

## Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Rohde H, Qin J, Cui Y, et al. Open-source genomic analysis of Shiga-toxin-producing *E. coli* O104:H4. N Engl J Med 2011;365:718-24. DOI: 10.1056/NEJMoa1107643.

## **Supplementary Appendix: Open-source genomics of a Shiga-toxin-producing *Escherichia coli* O104:H4**

### **Crowd-sourcing consortium**

The following members of the *E. coli* O104:H4 Genome Analysis Crowd-sourcing consortium made contributions that influenced the analyses reported here: Kathryn E. Holt, David J. Studholme, Michael Feldgarden and Marina Manrique.

A full account of crowd-sourcing efforts can be accessed here: <https://github.com/ehc-outbreak-crowdsourced/BGI-data-analysis/wiki/>

### **Methods and Results**

#### *Ion Torrent library construction and sequencing*

Genomic DNA was extracted and purified using a conventional SDS lysis and phenol-chloroform method. 5µg of DNA (OD260/OD280 = 1.85) was dissolved in TE buffer to a total volume of 100 µl and fragmented by sonication (Covaris S2, Massachusetts, USA) to a size distribution of 50-300 bp. Library preparation and template preparation of live Ion Sphere™ Particles was performed according to the manufacturer's protocol (Ion Torrent, USA). During the library preparation, nick-translation was followed by 5 cycles of PCR amplification. Finally the sequencing was performed on the PGM Sequencer. Seven 314 chips were run to generate 79.1 Mb of sequence, with average length of 101 bp.

#### *Illumina library construction and sequencing*

Whole-genome sequencing was performed using Illumina HiSeq 2000 (Illumina Inc. U.S.A) by generating paired-end libraries with an average insert size of 470 bp, 2 kbp and 6 kbp

following the manufacturer's instruction. The read lengths were 90bp, 50bp, 50bp and 1Gb, 576Mb and 576Mb high quality data were generated from each library respectively.

*Creation of a draft genome assembly using Ion Torrent PGM data (2nd June 2011)*

An assembly was performed using MIRA 3.2.1.17\_dev using command-line parameters --*job=denovo,genome,accurate,iontor -GE:not=1*. The Ion Torrent PGM assembly from 5 chips of Ion Torrent 314 data produced an assembly of 3,057 contigs, total bases: 5,491,032 with an N50 value of 3,675.

*Creation of a hybrid assembly using Ion Torrent PGM data and Illumina single-end data (6th June 2011)*

Ion Torrent and Illumina read data were quality filtered before assembly including removal of adapter contamination. The Ion Torrent PGM assembly from 7 chips of Ion Torrent 314 data were assembled with Newbler 2.0.00.22. Illumina single-end data (taken from the in-progress paired-end run) were assembled using SOAPdenovo 1.06<sup>1</sup> (with *k*-mer of 51 and parameters "-d 1, -R". Assemblies were combined using AMOS minimus2 1.59 with parameters REFCOUNT=0, OVERLAP=50, MINID=94, MAXTRIM=10<sup>2</sup>. The resulting assembly consisted of 451 contigs greater than 200bp with an N50 of 53266bp. The largest contig was 204342bp.

*Creation of a draft genome scaffold assembly using Illumina paired-end and mate-pair reads*

A draft *de novo* assembly was produced using SOAPdenovo version 1.05. Contigs were first assembled using the 470bp paired-end library initially using a *k*-mer value of 45 for de Bruijn graph construction. These were subsequently scaffolded in a hierarchical fashion using 2kb followed by 6kb mate-pair libraries by way of the rank parameter in the SOAPdenovo configuration file. Other parameters supplied to SOAPdenovo included -F to attempt to fill

gaps in scaffolds. Where possible, in order to fill remaining scaffold gaps, local information available from the abundant mate-pair data was utilised by the GapCloser utility which was run over the assembly output with a *k*-mer size of 23. Both scaffolds and un-scaffolded contigs were used in further analysis, with the exception of contigs smaller than 200bp which were excluded.

*De novo* assembly produced 24 scaffolds plus 75 un-scaffolded contigs. The largest scaffold was 757969bp, the smallest was 552bp. Scaffold N50 was 403980bp. After gap filling the scaffolds contained 143 distinct stretches of gaps (represented as ambiguous 'N' bases) comprising 94491bp of sequence.

#### *Insert sizes*

The estimated insert size with standard deviations predicted by SOAPdenovo are demonstrated in Table S1.

**Table S1. The estimated insert size determined by the *de novo* assembly process.**

Library	Estimated insert size	Standard deviation
470bp	468	31
2kb	2548	246
6kb	6193	566

#### *Determination of closest reference by average nucleotide identity (ANI)*

Average nucleotide identity with all complete *E. coli* genomes available in GenBank was calculated using the ANIb algorithm which uses BLAST as the underlying alignment method<sup>3-4</sup>. Scrutiny of results (Table S2) revealed that *E. coli* 55989 showed the highest nucleotide identity with an ANI of 99.8% between the TY2482 draft chromosome and *E. coli*

55989. The ANIb algorithm shreds sequences into 1kb segments. BLAST alignments needed to be longer than 700bp and have >70% nucleotide identity to count towards ANIb calculation. ANIb parameters to BLAST were "-F F -e 0.001 -v 1 -b 1 -X 150 -q -1".

**Table S2. Average nucleotide identities for TY2482 compared against all complete *E. coli* genomes**

<b>TY2482 vs</b>	<b>ANIb</b>
Escherichia coli 55989	99.84
Escherichia coli IAI1	99.2
Escherichia coli W	99.14
Escherichia coli E24377A	99.09
Escherichia coli SE11	99.09
Escherichia coli O103:H2 str. 12009	98.95
Escherichia coli O26:H11 str. 11368	98.98
Escherichia coli O111:H- str. 11128	98.85
Escherichia coli HS	98.67
Escherichia coli ATCC 8739	98.55
Escherichia coli str. K-12 substr. W3110	98.54
Escherichia coli str. K-12 substr. MG1655	98.54
Escherichia coli DH1	98.54
Escherichia coli BL21-Gold(DE3)plyss AG	98.53
Escherichia coli BL21(DE3)	98.53
Escherichia coli BL21(DE3)	98.53
Escherichia coli B str. REL606	98.53
Escherichia coli BW2952	98.49
Escherichia coli str. K-12 substr. DH10B	98.5
Escherichia coli H10407	98.5
Escherichia coli ETEC H10407	98.5
Escherichia coli O55:H7 str. CB9615	97.92
Escherichia coli O157:H7 str. TW14359	97.86
Escherichia coli O157:H7 str. Sakai	97.87
Escherichia coli O157:H7 str. EC4115	97.86
Escherichia coli O157:H7 str. EDL933	97.82
Escherichia coli 042	97.45
Escherichia coli UMN026	97.39
Escherichia coli IAI39	97.3
Escherichia coli SMS-3-5	97.21
Escherichia coli SE15	97.06
Escherichia coli CFT073	97.02
Escherichia coli S88	96.97

Escherichia coli O83:H1 str. NRG 857C	96.99
Escherichia coli O127:H6 str. E2348/69	96.95
Escherichia coli UM146	96.94
Escherichia coli 536	96.95
Escherichia coli UTI89	96.93
Escherichia coli APEC O1	96.98
Escherichia coli ED1a	96.82

#### *Annotation of putative regions of difference between TY2482 and 59989*

The TY2482 scaffold assembly was aligned against *E. coli* 55989 using progressiveMauve<sup>5</sup> (part of Mauve 2.3.1) using default settings. For ease of viewing, scaffolds were moved and where necessary reverse complemented to fit the order of the *E. coli* 55989 chromosome using the Mauve contig mover, again run with default parameters. Unaligned regions of the TY2482  $\geq 5$ kb were examined as putative regions of difference. Gene prediction within these regions was performed using Glimmer 3.02<sup>6</sup> using the g3-iterated.sh workflow with default options. Genes with a raw score of  $\geq 1.0$  were extracted for further analysis. Due to Glimmer mis-predictions when run on the plasmid sequences, plasmids pESBL and pAA instead used Heuristic GeneMark.hmm<sup>7</sup> PROKARYOTIC (version 2.8a) for gene calling. This was run with default settings and model file "heuristic\_no\_rbs.mat" ([http://opal.biology.gatech.edu/GeneMark/heuristic\\_hmm2.cgi](http://opal.biology.gatech.edu/GeneMark/heuristic_hmm2.cgi)). Putative protein products  $\geq 50$  aa in length were searched against the Genbank non-redundant protein database using PHMMER using HMMER (<http://hmmer.janelia.org/>). Genome visualisation plots were generated using CGview<sup>8</sup>.

Manual inspection of scaffolds revealed each plasmid was contained within a single scaffold. Manual curation of pG2011 demonstrated an  $\sim 1.5$ kb plasmid with  $>99\%$  nucleotide identity to *E. coli* strain H30 plasmid pO26-S1. This plasmid sequence was present as a 2-copy tandem repeat in the assembly, likely an artefact of the mate-pair assembly process (as insert

sizes longer than the plasmid were used) and has been manually edited to form a single copy.

The location of the plasmids in the assembly are as follows: pESBL-TY2482 = scaffold19, pAA-TY2482 = scaffold16, pG2011 = scaffold21.

### **Accession numbers**

The sequencing reads have been deposited into NCBI's Short Read Archive with accession numbers SRR227300, SRR227337, SRR227338, SRR227339, SRR227340, SRR231653, SRR231654 (Ion Torrent) and SRX079806 (Illumina mate-pair), SRX079805 (Illumina mate-pair) , SRX079804 (Illumina paired-end).

The scaffolded assembly and annotation has been deposited to Genbank, accession number AFVR000000000 (draft Illumina scaffold assembly), AFVS000000000 (Ion Torrent assembly) and AFOG010000000 (hybrid Ion Torrent and Illumina single-end assembly).

### **References**

1. Li R, Zhu H, Ruan J, et al. De novo assembly of human genomes with massively parallel short read sequencing. *Genome Res* 2010;20:265-72.
2. Sommer DD, Delcher AL, Salzberg SL, Pop M. Minimus: a fast, lightweight genome assembler. *BMC Bioinformatics*. 2007;8:64.
3. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* 2007;57:81-91.
4. Richter M, Rossello-Mora R. Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci U S A* 2009;106:19126-31.
5. Darling AE, Mau B, Perna NT. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* 2010;5:e11147.

6. Delcher AL, Bratke KA, Powers EC, Salzberg SL. Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics* 2007;23:673-679
7. Besemer J, Borodovsky M. Heuristic approach to deriving models for gene finding. *Nucleic Acids Res* 1999;27:3911-20.
8. Stothard P, Wishart DS. Circular genome visualization and exploration using CGView. *Bioinformatics* 2005;21:537-9.



**Table S3. Annotated genes on the RODs and plasmids of TY2482**

ROD_ID	CDS_ID	Best hit (NR)	Curated annotation	Best hit (species)
I-ROD1	irod1_orf00001	conserved domain protein		Escherichia coli MS 84-1
I-ROD1	irod1_orf00002	hypothetical protein ERKG_00886		Escherichia coli H252
I-ROD1	irod1_orf00003	transposon Tn21 resolvase		Escherichia coli B7A
I-ROD1	irod1_orf00005	hypothetical protein ECoL_00180		Escherichia coli EC4100B
I-ROD1	irod1_orf00007	conserved domain protein		Escherichia coli MS 187-1
I-ROD1	irod1_orf00008	conserved domain protein		Escherichia coli MS 187-1
I-ROD1	irod1_orf00009	hypothetical protein HMPREF9550_01817		Escherichia coli MS 187-1
I-ROD1	irod1_orf00011	putative acyl-carrier-protein S-malonyltransferase		Escherichia coli B7A
I-ROD1	irod1_orf00012	hypothetical protein EcB7A_3346		Escherichia coli B7A
I-ROD1	irod1_orf00013	hypothetical protein HMPREF9550_01813		Escherichia coli MS 187-1
I-ROD1	irod1_orf00015	hypothetical protein ECoL_00172		Escherichia coli EC4100B
I-ROD1	irod1_orf00016	hypothetical protein HMPREF9542_01440		Escherichia coli MS 117-3
I-ROD1	irod1_orf00017	phage integrase		Escherichia coli H252
I-ROD2	irod2_orf00001	integrase		Escherichia coli O26:H11 str. 11368
I-ROD2	irod2_orf00002	hypothetical bacteriophage protein		Escherichia coli O157:H7 str. EC4501
I-ROD2	irod2_orf00004	conserved hypothetical protein		Escherichia coli O157:H7 str. EC4501
I-ROD2	irod2_orf00006	putative antirepressor		Escherichia coli O103:H2 str. 12009
I-ROD2	irod2_orf00007	hypothetical protein ECO26_1566		Escherichia coli O26:H11 str. 11368
I-ROD2	irod2_orf00008	hypothetical protein ECO26_1570		Escherichia coli O26:H11 str. 11368
I-ROD2	irod2_orf00009	gp43		Escherichia coli B171
I-ROD2	irod2_orf00010	conserved hypothetical protein		Escherichia coli O157:H7 str. EC508
I-ROD2	irod2_orf00011	conserved hypothetical protein		Escherichia coli O157:H7 str. EC508
I-ROD2	irod2_orf00012	hypothetical protein ECO103_2866		Escherichia coli O103:H2 str. 12009
I-ROD2	irod2_orf00013	hypothetical protein ECO103_2865		Escherichia coli O103:H2 str. 12009
I-ROD2	irod2_orf00014	putative exonuclease		Escherichia coli O103:H2 str. 12009
I-ROD2	irod2_orf00015	RecT protein		Escherichia coli O157:H7 str. EC508
I-ROD2	irod2_orf00016	conserved hypothetical protein		Escherichia coli O157:H7 str. EC4501

I-ROD2	irod2_orf00017	hypothetical protein ECH7EC4501_4934		Escherichia coli O157:H7 str. EC4501
I-ROD2	irod2_orf00019	conserved domain protein		Escherichia coli O157:H7 str. EC4501
I-ROD2	irod2_orf00020	conserved hypothetical protein		Escherichia coli O157:H7 str. EC4501
I-ROD2	irod2_orf00021	phage regulatory protein, Rha family		Escherichia coli O157:H7 str. EC4501
I-ROD2	irod2_orf00022	type II restriction enzyme BsuBI		Escherichia coli O157:H7 str. TW14588
I-ROD2	irod2_orf00023	modification methylase BsuBI		Escherichia coli O157:H7 str. TW14588
I-ROD2	irod2_orf00025	repressor protein CI		Escherichia coli O157:H7 str. EC4501
I-ROD2	irod2_orf00026	hypothetical protein SDY_1924		Shigella dysenteriae Sd197
I-ROD2	irod2_orf00027	helicase domain protein		Escherichia coli O157:H7 str. EC4501
I-ROD2	irod2_orf00028	hypothetical protein ECOK1180_4044		Escherichia coli 1180
I-ROD2	irod2_orf00029	hypothetical protein ECO103_2848		Escherichia coli O103:H2 str. 12009
I-ROD2	irod2_orf00030	protein ninG		Escherichia coli 1180
I-ROD2	irod2_orf00033	DNA modification methylase		Stx2-converting phage 86
I-ROD2	irod2_orf00034	Shiga toxin 2 subunit A		Enterobacteria phage 933W
I-ROD2	irod2_orf00036	Shiga toxin 2 subunit B		Enterobacteria phage 933W
I-ROD2	irod2_orf00037	hypothetical protein		Shigella phage 7888
I-ROD2	irod2_orf00038	hypothetical protein ECs2970		Escherichia coli O157:H7 str. Sakai
I-ROD2	irod2_orf00039	conserved domain protein		Escherichia coli O157:H7 str. EC4113
I-ROD2	irod2_orf00040	lysozyme		Escherichia coli O157:H7 str. EC4501
I-ROD2	irod2_orf00041	anti-repressor protein Ant		Enterobacteria phage VT2phi_272
I-ROD2	irod2_orf00044	endopeptidase (Protein gp15)		Escherichia coli S88
I-ROD2	irod2_orf00045	Rha protein		Escherichia coli O157:H7 str. TW14588
I-ROD2	irod2_orf00048	putative terminase small subunit		Stx2-converting phage 86
I-ROD2	irod2_orf00049	hypothetical protein ECOK1180_4067		Escherichia coli 1180
I-ROD2	irod2_orf00050	large subunit terminase		Escherichia coli O157:H7 str. EC4113
I-ROD2	irod2_orf00051	putative phage portal protein		Stx2-converting phage 86
I-ROD2	irod2_orf00052	hypothetical protein 933Wp53		Enterobacteria phage 933W
I-ROD2	irod2_orf00053	hypothetical protein 933Wp54		Enterobacteria phage 933W
I-ROD2	irod2_orf00054	hypothetical protein 933Wp55		Enterobacteria phage 933W
I-ROD2	irod2_orf00055	hypothetical protein 933Wp56		Enterobacteria phage 933W
I-ROD2	irod2_orf00056	hypothetical protein ECs1226		Escherichia coli O157:H7 str. Sakai
I-ROD2	irod2_orf00057	hypothetical protein ECO103_2826		Escherichia coli O103:H2 str. 12009

I-ROD2	irod2_orf00059	tail fiber protein		Escherichia coli O157:H7 str. EC4113
I-ROD2	irod2_orf00061	hypothetical protein Stx2-86_gp25		Stx2-converting phage 86
I-ROD2	irod2_orf00063	conserved hypothetical protein		Escherichia coli O157:H7 str. EC4196
I-ROD2	irod2_orf00065	outer membrane protein Lom precursor		Enterobacteria phage 933W
I-ROD2	irod2_orf00066	conserved hypothetical protein		Escherichia coli O157:H7 str. EC4501
I-ROD2	irod2_orf00067	hypothetical protein 933Wp68		Enterobacteria phage 933W
I-ROD2	irod2_orf00068	conserved hypothetical protein		Escherichia coli O157:H7 str. EC4501
I-ROD2	irod2_orf00069	hypothetical protein 933Wp70		Enterobacteria phage 933W
I-ROD2	irod2_orf00071	hypothetical protein 933Wp71		Enterobacteria phage 933W
I-ROD2	irod2_orf00072	hypothetical protein Stx2-86_gp35		Stx2-converting phage 86
I-ROD2	irod2_orf00073	hypothetical protein		Enterobacteria phage Min27
I-ROD3	irod3_orf00001	predicted integrase		Escherichia sp. 4_1_40B
I-ROD3	irod3_orf00002	unknown		Shigella flexneri 2a
I-ROD3	irod3_orf00003	prophage CP4-57 regulatory protein alpA		Escherichia coli 3431
I-ROD3	irod3_orf00004	unknown		Shigella flexneri 2a
I-ROD3	irod3_orf00005	type III restriction enzyme, res subunit		Pelobacter propionicus DSM 2379
I-ROD3	irod3_orf00006	hypothetical protein E4_08923		Escherichia sp. 4_1_40B
I-ROD3	irod3_orf00007	Transposase		Shigella dysenteriae CDC 74-1112
I-ROD3	irod3_orf00008	IS66 family element, orf2		Shigella boydii CDC 3083-94
I-ROD3	irod3_orf00011	hypothetical protein		Escherichia coli
I-ROD3	irod3_orf00012	hypothetical protein HMPREF9552_03072		Escherichia coli MS 198-1
I-ROD3	irod3_orf00013	hypothetical protein E4_08823		Escherichia sp. 4_1_40B
I-ROD3	irod3_orf00014	hypothetical protein Z1185		Escherichia coli O157:H7 EDL933
I-ROD3	irod3_orf00015	conserved hypothetical protein		Escherichia coli O157:H7 str. TW14588
I-ROD3	irod3_orf00016	hypothetical protein Z1188		Escherichia coli O157:H7 EDL933
I-ROD3	irod3_orf00017	putative glucosyltransferase		Escherichia coli O157:H7 EDL933
I-ROD3	irod3_orf00018	putative ferric enterochelin esterase McmK		Escherichia coli
I-ROD3	irod3_orf00019	MchS2 protein		Escherichia coli
I-ROD3	irod3_orf00020	hypothetical protein p1ECUMN_0112		Escherichia coli UMN026
I-ROD3	irod3_orf00022	MchC protein		Escherichia coli CFT073
I-ROD3	irod3_orf00023	microcin H47 secretion protein		Escherichia coli 042
I-ROD3	irod3_orf00024	MtfB		Escherichia coli

I-ROD3	irod3_orf00025	conserved hypothetical protein		Escherichia coli O157:H7 str. EC4196
I-ROD3	irod3_orf00026	hypothetical protein ECDG_03856		Escherichia coli B185
I-ROD3	irod3_orf00027	hypothetical protein ROD_49891		Citrobacter rodentium ICC168
I-ROD3	irod3_orf00028	hypothetical protein ROD_49911		Citrobacter rodentium ICC168
I-ROD3	irod3_orf00029	ImpA-related N- superfamily		Escherichia coli M605
I-ROD3	irod3_orf00030	hypothetical protein AHA_1063		Aeromonas hydrophila subsp. hydrophila ATCC 7966
I-ROD3	irod3_orf00031	immunoglobulin-binding regulator A		Escherichia coli M605
I-ROD3	irod3_orf00032	insertion element IS1 7 protein insA		Shigella dysenteriae 1617
I-ROD3	irod3_orf00034	hypothetical protein ECNA114_2538		Escherichia coli NA114
I-ROD3	irod3_orf00035	putative transposase		Shigella flexneri K-671
I-ROD3	irod3_orf00036	putative ATP synthase F0, A subunit		Escherichia coli MS 116-1
I-ROD3	irod3_orf00037	aspartate racemase		Shigella flexneri K-272
I-ROD3	irod3_orf00038	hypothetical protein HMPREF9541_00362		Escherichia coli MS 116-1
I-ROD3	irod3_orf00039	putative transcriptional regulator		Shigella flexneri 2a
I-ROD3	irod3_orf00042	conserved domain protein		Escherichia coli MS 116-1
I-ROD3	irod3_orf00043	predicted protein		Nematostella vectensis
I-ROD3	irod3_orf00044	protein kinase		Yersinia pseudotuberculosis IP 31758
I-ROD3	irod3_orf00045	hypothetical protein ESA_01782		Cronobacter sakazakii ATCC BAA-894
I-ROD3	irod3_orf00046	putative tellurium resistance protein	TerY3	Serratia marcescens
I-ROD3	irod3_orf00047	putative tellurium resistance protein	TerY2	Serratia marcescens
I-ROD3	irod3_orf00049	tellurium resistance protein	TerX	Serratia marcescens
I-ROD3	irod3_orf00050	putative tellurium resistance protein TerY	TerY1	Enterobacter cloacae subsp. cloacae ATCC 13047
I-ROD3	irod3_orf00051	terW	TerW	Citrobacter sp. 30_2
I-ROD3	irod3_orf00052	hypothetical protein SMR0069		Serratia marcescens
I-ROD3	irod3_orf00053	hypothetical protein Z1166		Escherichia coli O157:H7 EDL933
I-ROD3	irod3_orf00054	putative ATP-binding protein		Escherichia coli APEC O1
I-ROD3	irod3_orf00055	hypothetical protein APECO1_O1R65		Escherichia coli APEC O1
I-ROD3	irod3_orf00056	hypothetical protein APECO1_O1R66		Escherichia coli APEC O1
I-ROD3	irod3_orf00057	hypothetical protein APECO1_O1R67		Escherichia coli APEC O1
I-ROD3	irod3_orf00058	putative phage inhibition, colicin resistance and tellurite resistance protein	TerZ	Escherichia coli O157:H7 EDL933

I-ROD3	irod3_orf00059	putative phage inhibition, colicin resistance and tellurite resistance protein	TerA	Escherichia coli O157:H7 EDL933
I-ROD3	irod3_orf00060	putative phage inhibition, colicin resistance and tellurite resistance protein	TerB	Escherichia coli O157:H7 EDL933
I-ROD3	irod3_orf00061	putative phage inhibition, colicin resistance and tellurite resistance protein	TerC	Escherichia coli O157:H7 EDL933
I-ROD3	irod3_orf00063	putative phage inhibition, colicin resistance and tellurite resistance protein	TerD	Escherichia coli O157:H7 EDL933
I-ROD3	irod3_orf00064	putative phage inhibition, colicin resistance and tellurite resistance protein	TerE	Escherichia coli O157:H7 EDL933
I-ROD3	irod3_orf00065	putative tellurium resistance protein TerF	TerF	Escherichia coli O103:H2 str. 12009
I-ROD3	irod3_orf00067	putative GTP-binding protein		Escherichia coli 042
I-ROD3	irod3_orf00068	antigen 43 precursor		Escherichia coli
I-ROD3	irod3_orf00071	putative autotransporter		Shigella sp. D9
I-ROD3	irod3_orf00072	hypothetical protein EscherichiacoliO157_22726		Escherichia coli O157:H7 str. FRIK2000
I-ROD3	irod3_orf00073	hypothetical protein ECS88_2092		Escherichia coli S88
I-ROD3	irod3_orf00074	hypothetical protein ECS88_2092		Escherichia coli S88
I-ROD3	irod3_orf00075	conserved hypothetical protein		Escherichia coli H591
I-ROD3	irod3_orf00077	hypothetical protein SD1617_3951		Shigella dysenteriae 1617
I-ROD3	irod3_orf00078	hypothetical protein ECS88_2094		Escherichia coli S88
I-ROD3	irod3_orf00079	hypothetical protein APECO1_1098		Escherichia coli APEC O1
I-ROD3	irod3_orf00080	hypothetical protein ECO103_3758		Escherichia coli O103:H2 str. 12009
I-ROD3	irod3_orf00081	hypothetical protein ECNA114_2131		Escherichia coli NA114
I-ROD3	irod3_orf00083	toxin of the YeeV-YeeU toxin-antitoxin system		Escherichia sp. 4_1_40B
I-ROD3	irod3_orf00084	conserved hypothetical protein		Escherichia coli ETEC H10407
I-ROD3	irod3_orf00086	hypothetical protein UTI89_C4999		Escherichia coli UTI89
I-ROD3	irod3_orf00087	hypothetical protein Z1226		Escherichia coli O157:H7 EDL933
I-ROD4	irod4_orf00001	AntB		Escherichia coli
I-ROD4	irod4_orf00003	conserved hypothetical protein		Escherichia coli FVEC1302
I-ROD4	irod4_orf00004	valyl-tRNA synthetase		Escherichia coli E110019
I-ROD4	irod4_orf00006	hypothetical protein Stx2-86_gp35		Stx2-converting phage 86
I-ROD4	irod4_orf00008	hypothetical protein SDY_1670		Shigella dysenteriae Sd197
I-ROD4	irod4_orf00010	hypothetical protein ECED1_1152		Escherichia coli ED1a
I-ROD4	irod4_orf00011	hypothetical protein ECED1_1151		Escherichia coli ED1a
I-ROD4	irod4_orf00012	hypothetical protein 933Wp68		Enterobacteria phage 933W
I-ROD4	irod4_orf00013	hypothetical protein Stx2-86_gp30		Stx2-converting phage 86
I-ROD4	irod4_orf00014	putative outer membrane precursor Lom		Escherichia coli O103:H2 str. 12009

I-ROD4	irod4_orf00016	hypothetical protein Stx2Ip034		Stx2 converting phage I
I-ROD4	irod4_orf00018	hypothetical protein Stx2-86_gp25		Stx2-converting phage 86
I-ROD4	irod4_orf00020	putative long tail fiber protein		Stx2-converting phage 86
I-ROD4	irod4_orf00021	hypothetical protein ECED1_1137		Escherichia coli ED1a
I-ROD4	irod4_orf00022	hypothetical protein ECED1_1136		Escherichia coli ED1a
I-ROD4	irod4_orf00023	hypothetical protein ECED1_1135		Escherichia coli ED1a
I-ROD4	irod4_orf00024	hypothetical protein Stx2-86_gp17		Stx2-converting phage 86
I-ROD4	irod4_orf00025	hypothetical protein ECED1_1133		Escherichia coli ED1a
I-ROD4	irod4_orf00026	hypothetical protein ECED1_1132		Escherichia coli ED1a
I-ROD4	irod4_orf00027	putative phage portal protein		Stx2-converting phage 86
I-ROD4	irod4_orf00028	hypothetical protein ECOK1180_4067		Escherichia coli 1180
I-ROD4	irod4_orf00029	large subunit terminase		Escherichia coli O157:H7 str. EC4113
I-ROD4	irod4_orf00030	putative terminase small subunit		Stx2-converting phage 86
I-ROD4	irod4_orf00033	bacteriophage lysis protein		Shigella dysenteriae 1012
I-ROD4	irod4_orf00036	putative endolysin		Shigella dysenteriae Sd197
I-ROD4	irod4_orf00037	protein S		Enterobacteria phage 933W
I-ROD4	irod4_orf00038	conserved hypothetical protein		Shigella dysenteriae 1617
I-ROD4	irod4_orf00039	hypothetical protein SGF_04061		Shigella flexneri CDC 796-83
I-ROD4	irod4_orf00040	YjhS		Shigella boydii CDC 3083-94
I-ROD4	irod4_orf00041	putative NinH protein		Phage BP-4795
I-ROD4	irod4_orf00042	crossover junction endodeoxyribonuclease		Escherichia coli ED1a
I-ROD4	irod4_orf00044	hypothetical protein E2348C_2522		Escherichia coli O127:H6 str. E2348/69
I-ROD4	irod4_orf00045	putative ninB protein		Escherichia coli ED1a
I-ROD4	irod4_orf00046	putative antirepressor protein Ant from prophage		Escherichia coli ED1a
I-ROD4	irod4_orf00047	hypothetical protein ECO26_2262		Escherichia coli O26:H11 str. 11368
I-ROD4	irod4_orf00048	death-on-curing family protein		Escherichia coli STEC_7v
I-ROD4	irod4_orf00049	hypothetical protein ECSTEC7V_1837		Escherichia coli STEC_7v
I-ROD4	irod4_orf00050	hypothetical protein ECO111_1061		Escherichia coli O111:H- str. 11128
I-ROD4	irod4_orf00051	hypothetical protein G2583_1712		Escherichia coli O55:H7 str. CB9615
I-ROD4	irod4_orf00052	hypothetical protein EcE24377A_1426		Escherichia coli E24377A
I-ROD4	irod4_orf00053	hypothetical protein ECO103_1369		Escherichia coli O103:H2 str. 12009
I-ROD4	irod4_orf00055	putative replication protein		Escherichia coli ED1a

I-ROD4	irod4_orf00056	hypothetical protein ECED1_1103		Escherichia coli ED1a
I-ROD4	irod4_orf00057	hypothetical protein ECED1_1102		Escherichia coli ED1a
I-ROD4	irod4_orf00058	regulatory protein CII from prophage		Escherichia coli ED1a
I-ROD4	irod4_orf00059	prophage repressor CI		Enterobacteria phage HK97
I-ROD4	irod4_orf00060	hypothetical protein ECED1_1098		Escherichia coli ED1a
I-ROD4	irod4_orf00061	hypothetical protein ECED1_1097		Escherichia coli ED1a
I-ROD4	irod4_orf00063	monocarboxylate transporter		Culex quinquefasciatus
I-ROD4	irod4_orf00064	hypothetical protein ECED1_1095		Escherichia coli ED1a
I-ROD4	irod4_orf00065	hypothetical protein ECED1_1094		Escherichia coli ED1a
I-ROD4	irod4_orf00067	FtsZ inhibitor protein		Escherichia coli ED1a
I-ROD4	irod4_orf00068	hypothetical protein ECED1_1091		Escherichia coli ED1a
I-ROD4	irod4_orf00069	Exodeoxyribonuclease VIII (putative partial) from phage origin		Escherichia coli ED1a
I-ROD4	irod4_orf00070	putative host-nuclease inhibitor protein Gam		Shigella dysenteriae Sd197
I-ROD4	irod4_orf00071	Recombination protein bet from phage origin		Escherichia coli ED1a
I-ROD4	irod4_orf00072	putative exonuclease encoded by prophage CP-933K		Escherichia coli O157:H7 EDL933
I-ROD4	irod4_orf00074	prophage DLP12 integrase		Escherichia coli 101-l
I-ROD5	irod5_orf00001	hypothetical protein SSON_1273		Shigella sonnei Ss046
I-ROD5	irod5_orf00002	hypothetical protein EC55989_1079		Escherichia coli 55989
I-ROD5	irod5_orf00005	triple helix repeat-containing collagen		Clostridium beijerinckii NCIMB 8052
I-ROD5	irod5_orf00006	hypothetical protein SD15574_2985		Shigella dysenteriae 155-74
I-ROD5	irod5_orf00008	hypothetical protein ECE128010_5420		Escherichia coli E128010
I-ROD5	irod5_orf00011	Putative tail component of prophage		Escherichia coli NA114
I-ROD5	irod5_orf00012	hypothetical protein ECLG_05105		Escherichia coli TA271
I-ROD5	irod5_orf00015	Superoxide dismutase (Cu-Zn)		Escherichia coli O55:H7 str. CB9615
I-ROD5	irod5_orf00019	minor tail protein		Escherichia coli UTI89
I-ROD5	irod5_orf00021	minor tail protein		Escherichia coli UTI89
I-ROD5	irod5_orf00022	putative tail fiber component H of prophage CP-933U		Escherichia coli O157:H7 EDL933
I-ROD5	irod5_orf00024	Phage minor tail protein		Escherichia coli EC4100B
I-ROD5	irod5_orf00025	Phage minor tail protein		Escherichia coli EC4100B
I-ROD5	irod5_orf00026	phage major tail protein		Escherichia coli 042
I-ROD5	irod5_orf00027	hypothetical protein DAPPUDRAFT_279812		Daphnia pulex
I-ROD5	irod5_orf00029	polysaccharide Transporter, PST family		Enterococcus faecium E1679

I-ROD5	irod5_orf00032	hypothetical protein		Arthrospira platensis NIES-39
I-ROD5	irod5_orf00034	terminase large subunit domain protein		Escherichia coli RN587/1
I-ROD5	irod5_orf00035	conserved hypothetical protein		Escherichia albertii TW07627
I-ROD5	irod5_orf00036	phage major capsid protein E		Escherichia coli H489
I-ROD5	irod5_orf00039	Hypothetical protein CBG02325		Caenorhabditis briggsae
I-ROD5	irod5_orf00041	conserved domain protein		Escherichia coli MS 153-1
I-ROD5	irod5_orf00042	hypothetical protein c1457		Escherichia coli CFT073
I-ROD5	irod5_orf00043	Phage minor tail protein		Escherichia coli EC4100B
I-ROD5	irod5_orf00045	hypothetical protein MK0973		Methanopyrus kandleri AV19
I-ROD5	irod5_orf00047	hypothetical protein SCA50_1305		Salmonella enterica subsp. enterica serovar Choleraesuis str. SCSA50
I-ROD5	irod5_orf00049	hypothetical protein ECOK1_1278		Escherichia coli IHE3034
I-ROD5	irod5_orf00052	phage DNA packaging protein Nu1		Escherichia coli MS 21-1
I-ROD5	irod5_orf00053	putative phage protein		Escherichia coli 042
I-ROD5	irod5_orf00055	hypothetical protein ECS88_0566		Escherichia coli S88
I-ROD5	irod5_orf00056	endopeptidase		Escherichia coli 2362-75
I-ROD5	irod5_orf00058	hypothetical protein SBO_1923		Shigella boydii Sb227
I-ROD5	irod5_orf00059	putative membrane-associated lysozyme; Qin prophage		Escherichia coli 55989
I-ROD5	irod5_orf00061	hypothetical protein Stx2-86_gp06		Stx2-converting phage 86
I-ROD5	irod5_orf00062	hypothetical protein Stx2-86_gp05		Stx2-converting phage 86
I-ROD5	irod5_orf00063	lysis protein S		Stx2-converting phage 86
I-ROD5	irod5_orf00066	hypothetical protein DAPPUDRAFT_52038		Daphnia pulex
I-ROD5	irod5_orf00068	heterokaryon incompatibility protein		Glomerella graminicola M1.001
I-ROD5	irod5_orf00070	DNA methylase family protein		Shigella flexneri J1713
I-ROD5	irod5_orf00071	hypothetical protein HMPREF9542_00842		Escherichia coli MS 117-3
I-ROD5	irod5_orf00072	hypothetical protein EcF11_4284		Escherichia coli F11
I-ROD5	irod5_orf00075	hypothetical protein ECRN5871_4170		Escherichia coli RN587/1
I-ROD5	irod5_orf00076	hypothetical protein E4_10746		Escherichia sp. 4_1_40B
I-ROD5	irod5_orf00077	endodeoxyribonuclease RusA family protein		Escherichia coli STEC_7v
I-ROD5	irod5_orf00078	LexA repressor		Escherichia coli S88
I-ROD5	irod5_orf00079	DNA adenine methylase		Escherichia coli UTI89
I-ROD5	irod5_orf00080	hypothetical protein ECS88_0547		Escherichia coli S88



I-ROD5	irod5_orf00081	hypothetical protein PcdtI_gp46		Phage cdtI
I-ROD5	irod5_orf00082	putative antirepressor		Escherichia coli EC4100B
I-ROD5	irod5_orf00083	nucleic acid-binding protein; e14 prophage		Escherichia coli S88
I-ROD5	irod5_orf00084	hypothetical protein ECD227_2469		Escherichia fergusonii ECD227
I-ROD5	irod5_orf00085	regulatory protein cI		Escherichia coli EC4100B
I-ROD5	irod5_orf00086	hypothetical protein ECoL_03975		Escherichia coli EC4100B
I-ROD5	irod5_orf00087	hypothetical protein ECoL_03976		Escherichia coli EC4100B
I-ROD5	irod5_orf00089	Hypothetical protein yfdR		Escherichia coli EC4100B
I-ROD5	irod5_orf00090	hypothetical protein ShiD9_12075		Shigella sp. D9
I-ROD5	irod5_orf00091	conserved hypothetical protein		Escherichia coli E22
I-ROD5	irod5_orf00093	conserved hypothetical protein		Escherichia coli E22
I-ROD5	irod5_orf00095	Phage EaA protein		Escherichia coli EC4100B
I-ROD5	irod5_orf00096	Phage EaA protein		Escherichia coli EC4100B
I-ROD5	irod5_orf00097	Integrase		Escherichia coli EC4100B
I-ROD6	irod6_orf00001	molybdate metabolism regulator		Escherichia coli 536
I-ROD6	irod6_orf00003	hypothetical protein ECP_2154		Escherichia coli 536
I-ROD6	irod6_orf00005	yehL protein		Escherichia coli B088
I-ROD6	irod6_orf00006	hypothetical protein ECP_2157		Escherichia coli 536
I-ROD6	irod6_orf00007	hypothetical protein ECIAI1_2197		Escherichia coli IAI1
I-ROD6	irod6_orf00008	hypothetical protein ECP_2159		Escherichia coli 536
I-ROD7	irod7_orf00001	integrase		Escherichia coli
I-ROD7	irod7_orf00002	Evolved beta-D-galactosidase, beta subunit		Shigella dysenteriae CDC 74-1112
I-ROD7	irod7_orf00003	transposase TnpA		Corynebacterium glutamicum
I-ROD7	irod7_orf00004	resolvase for Tn21		Plasmid R100
I-ROD7	irod7_orf00006	Urf2 protein		Escherichia fergusonii ECD227
I-ROD7	irod7_orf00007	integrase		Plasmid R100
I-ROD7	irod7_orf00008	dihydrofolate reductase type A7		Salmonella enterica subsp. enterica serovar Weltevreden
I-ROD7	irod7_orf00010	putative transposase		Klebsiella pneumoniae subsp. pneumoniae MGH 78578
I-ROD7	irod7_orf00014	3-hydroxyisobutyrate dehydrogenase		Mycobacterium tuberculosis 210
I-ROD7	irod7_orf00016	protein RepC		Salmonella enterica subsp. enterica serovar Enteritidis

I-ROD7	irod7_orf00017	dihydropteroate synthase		Salmonella enterica subsp. enterica serovar Typhi str. CT18
I-ROD7	irod7_orf00018	aminoglycoside/hydroxyurea antibiotic resistance kinase		Escherichia coli MS 200-1
I-ROD7	irod7_orf00019	beta-lactamase		Escherichia coli 3431
I-ROD7	irod7_orf00021	hypothetical protein R100p008		Plasmid R100
I-ROD7	irod7_orf00022	putative mercury resistance protein		Plasmid R100
I-ROD7	irod7_orf00023	transcriptional regulator MerD		Plasmid R100
I-ROD7	irod7_orf00026	RecName: Full=Mercuric reductase; AltName: Full=Hg(II) reductase		
I-ROD7	irod7_orf00027	putative mercury transport protein MerC		Aeromonas salmonicida subsp. salmonicida A449
I-ROD7	irod7_orf00029	Tn501 orf, hypotheical		Shigella flexneri 5a
I-ROD7	irod7_orf00033	InsL		Escherichia coli 53638
I-ROD7	irod7_orf00034	hypothetical protein pFL129_4		Escherichia coli
I-ROD7	irod7_orf00036	TetA		Salmonella enterica subsp. enterica serovar Choleraesuis
I-ROD7	irod7_orf00037	integral membrane protein DUF6		Escherichia coli MS 78-1
I-ROD7	irod7_orf00038	hypothetical protein HMPREF9544_05491		Escherichia coli MS 153-1
I-ROD7	irod7_orf00039	conserved hypothetical protein		Escherichia coli ETEC H10407
I-ROD7	irod7_orf00040	conserved hypothetical protein		Escherichia coli SE15
I-ROD7	irod7_orf00042	hypothetical protein HMPREF9553_03865		Escherichia coli MS 200-1
I-ROD7	irod7_orf00044	putative regulatory protein		Escherichia coli 536
I-ROD7	irod7_orf00045	conserved hypothetical protein		Escherichia coli SE15
I-ROD7	irod7_orf00046	transposase		Escherichia coli SE15
I-ROD7	irod7_orf00048	hypothetical protein		Escherichia coli SE15
I-ROD7	irod7_orf00049	hypothetical protein ECUMN_4880		Escherichia coli UMN026
I-ROD7	irod7_orf00051	putative autotransporter		Escherichia coli 536
I-ROD7	irod7_orf00052	antigen 43 domain protein		Escherichia coli LT-68
I-ROD7	irod7_orf00053	hypothetical protein EcE24377A_4893		Escherichia coli E24377A
I-ROD7	irod7_orf00054	hypothetical protein ECNA114_2131		Escherichia coli NA114
I-ROD7	irod7_orf00056	conserved domain protein		Escherichia coli MS 187-1
I-ROD7	irod7_orf00058	conserved hypothetical protein		Escherichia coli SE15
I-ROD7	irod7_orf00059	putative radC-like protein yeeS		Escherichia coli CFT073
I-ROD7	irod7_orf00060	hypothetical protein c0272		Escherichia coli CFT073

I-ROD7	irod7_orf00061	unknown		Escherichia coli
I-ROD7	irod7_orf00063	DNA repair protein		Escherichia coli MS 78-1
I-ROD7	irod7_orf00064	hypothetical protein c4574		Escherichia coli CFT073
I-ROD7	irod7_orf00065	conserved hypothetical protein		Shigella dysenteriae 1617
I-ROD7	irod7_orf00067	hypothetical protein APECO1_3486		Escherichia coli APEC O1
I-ROD7	irod7_orf00068	hypothetical protein SF3000		Shigella flexneri 2a str. 301
I-ROD7	irod7_orf00069	hypothetical protein ECO103_3594		Escherichia coli O103:H2 str. 12009
I-ROD7	irod7_orf00070	hypothetical protein ECED1_4984		Escherichia coli ED1a
pESBL	scaffold19_orf0002	YciB		Escherichia coli
pESBL	scaffold19_orf0003	hypothetical protein pECBactecp21		Escherichia coli
pESBL	scaffold19_orf0004	hypothetical protein SC121		Salmonella enterica subsp. enterica serovar Choleraesuis str. SC-B67
pESBL	scaffold19_orf0005	single-stranded DNA-binding protein		Salmonella enterica subsp. enterica serovar Choleraesuis str. SC-B67
pESBL	scaffold19_orf0006	hypothetical protein pO157p50		Escherichia coli O157:H7 str. Sakai
pESBL	scaffold19_orf0007	plasmid SOS inhibition protein B		Escherichia coli
pESBL	scaffold19_orf0008	plasmid SOS inhibition protein A		Salmonella enterica subsp. enterica serovar Choleraesuis str. SC-B67
pESBL	scaffold19_orf0009	hypothetical protein SC115		Salmonella enterica subsp. enterica serovar Choleraesuis str. SC-B67
pESBL	scaffold19_orf0010	antirestriction protein		Escherichia coli MS 107-1
pESBL	scaffold19_orf0011	hypothetical protein ECSE_P1-0063		Escherichia coli SE11
pESBL	scaffold19_orf0012	hypothetical protein HMPREF9542_03988		Escherichia coli MS 117-3
pESBL	scaffold19_orf0013	hypothetical protein SeHA_A0062		Salmonella enterica subsp. enterica serovar Heidelberg str. SL476
pESBL	scaffold19_orf0014	hypothetical protein EcE22_3665		Escherichia coli E22
pESBL	scaffold19_orf0015	CcgAII protein		Salmonella enterica subsp. enterica serovar Choleraesuis str. SC-B67
pESBL	scaffold19_orf0016	putative transposase		Escherichia coli E22
pESBL	scaffold19_orf0018	hypothetical protein SC107		Salmonella enterica subsp. enterica serovar Choleraesuis str. SC-B67
pESBL	scaffold19_orf0019	hypothetical protein R64_p076		Salmonella enterica subsp. enterica serovar Typhimurium
pESBL	scaffold19_orf0020	hypothetical protein SC105		Salmonella enterica subsp. enterica serovar Choleraesuis str. SC-B67

pESBL	scaffold19_orf0021	hypothetical protein pECBactecp14		Escherichia coli
pESBL	scaffold19_orf0022	hypothetical protein LH0067		Escherichia coli
pESBL	scaffold19_orf0023	relaxosome component		Plasmid Collb-P9
pESBL	scaffold19_orf0024	NikB		Escherichia coli O157:H7 str. Sakai
pESBL	scaffold19_orf0025	hypothetical protein EcE24377A_D0057		Escherichia coli E24377A
pESBL	scaffold19_orf0026	hypothetical protein pECBactecp09		Escherichia coli
pESBL	scaffold19_orf0027	hypothetical protein pECBactecp08		Escherichia coli
pESBL	scaffold19_orf0028	putative protein FinQ		Escherichia coli MS 84-1
pESBL	scaffold19_orf0029	counter protein for PndA		Escherichia coli
pESBL	scaffold19_orf0030	hypothetical protein SC084		Salmonella enterica subsp. enterica serovar Choleraesuis str. SC-B67
pESBL	scaffold19_orf0031	conserved hypothetical protein		Escherichia coli MS 107-1
pESBL	scaffold19_orf0032	hypothetical protein ECSE_P1-0081		Escherichia coli SE11
pESBL	scaffold19_orf0033	putative regulator protein		Escherichia coli SE11
pESBL	scaffold19_orf0034	exclusion-determining family protein		Escherichia coli MS 84-1
pESBL	scaffold19_orf0035	TraY		Escherichia coli O157:H7 str. EC4486
pESBL	scaffold19_orf0036	F pilin acetylation protein		Escherichia coli
pESBL	scaffold19_orf0037	F pilus assembly		Escherichia coli
pESBL	scaffold19_orf0038	F pilus assembly		Escherichia coli
pESBL	scaffold19_orf0039	TraU		Escherichia coli O157:H7 str. EC4401
pESBL	scaffold19_orf0040	hypothetical protein HMPREF9542_01329		Escherichia coli MS 117-3
pESBL	scaffold19_orf0041	TraR protein		Escherichia coli
pESBL	scaffold19_orf0042	hypothetical protein Collb-P9_p070		Plasmid Collb-P9
pESBL	scaffold19_orf0043	hypothetical protein Collb-P9_p071		Plasmid Collb-P9
pESBL	scaffold19_orf0044	hypothetical protein Collb-P9_p072		Plasmid Collb-P9
pESBL	scaffold19_orf0045	hypothetical protein Collb-P9_p073		Plasmid Collb-P9
pESBL	scaffold19_orf0046	hypothetical protein Collb-P9_p074		Plasmid Collb-P9
pESBL	scaffold19_orf0047	thick pilus signal peptide		Escherichia coli W
pESBL	scaffold19_orf0048	DNA primase		Escherichia coli O157:H7 str. EC4401
pESBL	scaffold19_orf0049	EDTA-resistant nuclease		Escherichia coli
pESBL	scaffold19_orf0051	ATP-binding protein		Plasmid Collb-P9
pESBL	scaffold19_orf0052	lipoprotein		Salmonella enterica subsp. enterica serovar

				Typhimurium
pESBL	scaffold19_orf0053	hypothetical protein Collb-P9_p082		Plasmid Collb-P9
pESBL	scaffold19_orf0054	hypothetical protein Collb-P9_p083		Plasmid Collb-P9
pESBL	scaffold19_orf0055	F pilus assembly		Escherichia coli
pESBL	scaffold19_orf0056	TraE protein		Escherichia coli
pESBL	scaffold19_orf0057	shufflon-specific DNA recombinase		Escherichia coli AA86
pESBL	scaffold19_orf0058	hypothetical protein HMPREF9536_01879		Escherichia coli MS 84-1
pESBL	scaffold19_orf0059	conserved hypothetical protein		Escherichia coli MS 107-1
pESBL	scaffold19_orf0060	hypothetical protein R64_p118		Salmonella enterica subsp. enterica serovar Typhimurium
pESBL	scaffold19_orf0061	shufflon protein C'		Escherichia coli O157:H7 str. EC4486
pESBL	scaffold19_orf0062	conserved hypothetical protein		Escherichia coli MS 107-1
pESBL	scaffold19_orf0063	shufflon protein A		Salmonella enterica subsp. enterica serovar Kentucky str. CVM29188
pESBL	scaffold19_orf0064	peptidase A24A prepilin type IV		Escherichia coli W
pESBL	scaffold19_orf0065	type IV prepilin cluster		Escherichia coli
pESBL	scaffold19_orf0066	type IV prepilin cluster; prepilin		Escherichia coli
pESBL	scaffold19_orf0067	integral membrane protein		Escherichia coli E22
pESBL	scaffold19_orf0068	ATP-binding protein PilQ		Escherichia coli SE11
pESBL	scaffold19_orf0069	IncI1 conjugal transfer protein PilP		Escherichia coli
pESBL	scaffold19_orf0070	IncI1 conjugal transfer protein PilO		Escherichia coli
pESBL	scaffold19_orf0071	lipoprotein PilN		Escherichia coli SE11
pESBL	scaffold19_orf0072	hypothetical protein Collb-P9_p101		Plasmid Collb-P9
pESBL	scaffold19_orf0073	IncI1 conjugal transfer protein PilL		Escherichia coli
pESBL	scaffold19_orf0074	predicted protein		Nematostella vectensis
pESBL	scaffold19_orf0075	IncI1 conjugal transfer protein TraC		Escherichia coli
pESBL	scaffold19_orf0076	transcription termination factor NusG		Escherichia coli MS 84-1
pESBL	scaffold19_orf0077	TraA protein		Escherichia coli SE11
pESBL	scaffold19_orf0078	replication initiation protein		Salmonella enterica subsp. enterica serovar Kentucky str. CVM29188
pESBL	scaffold19_orf0079	hypothetical protein ND12IncI1_3		Escherichia coli
pESBL	scaffold19_orf0080	hypothetical protein pECBactecp34		Escherichia coli
pESBL	scaffold19_orf0081	YagA		Escherichia coli O157:H7 str. EC4486

pESBL	scaffold19_orf0082	transposase		Salmonella enterica subsp. enterica serovar Infantis
pESBL	scaffold19_orf0083	conserved hypothetical protein		Escherichia coli MS 21-1
pESBL	scaffold19_orf0084	hypothetical protein		Escherichia coli
pESBL	scaffold19_orf0085	hypothetical protein pC15-1a_016	blaCTX-M-15	Escherichia coli
pESBL	scaffold19_orf0086	ISEcp1 transposase		Escherichia coli
pESBL	scaffold19_orf0087	transposase for transposon Tn3		Escherichia coli
pESBL	scaffold19_orf0088	hypothetical protein pC15-1a_019		Escherichia coli
pESBL	scaffold19_orf0089	TEM-1 beta-lactamase	blaTEM-1	Salmonella enterica subsp. enterica serovar Montevideo
pESBL	scaffold19_orf0090	conserved domain protein		Escherichia coli MS 21-1
pESBL	scaffold19_orf0091	cobyrinic acid a,c-diamide synthase		Escherichia coli
pESBL	scaffold19_orf0093	protein impB domain protein		Escherichia coli 1357
pESBL	scaffold19_orf0094	hypothetical protein Collb-P9_p029		Plasmid Collb-P9
pESBL	scaffold19_orf0095	DinI-like family protein		Escherichia coli MS 21-1
pESBL	scaffold19_orf0096	hypothetical protein p026VIR_p092		Escherichia coli
pESBL	scaffold19_orf0097	hypothetical protein ECO103_p71		Escherichia coli O103:H2 str. 12009
pESBL	scaffold19_orf0098	conserved hypothetical protein		Escherichia coli H299
pESBL	scaffold19_orf0099	hypothetical protein ND12IncI1_24		Escherichia coli
pESBL	scaffold19_orf0100	conserved hypothetical protein		Escherichia coli W
pAA	scaffold16_orf0001	putative secreted protein		Streptomyces hygroscopicus ATCC 53653
pAA	scaffold16_orf0002	hypothetical protein c3579		Escherichia coli CFT073
pAA	scaffold16_orf0003	unknown protein encoded in ISEc8		Escherichia coli O157:H7 EDL933
pAA	scaffold16_orf0004	hypothetical protein SbBS512_A0019		Shigella boydii CDC 3083-94
pAA	scaffold16_orf0005	AggA457 protein	AggA	Escherichia coli
pAA	scaffold16_orf0006	RecName: Full=Protein AggB; Flags: Precursor	AggB	
pAA	scaffold16_orf0007	HdaC, HUS-associated diffuse adherence	AggC	Escherichia coli
pAA	scaffold16_orf0008	RecName: Full=Chaperone protein AggD; Flags: Precursor	AggD	
pAA	scaffold16_orf0010	putative resolvase		Escherichia coli
pAA	scaffold16_orf0011	3-hydroxyisobutyrate dehydrogenase		Mycobacterium tuberculosis 210
pAA	scaffold16_orf0012	hypothetical protein Collb-P9_p027		Plasmid Collb-P9
pAA	scaffold16_orf0013	StbA protein		Escherichia coli MS 84-1

pAA	scaffold16_orf0015	putative 60 kDa chaperonin		Escherichia coli E24377A
pAA	scaffold16_orf0016	hypothetical protein ColIb-P9_p024		Plasmid ColIb-P9
pAA	scaffold16_orf0017	resolvase		Salmonella enterica subsp. enterica serovar Kentucky str. CVM29188
pAA	scaffold16_orf0018	plasmid maintenance protein CcdB		Escherichia coli
pAA	scaffold16_orf0019	plasmid maintenance protein CcdA		Escherichia coli
pAA	scaffold16_orf0021	hypothetical protein E4_23171		Escherichia sp. 4_1_40B
pAA	scaffold16_orf0022	hypothetical protein p1ECUMN_0160		Escherichia coli UMN026
pAA	scaffold16_orf0024	orf906		Escherichia coli
pAA	scaffold16_orf0026	phage integrase		Escherichia coli M863
pAA	scaffold16_orf0027	COG1506: Dipeptidyl aminopeptidases/acylaminoacyl-peptidases		Magnetospirillum magnetotacticum MS-1
pAA	scaffold16_orf0028	hypothetical protein pECL46p020		Escherichia coli
pAA	scaffold16_orf0029	hypothetical protein pEC55989_0007		Escherichia coli 55989
pAA	scaffold16_orf0030	hypothetical protein IPF_103		Escherichia coli 1520
pAA	scaffold16_orf0031	incFII family plasmid replication initiator RepA		Escherichia coli MS 78-1
pAA	scaffold16_orf0032	replication initiation protein		Escherichia coli E128010
pAA	scaffold16_orf0033	replication protein		Escherichia sp. 4_1_40B
pAA	scaffold16_orf0034	conjugal transfer pilus acetylation protein TraX		Shigella flexneri 2a str. 301
pAA	scaffold16_orf0035	hypothetical protein pYT1_p113		Salmonella enterica subsp. enterica serovar Typhimurium
pAA	scaffold16_orf0036	DNA helicase TraI		Escherichia coli MS 57-2
pAA	scaffold16_orf0037	conserved hypothetical protein		Salmonella enterica subsp. enterica serovar Kentucky
pAA	scaffold16_orf0038	hypothetical protein c3659		Escherichia coli CFT073
pAA	scaffold16_orf0039	hypothetical protein c3661		Escherichia coli CFT073
pAA	scaffold16_orf0040	hypothetical protein pB171_031		Escherichia coli
pAA	scaffold16_orf0041	conserved hypothetical protein		Escherichia coli H299
pAA	scaffold16_orf0042	conjugal transfer fertility inhibition protein FinO		Escherichia coli
pAA	scaffold16_orf0043	conjugal transfer pilus acetylation protein TraX		Salmonella enterica subsp. enterica serovar Kentucky str. CVM29188
pAA	scaffold16_orf0044	hypothetical protein pYT1_p113		Salmonella enterica subsp. enterica serovar Typhimurium
pAA	scaffold16_orf0045	conjugal transfer nickase/helicase TraI		Escherichia coli

pAA	scaffold16_orf0046	conjugal transfer nickase/helicase TraI		Salmonella enterica subsp. enterica serovar Kentucky str. CVM29188
pAA	scaffold16_orf0047	hypothetical protein R100p115.2br		Plasmid R100
pAA	scaffold16_orf0048	Protein traJ		Escherichia coli 55989
pAA	scaffold16_orf0049	TraM		Escherichia coli
pAA	scaffold16_orf0050	putative lytic transglycosylase		Escherichia coli ETEC H10407
pAA	scaffold16_orf0051	conserved hypothetical protein		Escherichia coli MS 185-1
pAA	scaffold16_orf0052	putative recombinase		Escherichia coli
pAA	scaffold16_orf0053	SepA		Escherichia coli 536
pAA	scaffold16_orf0054	putative transposase		Escherichia coli
pAA	scaffold16_orf0057	conserved hypothetical protein		Escherichia coli MS 153-1
pAA	scaffold16_orf0058	AatD		Escherichia sp. 4_1_40B
pAA	scaffold16_orf0059	AatC ATB binding protein of ABC transporter		Escherichia coli 55989
pAA	scaffold16_orf0060	AatB		Escherichia coli 55989
pAA	scaffold16_orf0061	AatA outermembrane protein		Escherichia coli 55989
pAA	scaffold16_orf0062	AatP permease		Escherichia sp. 4_1_40B
pAA	scaffold16_orf0063	serine protease eatA		Shigella dysenteriae 1617
pAA	scaffold16_orf0064	protease IgA1		Escherichia coli
pAA	scaffold16_orf0065	hypothetical protein E4_23001		Escherichia sp. 4_1_40B
pAA	scaffold16_orf0066	Serine protease sat precursor (Secreted autotransporter toxin sat) (fragment)		Escherichia coli 55989
pAA	scaffold16_orf0067	ISPsy2, transposase		Escherichia coli MS 185-1
pAA	scaffold16_orf0069	14 kDa aggregative adherence fimbriae I protein (Fragment) (modular protein)		Escherichia coli 55989
pAA	scaffold16_orf0070	putative transposase domain protein		Escherichia coli 3431
pAA	scaffold16_orf0071	Serine protease sepA precursor (fragment)		Escherichia sp. 4_1_40B
pAA	scaffold16_orf0072	IS186 transposase		Escherichia coli UMNK88
pAA	scaffold16_orf0073	CvaB, IS186 transposase		Escherichia coli BW2952
pAA	scaffold16_orf0074	hypothetical protein		Escherichia coli
pAA	scaffold16_orf0075	hypothetical protein Mtub2_09757		Mycobacterium tuberculosis 210
pAA	scaffold16_orf0076	putative IS639 ORF1		Escherichia coli ETEC 1392/75
pAA	scaffold16_orf0077	putative transcriptional activator aggR (AAF-III) regulatory protein)		Escherichia coli 55989
pAA	scaffold16_orf0078	transposase ORF A, IS1		Escherichia coli 55989



pAA	scaffold16_orf0079	transposase		Escherichia coli M863
pAA	scaffold16_orf0080	hypothetical protein Mtub2_09757		Mycobacterium tuberculosis 210
pAA	scaffold16_orf0081	hypothetical protein E4_23056		Escherichia sp. 4_1_40B
pAA	scaffold16_orf0083	putative transposase (fragment)		Escherichia coli 55989
pAA	scaffold16_orf0084	putative Isopentenyl-diphosphate delta-isomerase (IPP isomerase) (Isopentenyl pyrophosphate isomerase) (IPP:DMAPP isomerase)		Escherichia coli 55989
pAA	scaffold16_orf0085	hypothetical protein pEC55989_0080		Escherichia coli 55989
pAA	scaffold16_orf0086	conserved hypothetical protein		Escherichia coli MS 119-7
pAA	scaffold16_orf0087	transposase		Escherichia coli M863
pAA	scaffold16_orf0088	putative transposase insK for insertion sequence element IS150		Shigella sonnei 53G
pAA	scaffold16_orf0089	putative protein encoded within IS		Shigella sonnei Ss046