonella typhi
husec41_c98	85976	662	3129	+	fimbria	C8TH67	Predicted outer membrane usher protein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_rep_c2402	74160	2	178	+	fimbria	D3GT73	Putative fimbrial protein	Escherichia coli O44:H18 (strain 042 / EAEC)



**Table 6: Mercuric resistance plasmid**

Contig ID	Gen ID	start	end	S	tags	Similar to	Protein names	Organism
husec41_c784	92895	725	294	-	Mercuric resistance	C8UQM8	Mercuric resistance operon regulatory protein MerR	Escherichia coli O111:H- (strain 11128 / EHEC)
husec41_c784	101266	797	1144	+	Mercuric resistance, plasmid	C8UQM9	Mercuric ion transport protein MerT	Escherichia coli O111:H- (strain 11128 / EHEC)
husec41_c784	92617	1161	1433	+	Mercuric resistance	D3H375	Mercuric ion transport protein	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c784	8662	1433	1891	+	Mercuric resistance, plasmid	Q0ZKU6	Mercuric resistance protein MerC (Mercury resistance operon transport protein MerC) (Putative uncharacterized protein)	Escherichia coli
husec41_c784	109956	1946	3639	+	Mercuric	Q935L3	Putative mercuric reductase (EC 1.16.1.1)	Salmonella typhi
husec41_c784	112267	3660	4019	+	Mercuric resistance	D3H372	MerR-family transcriptional regulator	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c784	23874	4075	4953	+	plasmid	Q5J458	Urf2	Salmonella choleraesuis

The predicted proteins included in this table are all located in the same contig and probably are in a plasmid forming a functional operon with a MerR family regulator in a divergent orientation to the rest of the components of the operon.

**Table 7: Tellurium resistance**

Contig ID	Gen ID	start	end	S	tags	Similar to	Protein names	Organism
husec41_c1152	96081	28	767	+	Tellurium	C8TNH6	Putative tellurium resistance protein TerC	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1243	99769	532	1110	+	Tellurium	C8TNH3	Putative tellurium resistance protein TerZ	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1243	73496	1113	2266	+	Tellurium	C8TNH4	Putative tellurium resistance protein TerA	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1243	66160	2292	2741	+	Tellurium	C8TNH5	Putative tellurium resistance protein TerB	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1512	103519	1566	327	-	Tellurium	C8TX21	Putative tellurium resistance protein TerF	Escherichia coli O103:H2 (strain 12009 / EHEC)
husec41_c1713	96082	1	390	+	Tellurium	C8TNH6	Putative tellurium resistance protein TerC	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1713	82503	442	851	+	Tellurium	C8TNH7	Putative tellurium resistance protein TerD	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c273	26702	193	1230	+	Tellurium resistance	Q19NM8	TerY3	Escherichia coli O1:K1 / APEC
husec41_c2960	27515	6	444	+	Tellurium resistance, plasmid	Q19NL2	TerE	Escherichia coli O1:K1 / APEC
husec41_c750	24372	541	5	-	Tellurium resistance	Q19NM7	TerY2	Escherichia coli O1:K1 / APEC
husec41_c750	5509	1149	541	-	Tellurium resistance, plasmid	P75012	Tellurium resistance protein TerX	Serratia marcescens
husec41_c750	23668	1919	1131	-	Tellurium resistance	Q19NM5	TerY1	Escherichia coli O1:K1 / APEC
husec41_c750	5508	2771	2306	-	Tellurium resistance, plasmid	P75010	Tellurium resistance protein terW	Serratia marcescens
husec41_c794	111062	423	127	-	Tellurium	C8TNH8	Putative tellurium resistance protein TerE	Escherichia coli O26:H11 (strain 11368 / EHEC)

In addition to genes involved in Mercuric resistance and Tellurium resistance we have predicted and annotated in this genome many genes involved in resistance to other metals (See complete annotation)

**Table 8: Transposases**

Contig ID	Gen ID	start	end	S	tags	Similar to	Protein names	Organism
husec41_c1004	111932	3	1581	+	Transposase	D3H358	Transposase	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c1066	6571	2695	1796	-	Transposase	P30192	Putative uncharacterized protein ychG	Escherichia coli (strain K12)
husec41_c1147	110427	1338	1068	-	Transposase	B7L8U5	Putative transposase ORF A, IS609 family	Escherichia coli (strain 55989 / EAEC)
husec41_c1214	21769	66	377	+	Transposase	C2DME2	Possible transposase insP for IS630	Escherichia coli 83972
husec41_c130	71084	1	1366	+	Transposase	B7L5X1	Tn7-like transposition protein TnsB	Escherichia coli (strain 55989 / EAEC)
husec41_c130	63306	1369	2664	+	Transposase	B7L5W7	Tn7-like transposition protein TnsC	Escherichia coli (strain 55989 / EAEC)
husec41_c1390	13776	3	491	+	Transposase	E9Z073	Transposase	Escherichia coli M863
husec41_c1412	25246	1462	2	-	Transposase	C8TNV4	Predicted transposase	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1539	24646	713	237	-	Transposase	Q8XC14	Putative IS encoded protein encoded within prophage CP-9330	Escherichia coli O157:H7
husec41_c1554	7744	908	1933	+	Transposase	P16943	Insertion element IS630 uncharacterized 39 kDa protein (ISO-IS200 39 kDa protein)	Shigella sonnei
husec41_c1562	102568	214	855	+	Transposase, plasmid	B7LWU0	Transposon Tn3 resolvase	Escherichia coli (strain 55989 / EAEC)

husec41_c1562	8933	296	3	-	Transposase, plasmid	C7S9E2	Transposase for transposon Tn3 (Truncated TnpA)	Escherichia coli
husec41_c1569	16194	2	298	+	Transposase	B3HPU8	Transposase OrfA, ISEc8	Escherichia coli F11
husec41_c1595	21687	515	29	-	Transposase	Q6Q6S9	Transposase	Escherichia coli
husec41_c1628	110297	589	38	-	Transposase	B7LBW1	Transposase	Escherichia coli (strain 55989 / EAEC)
husec41_c1683	106988	233	526	+	Transposase	B7LHE6	Putative uncharacterized protein yncI	Escherichia coli (strain 55989 / EAEC)
husec41_c1687	21743	1866	1304	-	Transposase	B3HAL5	Transposon Tn21 resolvase	Escherichia coli B7A
husec41_c1689	5585	1187	3	-	Transposase	P03008	Transposase for transposon Tn3	Escherichia coli
husec41_c1781	111704	100	405	+	Transposase	B7LFE0	IS3 element protein InsE	Escherichia coli (strain 55989 / EAEC)
husec41_c1781	51670	405	1269	+	Transposase, plasmid	B7LA99	IS3 element protein InsF (Transposase ORF B, IS3)	Escherichia coli (strain 55989 / EAEC)
husec41_c1799	51569	317	2	-	Transposase	C8TKW2	Putative uncharacterized protein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_c1859	25841	3	428	+	Transposase	B6IAL2	H repeat-associated protein	Escherichia coli (strain SE11)
husec41_c1864	20072	3	582	+	Transposase, plasmid	C9WXH4	Transposase (Fragment)	Escherichia coli
husec41_c1864	108791	742	542	-	Transposase, plasmid	Q935I2	Putative transposase	Salmonella typhi
husec41_c188	66659	5009	5828	+	Transposase	D3GR79	Putative transposase	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c188	66660	5768	5911	+	Transposase	D3GR79	Putative transposase	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c1905	83869	564	244	-	Transposase, plasmid	B7LWW4	Transposase ORF A, IS1	Escherichia coli (strain 55989 / EAEC)

husec41_c1921	111933	1121	217	-	Transposase	D3H358	Transposase	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c1974	67429	454	699	+	Transposase	B7LWV0	Transposase ORF A, IS629	Escherichia coli (strain 55989 / EAEC)
husec41_c1974	76522	845	1078	+	Transposase, plasmid	B7LWV1	Putative uncharacterized protein	Escherichia coli (strain 55989 / EAEC)
husec41_c199	5586	1	870	+	Transposase	P03008	Transposase for transposon Tn3	Escherichia coli
husec41_c1995	69269	349	1	-	Transposase	B7LFN9	Transposase ORF A, IS3 family	Escherichia coli (strain 55989 / EAEC)
husec41_c1998	55463	3	224	+	Transposase	D3GV20	Transposase (Fragment)	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c2024	12883	504	1	-	Transposase	B5AXD2	Putative transposase	Escherichia coli
husec41_c2049	21688	188	625	+	Transposase	Q6Q6S9	Transposase	Escherichia coli
husec41_c2055	22558	320	3	-	Transposase	A8A1D5	Transposase, IS605 family	Escherichia coli O9:H4 (strain HS)
husec41_c2063	69134	2	481	+	Transposase	B7L8E6	Putative transposase	Escherichia coli (strain 55989 / EAEC)
husec41_c2064	21598	322	3	-	Transposase	E7H9R9	Transposase IS66 family protein	Escherichia coli EPECa14
husec41_c2066	71085	3	332	+	Transposase	B7L5X1	Tn7-like transposition protein TnsB	Escherichia coli (strain 55989 / EAEC)
husec41_c211	78103	3309	2437	-	Transposase	B7L4U1	Putative transposase	Escherichia coli (strain 55989 / EAEC)
husec41_c2125	53014	610	1044	+	Transposase, plasmid	B7LA98	Putative uncharacterized protein	Escherichia coli (strain 55989 / EAEC)
husec41_c2125	88822	1044	1392	+	Transposase	D3GV17	Transposase	Escherichia coli O44:H18 (strain 042 / EAEC)

husec41_c2125	101185	4237	3107	-	Transposase, plasmid	B7LX19	Putative transposase, IS110 family	Escherichia coli (strain 55989 / EAEC)
husec41_c2132	91471	296	866	+	Transposase	B7LFP9	Transposase ORF B, IS629	Escherichia coli (strain 55989 / EAEC)
husec41_c2176	13522	341	839	+	Transposase	E9Z008	Transposase	Escherichia coli M863
husec41_c2181	91472	2	478	+	Transposase	B7LFP9	Transposase ORF B, IS629	Escherichia coli (strain 55989 / EAEC)
husec41_c2203	35831	2	262	+	Transposase, plasmid	F2W481	ISL3 family transposase	Escherichia coli
husec41_c226	22296	2352	2906	+	Transposase, plasmid	B7UHC5	Transposase of ISEc13 of IS110 family (Transposase of ISEc21)	Escherichia coli O127:H6 (strain E2348/69 / EPEC)
husec41_c226	17778	3420	2999	-	Transposase	B3IA39	Truncated transposase	Escherichia coli E22
husec41_c230	104559	2241	3141	+	Transposase	B7LBF9	Putative transposase	Escherichia coli (strain 55989 / EAEC)
husec41_c273	9423	4484	4966	+	Transposase	C4HU49	Transposase	Yersinia pestis Pestoides A
husec41_c30	5587	692	3	-	Transposase	P03008	Transposase for transposon Tn3	Escherichia coli
husec41_c30	8689	3454	2163	-	Transposase, plasmid	D9Z5B6	TnpA	Escherichia coli
husec41_c37	14165	683	282	-	Transposase	D7ZCL8	Transposase, IS4 family (Fragment)	Escherichia coli MS 69-1
husec41_c42	25291	3	656	+	Transposase	A7ZMR8	Transposase, IS605 family	Escherichia coli O139:H28 (strain E24377A / ETEC)
husec41_c524	14380	1798	2407	+	Transposase, plasmid	C7S9T2	Transposon Tn21 modulator protein	Escherichia coli
husec41_c524	90514	2536	3093	+	Transposase	D3H359	Transposon Tn21 resolvase	Escherichia coli O44:H18 (strain 042 / EAEC)

husec41_c524	111934	3099	3623	+	Transposase	D3H358	Transposase	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c526	74046	84	436	+	Transposase	D3GX03	Insertion sequence IS100, ATP-binding protein	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c526	32500	1588	2044	+	Transposase	P59697	Transposase for insertion sequence element IS200	Salmonella typhi
husec41_c6	82187	2934	2260	-	Transposase, plasmid	D3H553	Transposase	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_c642	10032	1	264	+	Transposase	B3WUE9	TniA transposase protein	Escherichia coli B171
husec41_c66	13681	1012	222	-	Transposase, plasmid	Q0H058	IS1N transposase	Escherichia coli
husec41_c662	24374	5	1458	+	Transposase	Q19NI3	TnsD	Escherichia coli O1:K1 / APEC
husec41_c699	106986	2	535	+	Transposase?	B7LHE6	Putative uncharacterized protein yncI	Escherichia coli (strain 55989 / EAEC)
husec41_c826	17526	1060	1440	+	Transposase	D8AYZ0	Transposase	Escherichia coli MS 175-1
husec41_c826	73020	1440	1784	+	Transposase	B7L939	Putative transposase ORF 2, IS66 family	Escherichia coli (strain 55989 / EAEC)
husec41_c835	76478	3	323	+	Transposase	B7LDR0	Transposase ORF B, IS1	Escherichia coli (strain 55989 / EAEC)
husec41_c838	11493	2	500	+	Transposase	E7SID2	Transposase	Shigella dysenteriae CDC 74-1112
husec41_c87	100390	1003	161	-	Transposase	B7L5X0	Tn7-like transposase TnsA	Escherichia coli (strain 55989 / EAEC)
husec41_c885	13775	46	558	+	Transposase	E9Z073	Transposase	Escherichia coli M863
husec41_c887	55149	1100	357	-	Transposase	C8TWQ4	Predicted IS602 transposase OrfB	Escherichia coli O103:H2 (strain 12009 / EHEC)

husec41_rep_c2269	51839	1109	132	-	Transposase	C8TGN1	Putative IS621 transposase	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_rep_c2271	5427	66	1175	+	Transposase	P11901	Transposase for insertion sequence element IS421	Escherichia coli
husec41_rep_c2273	13754	784	90	-	Transposase	E1U309	Transposase InsAB'	Escherichia coli
husec41_rep_c2275	48940	760	68	-	Transposase, plasmid	Q7AQT7	Putative transposase	Salmonella typhi
husec41_rep_c2276	59489	43	819	+	Transposase	B7LDT8	Putative uncharacterized protein	Escherichia coli (strain 55989 / EAEC)
husec41_rep_c2286	69133	49	592	+	Transposase	B7L8E6	Putative transposase	Escherichia coli (strain 55989 / EAEC)
husec41_rep_c2306	14601	85	282	+	Transposase	E9YLK7	Transposase (Fragment)	Escherichia coli TA007
husec41_rep_c2315	14599	351	1	-	Transposase	E9YLK7	Transposase (Fragment)	Escherichia coli TA007
husec41_rep_c2317	24645	1	495	+	Transposase	Q8XC14	Putative IS encoded protein encoded within prophage CP-9330	Escherichia coli O157:H7
husec41_rep_c2317	52995	548	1225	+	Transposase	C6UPH5	Transposase ISEc8 (Transposase, ISEc8)	Escherichia coli O157:H7 (strain TW14359 / EHEC)
husec41_rep_c2321	81376	3	362	+	Transposase	C8UCH9	Putative uncharacterized protein	Escherichia coli O111:H- (strain 11128 / EHEC)
husec41_rep_c2323	73469	311	15	-	Transposase	C8TND9	Putative uncharacterized protein	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_rep_c2336	22560	201	1	-	Transposase	A8A1D5	Transposase, IS605 family	Escherichia coli O9:H4 (strain HS)
husec41_rep_c2336	55285	239	1385	+	Transposase	C8TP26	Putative IS609 transposase TnpB	Escherichia coli O26:H11 (strain 11368 / EHEC)
husec41_rep_c2356	78317	3	632	+	Transposase	B7L8Y0	IS30 transposase; KpLE2 phage-like element	Escherichia coli (strain 55989 / EAEC)



husec41_rep_c2370	12154	446	78	-	Transposase	Q8VRB2	Putative transposase	Escherichia coli
husec41_rep_c2372	63307	487	248	-	Transposase	B7L5W7	Tn7-like transposition protein TnsC	Escherichia coli (strain 55989 / EAEC)
husec41_rep_c2408	20782	183	1	-	Transposase	E8IR83	IS66 transposase	Escherichia coli O55:H7 str. USDA 5905
husec41_rep_c2429	5589	3	323	+	Transposase	P03008	Transposase for transposon Tn3	Escherichia coli
husec41_rep_c2437	83870	89	268	+	Transposase, plasmid	B7LWW4	Transposase ORF A, IS1	Escherichia coli (strain 55989 / EAEC)
husec41_rep_c2441	88832	46	481	+	Transposase	D3GV27	Transposase	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_rep_c2445	76479	4	246	+	Transposase	B7LDR0	Transposase ORF B, IS1	Escherichia coli (strain 55989 / EAEC)
husec41_rep_c2446	76523	233	8	-	Transposase, plasmid	B7LWV1	Putative uncharacterized protein	Escherichia coli (strain 55989 / EAEC)
husec41_rep_c2449	56116	509	3	-	Transposase	D3GV14	Transposase	Escherichia coli O44:H18 (strain 042 / EAEC)
husec41_rep_c2478	9000	328	38	-	Transposase	E7IJ32	Putative transposase	Escherichia coli 1180
husec41_rep_c2487	13756	271	3	-	Transposase	E1U309	Transposase InsAB'	Escherichia coli
husec41_rep_c2498	91473	309	3	-	Transposase	B7LFP9	Transposase ORF B, IS629	Escherichia coli (strain 55989 / EAEC)
husec41_rep_c2499	15960	1	468	+	Transposase	E3Y631	ISL3 family transposase domain protein	Shigella flexneri 2a str. 2457T
husec41_rep_c2515	6352	246	58	-	Transposase	Q2EEQ8	Putative defective transposase ybfQ	Escherichia coli (strain K12)
husec41_rep_c2516	14290	16	332	+	Transposase	E3XVP6	Putative transposase	Escherichia coli 2362-75
husec41_rep_c2520	26207	451	35	-	Transposase	B6ICP6	Truncated transposase	Escherichia coli (strain SE11)

husec41_rep_c2540	2737	1	222	+	Transposase	P76102	Putative transposase InsQ for insertion sequence element IS609	Escherichia coli (strain K12)
husec41_rep_c2544	17058	19	291	+	Transposase	D6HU78	Putative uncharacterized protein	Escherichia coli B088
husec41_rep_c2565	83426	5	472	+	Transposase	B7L6S2	Putative transposase	Escherichia coli (strain 55989 / EAEC)
husec41_rep_c2570	17060	1	246	+	Transposase	D6HU78	Putative uncharacterized protein	Escherichia coli B088
husec41_rep_c2572	108075	3	407	+	Transposase	C8ULY2	Putative IS609 transposase TnpA	Escherichia coli O111:H- (strain 11128 / EHEC)
husec41_rep_c2577	13700	429	172	-	Transposase	Q4E6D2	Transposase (IS4 family)	Wolbachia endosymbiont of Drosophila simulans
husec41_rep_c2578	21434	279	1	-	Transposase	E3Y0F9	Transposase, IS605 OrfB family	Shigella flexneri 2a str. 2457T
husec41_rep_c2639	100431	288	3	-	Transposase	B7LBV5	Transposase, IS110 family	Escherichia coli (strain 55989 / EAEC)
husec41_rep_c2668	17698	206	3	-	Transposase	E9YA15	Transposase	Escherichia coli TA007
husec41_rep_c2679	27205	233	2	-	Transposase	E8Y4A6	Transposase IS200-family protein	Escherichia coli (strain ATCC 55124 / KO11)
husec41_rep_c2743	79963	191	3	-	Transposase, plasmid	B5YPI1	ISSd1, transposase OrfB	Escherichia coli O157:H7 (strain EC4115 / EHEC)
husec41_rep_c2812	83428	290	2	-	Transposase	B7L6S2	Putative transposase	Escherichia coli (strain 55989 / EAEC)
husec41_rep_c2873	12156	200	3	-	Transposase	Q8VRB2	Putative transposase	Escherichia coli
husec41_rep_c2875	2698	934	2	-	Transposase	P0CE57	Transposase insH for insertion sequence element IS5R	Escherichia coli (strain K12)
husec41_rep_c2889	20091	181	2	-	Transposase	D8B2Q9	ISPsy10, transposase family protein (Fragment)	Escherichia coli MS 175-1

husec41_rep_c2890	22348	3	245	+	Transposase	B1LJZ9	IS5 transposase	Escherichia coli (strain SMS-3-5 / SECEC)
husec41_rep_c2903	17062	187	2	-	Transposase	D6HU78	Putative uncharacterized protein	Escherichia coli B088
husec41_rep_c2907	83429	3	269	+	Transposase	B7L6S2	Putative transposase	Escherichia coli (strain 55989 / EAEC)

It seems that there are 121 putative transposases in this genome. It probably implies a high genomic plasticity and flexibility for adaptation to changing environments. In the next table we have collected some comments about these set of transposases of *E. coli* EHEC genome.

UNIPROT	Annotation	COMMENTS	FAMILY
D3H358	Transposase	Transposase_7. Segun interpro	Tn3
P30192	Putative uncharacterized protein ychG	Caution, it can be a pseudogene	
B7L8U5	Putative transposase ORF A, IS609 family	Putative	IS609
C2DME2	Possible transposase insP for IS630	Length 115 aa	IS630
E9Z073	Transposase	REP-associated transposases (Pubmed: 16563168)	IS110
C8TNV4	Predicted transposase		ISL3
B7LWU0	Transposon Tn3 resolvase	On plasmid, in some cases is associated to mercury resistance	Tn3
C7S9E2	Transposase for transposon Tn3 (Truncated TnpA)	125 aa on plasmid ?	Tn3
B3HPU8	Transposase OrfA, ISEc8	120 aa	ISEc8
Q6Q6S9	Transposase	Molecular characterization of ceftioxin-resistant Escherichia coli from Canadian hospitals	IS911
B7LBW1	Transposase	Pubmed 19165319	
B7LHE6	Putative uncharacterized protein yncI	REP-associated transposases (Pubmed: 16563168)	IS4
P03008	Transposase for transposon Tn3		Tn3
B7LFE0	IS3 element protein InsE	Pubmed 19165319	IS3

B7LA99	IS3 element protein InsF (Transposase ORF B, IS3)		IS3
C8TKW2	Putative uncharacterized protein		IS66
B6IAL2	H repeat-associated protein	REP-associated transposases (Pubmed: 16563168)	IS4
C9WXH4	Transposase (Fragment)		Tni?
Q935I2	Putative transposase		IS26
D3GR79	Putative transposase		
D3GR79	Putative transposase		
B7LWW4	Transposase ORF A, IS1	REP-associated transposases (Pubmed: 16563168)	IS110
D3H358	Transposase	On plasmid, in some cases is associated to mercury resistance	Tn3
B7LWV0	Transposase ORF A, IS629		IS629
B7LWV1	Putative uncharacterized protein		
P03008	Transposase for transposon Tn3	On plasmid, in some cases is associated to mercury resistance	Tn3
B7LFN9	Transposase ORF A, IS3 family		IS911
D3GV20	Transposase (Fragment)		Mutator
B5AXD2	Putative transposase		IS30
Q6Q6S9	Transposase		IS911
A8A1D5	Transposase, IS605 family		IS609
B7L8E6	Putative transposase		ISEc3
E7H9R9	Transposase IS66 family protein		IS66
B7L5X1	Tn7-like transposition protein TnsB		Tn7
B7L4U1	Putative transposase		YhgA
B7LA98	Putative uncharacterized protein		
D3GV17	Transposase		IS66
B7LX19	Putative transposase, IS110 family	IS116/IS110/IS902 family REP-associated transposases (Pubmed: 16563168)	IS110

B7LFP9	Transposase ORF B, IS629		IS629
E9Z008	Transposase		
B7LFP9	Transposase ORF B, IS629		IS629
F2W481	ISL3 family transposase		ISL3
B7UHC5	Transposase of ISEc13 of IS110 family (Transposase of ISEc21)	IS116/IS110/IS902 family REP-associated transposases (Pubmed: 16563168)	IS110
B3IA39	Truncated transposase		
B7LBF9	Putative transposase		YhgA
C4HU49	Transposase		
P03008	Transposase for transposon Tn3	On plasmid, in some cases is associated to mercury resistance	Tn3
D9Z5B6	TnpA		ISEcp1 TnpA
D7ZCL8	Transposase, IS4 family (Fragment)	REP-associated transposases (Pubmed: 16563168)	IS4
A7ZMR8	Transposase, IS605 family		IS605
D3H358	Transposase	In some cases is associated to mercury resistance	Tn3
P59697	Transposase for insertion sequence element IS200	TpnA2 related	IS200
D3H553	Transposase		IS66
B3WUE9	TniA transposase protein		
Q0H058	IS1N transposase		ISN1
B7LHE6	Putative uncharacterized protein yncI	REP-associated transposases (Pubmed: 16563168)	IS4
D8AYZ0	Transposase		
B7L939	Putative transposase ORF 2, IS66 family		IS66
B7LDR0	Transposase ORF B, IS1		IS1
E7SID2	Transposase		
B7L5X0	Tn7-like transposase TnsA		Tn7-like
E9Z073	Transposase	REP-associated transposases (Pubmed: 16563168)	IS110

C8TWQ4	Predicted IS602 transposase OrfB		IS911
C8TGN1	Putative IS621 transposase	IS116/IS110/IS902 family	IS621
P11901	Transposase for insertion sequence element IS421	IS4 related	IS186
E1U309	Transposase InsAB'		IS1
Q7AQT7	Putative transposase	IS1216 related	TnpA - IS26
B7LDT8	Putative uncharacterized protein	ISSd1 related	IS600
B7L8E6	Putative transposase		ISEc3
E9YLK7	Transposase (Fragment)		
E9YLK7	Transposase (Fragment)		
C6UPH5	Transposase ISEc8 (Transposase, ISEc8)		IS66
C8UCH9	Putative uncharacterized protein	IS682 related	IS66
C8TND9	Putative uncharacterized protein		ISEc8 ??
A8A1D5	Transposase, IS605 family		IS609
C8TP26	Putative IS609 transposase TnpB		IS605
B7L8Y0	IS30 transposase; KpLE2 phage-like element		IS30
Q8VRB2	Putative transposase		
E8IR83	IS66 transposase		IS66
P03008	Transposase for transposon Tn3		Tn3 - TnpA
B7LWW4	Transposase ORF A, IS1		IS1
D3GV27	Transposase		IS30
B7LDR0	Transposase ORF B, IS1		IS1
B7LWV1	Putative uncharacterized protein		
D3GV14	Transposase	IS116/IS110/IS902 family	IS110
E7IJ32	Putative transposase		
E1U309	Transposase InsAB'		IS1

B7LFP9	Transposase ORF B, IS629		IS629
E3Y631	ISL3 family transposase domain protein		ISL3
Q2EEQ8	Putative defective transposase ybfQ		ISEc2
E3XVP6	Putative transposase		
B6ICP6	Truncated transposase		
P76102	Putative transposase InsQ for insertion sequence element IS609		IS605
D6HU78	Putative uncharacterized protein		
B7L6S2	Putative transposase		IS605
D6HU78	Putative uncharacterized protein		
C8ULY2	Putative IS609 transposase TnpA	IS 200 related	IS609
Q4E6D2	Transposase (IS4 family)	InsH - IS4 related REP-associated transposases (Pubmed: 16563168)	IS5
E3Y0F9	Transposase, IS605 OrfB family		IS605
B7LBV5	Transposase, IS110 family	IS116/IS110/IS902	IS110
E9YA15	Transposase		IS1 1/5/6
E8Y4A6	Transposase IS200-family protein		IS200
B5YPI1	ISSd1, transposase OrfB		ISSd1
B7L6S2	Putative transposase	IS 609 related	IS605
Q8VRB2	Putative transposase		
P0CE57	Transposase insH for insertion sequence element IS5R	REP-associated transposases (Pubmed: 16563168)	IS5
D8B2Q9	ISPsy10, transposase family protein (Fragment)		ISpsy10
B1LJZ9	IS5 transposase	REP-associated transposases (Pubmed: 16563168)	IS5
D6HU78	Putative uncharacterized protein		
B7L6S2	Putative transposase	IS 609 related	IS605

## **Plasmids**

Around 246 predicted proteins appear to be related with plasmids. The tagged genome annotation table that we are updating continuously is available at:

[http://www.era7bioinformatics.com/docs/EHEC\\_E\\_COLI\\_GERMANY\\_OUTBREAK\\_Annotation\\_Era7\\_Bioinformatics\\_v1\\_5\\_6\\_2011.xls](http://www.era7bioinformatics.com/docs/EHEC_E_COLI_GERMANY_OUTBREAK_Annotation_Era7_Bioinformatics_v1_5_6_2011.xls)

This strain has many genetic capabilities that probably confer it a competitive advantage. Some important features that this strain bears in its genome:

- A restriction-modification system
- Many proteins involved in Fe transport and utilization. Siderophores: aerobactin, enterobactin.
- Lysozyme
- A general inhibitor of pancreatic serine proteases: inhibits chymotrypsin, trypsin, elastases, factor X, kallikrein as well as a variety of other proteases
- Proteins involved in anaerobic respiration
- Antimicrobial peptides
- Proteins involved in quorum-sensing and biofilm formation
- Proteins involved in Ni, Cu, Zn and Co resistance
- More than 170 phage proteins

**Annotation in Genbank format can be obtained at:**

[http://www.era7bioinformatics.com/docs/EHEC\\_E\\_COLI\\_GERMANY\\_OUTBREAK\\_Annotation\\_Era7\\_Bioinformatics\\_v1\\_5\\_6\\_2011.pdf](http://www.era7bioinformatics.com/docs/EHEC_E_COLI_GERMANY_OUTBREAK_Annotation_Era7_Bioinformatics_v1_5_6_2011.pdf)