Supplementary Figures

An ancient family of mobile genomic islands introducing cephalosporinase and carbapenemase genes in *Enterobacteriaceae*

Suruchi Nepal[1], Florian Bonn[2,†], Stefano Grasso[1], Tim Stobernack[1], Anne de Jong[3], Kai Zhou[1,4], Ronald Wedema[1], Sigrid Rosema[1], Dörte Becher[2], Andreas Otto[2], John W. Rossen[1], Jan Maarten van Dijl[1*#], and Erik Bathoorn[1*#]

[1]University of Groningen, University Medical Center Groningen, Department of Medical Microbiology, Hanzeplein 1, 9700 RB Groningen, The Netherlands. E-mail: s.nepal@umcg.nl; s.grasso@umcg.nl; t.stobernack01@umcg.nl; r.wedema@pl.hanze.nl; s.rosema@umcg.nl; j.w.a.rossen@rug.nl; j.m.van.dijl01@umcg.nl; d.bathoorn@umcg.nl.

[2]Institute for Microbiology, Ernst-Moritz-Arndt-University Greifswald, Greifswald, Germany. E-mail: bonn@med.uni-frankfurt.de; dbecher@uni-greifswald.de; andreas.otto@uni-greifswald.de;

[3]Department of Molecular Genetics, University of Groningen, Groningen Biomolecular Sciences and Biotechnology Institute, 9747 AG Groningen, the Netherlands. E-mail: anne.de.jong@rug.nl.

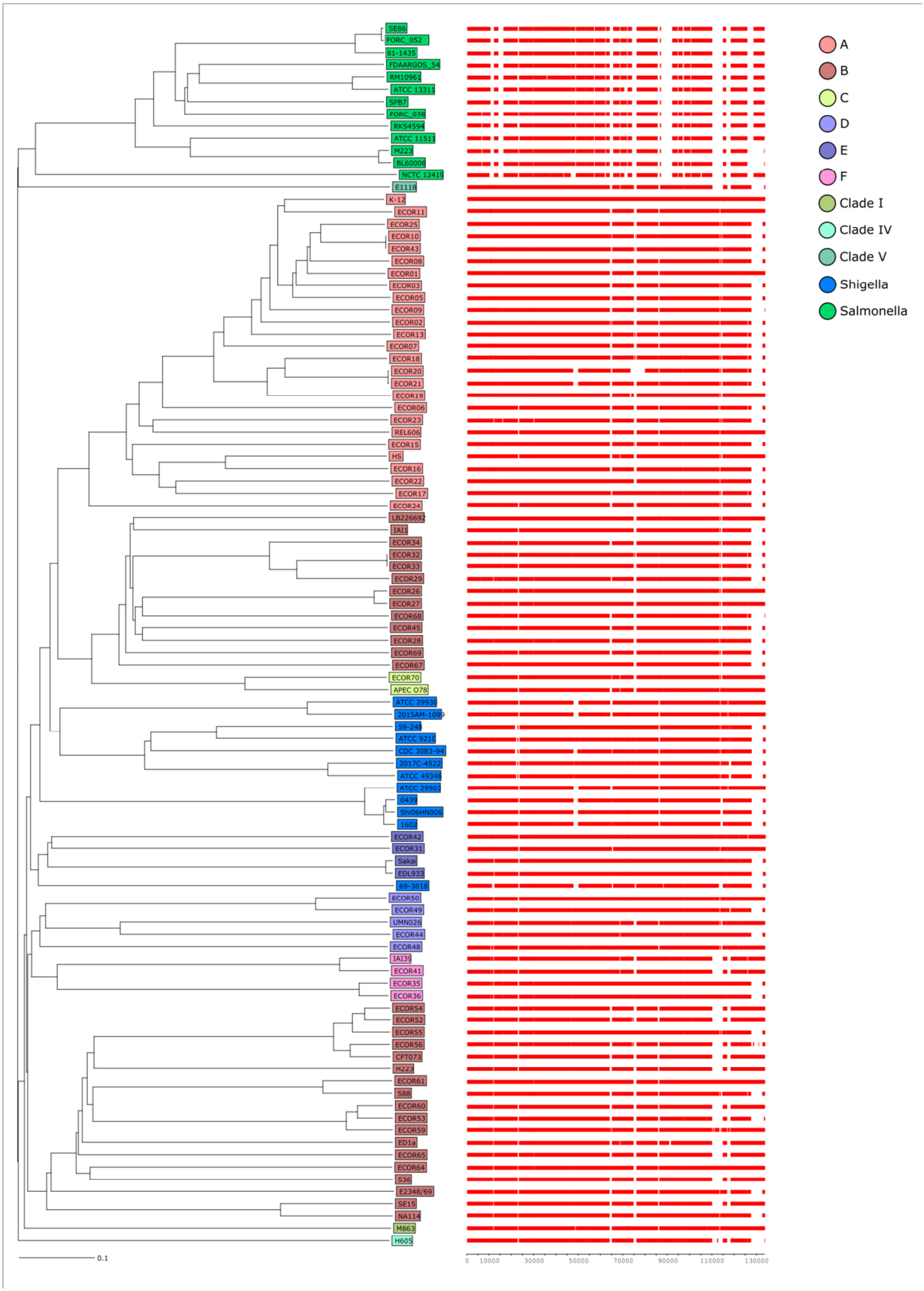
[4]State Key Laboratory for Diagnosis and Treatment of Infectious Disease, The First Affiliated Hospital, Zhejiang University, Hangzhou, China. E-mail: dr.kaizhou@qq.com.

[†]Present Address: Institute of Biochemistry 2, Goethe University Medical School, Frankfurt, Germany.

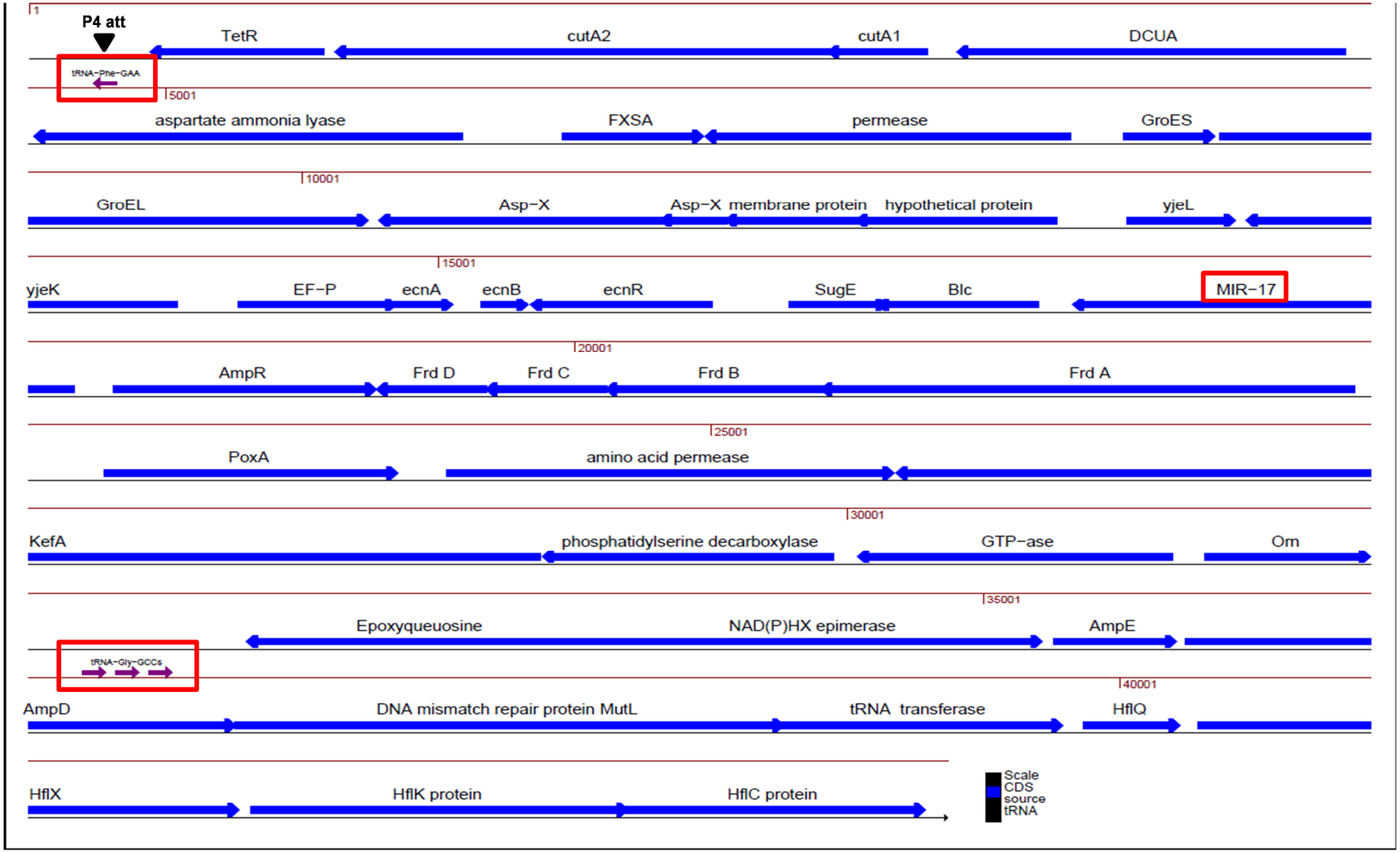
[*]These authors contributed equally to this work.

[#]Corresponding Authors.

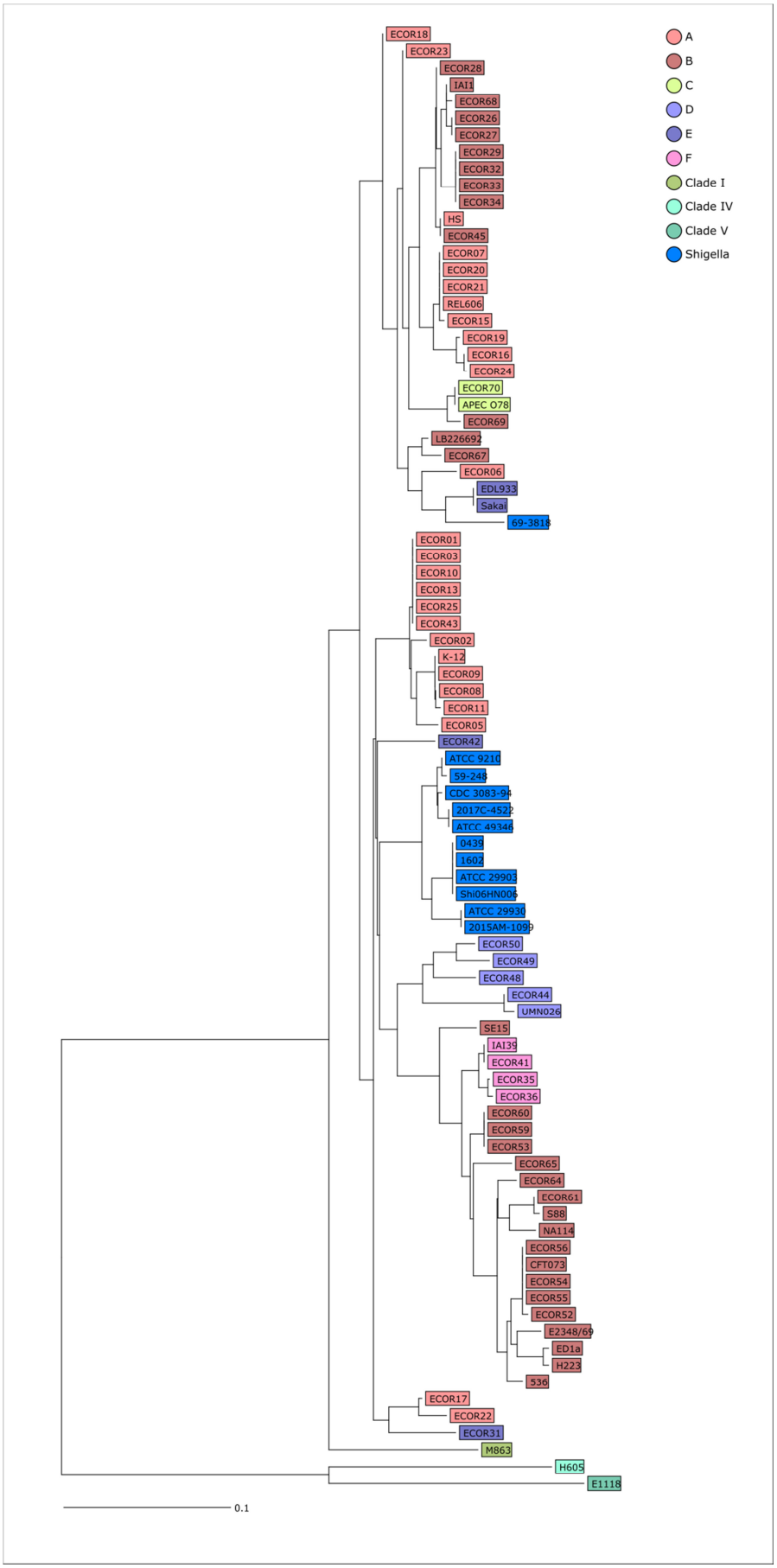
Running title: Cephalosporinase islands in *Enterobacteriaceae*


SE86
FORC_052
81-1435
FDAARGOS_54
RM10963
ATCC 13311
SPB7
FORC_036
RKS4594
ATCC 11511
M223
BL60006
NCTC 12419
E1118
K-12
ECOR11
ECOR25
ECOR10
ECOR43
ECOR08
ECOR01
ECOR03
ECOR05
ECOR09
ECOR02
ECOR13
ECOR07
ECOR18
ECOR20
ECOR21
ECOR19
ECOR06
ECOR23
REL606
ECOR15
HS
ECOR16
ECOR22
ECOR17
ECOR24
LB226692
IAI1
ECOR34
ECOR32
ECOR33
ECOR29
ECOR26
ECOR27
ECOR68
ECOR45
ECOR28
ECOR69
ECOR67
ECOR70
APEC O78
ATCC 29936
2015AM-1099
59-249
ATCC 9210
CDC 3083-94
2017C-4527
ATCC 49346
ATCC 29903
0439
Shi06HN006
1602
ECOR42
ECOR31
Sakai
EDL933
69-3818
ECOR50
ECOR49
UMN026
ECOR44
ECOR48
IAI39
ECOR41
ECOR35
ECOR36
ECOR54
ECOR52
ECOR55
ECOR56
CFT073
H223
ECOR61
S88
ECOR60
ECOR53
ECOR59
ED1a
ECOR65
ECOR64
536
E2348/69
SE15
NA114
M863
H605
0.1
0
10000
30000
50000
70000
90000
110000
130000
A
B
C
D
E
F
Clade I
Clade IV
Clade V
Shigella
Salmonella

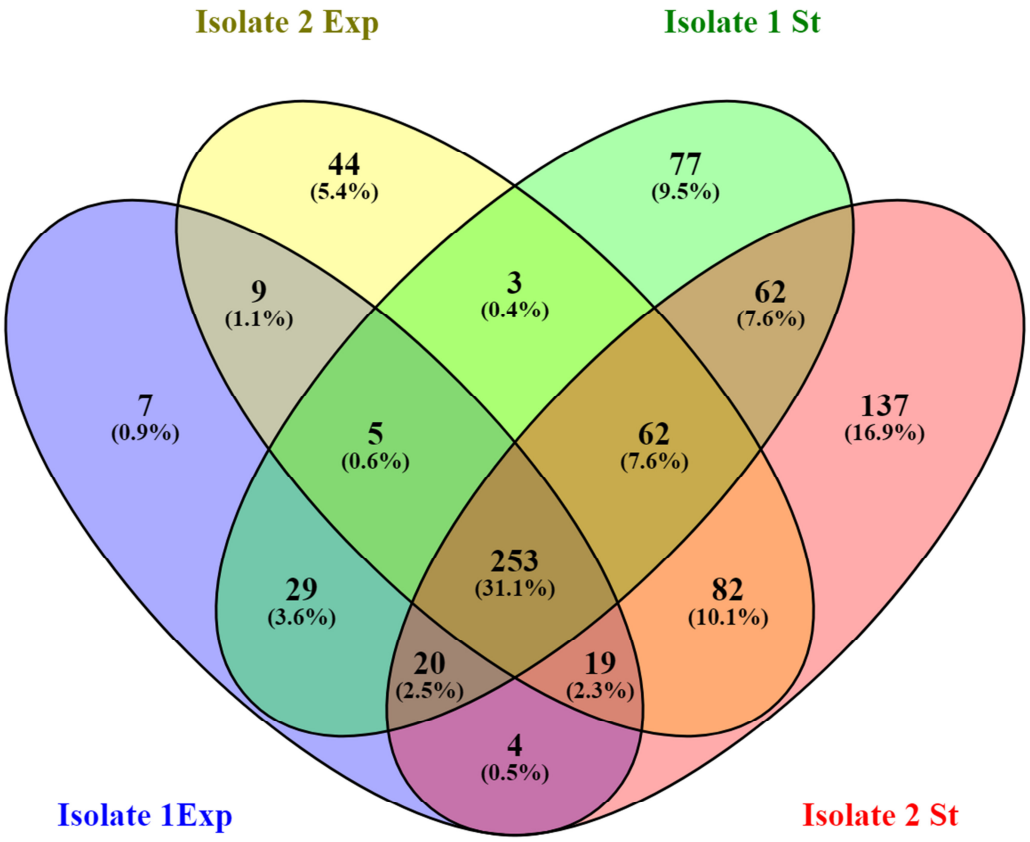
Supplemental Figure S1. Phylogenetic tree of *E. coli*, *Salmonella* and *Shigella*, and conservation of the MIR17-GI mobile genomic island. The phylogenetic tree of *E. coli* isolates, and *Salmonella* and *Shigella* species was created with the Ridom SeqSphere+ software v4.1.9 using a neighbor joining algorithm. It is built on a cgMLST analysis of 3839 targets from the reference genome of the *E. coli* K12 strain MG1655. The scale bar under the tree represents the phylogenetic distance (in %). The red bars indicate the conservation of the MIR17-GI using the respective MGI of *E. coli* K12 strain MG1655 as the reference. For the MGI comparisons, an 80% DNA similarity cut-off was used in DNA plotter. The scale bar under the MGI alignment indicates the sequence position of the reference MGI. The color codes mark publicly available sequences of *E. coli* isolates, and *Salmonella* and *Shigella* species that belong to particular phylogenetic groups or clades. The MIR17-GI was represented in all 100 investigated genomes, most of which were derived from a study by Clermont *et al*. (*Microbiology* 2015, 161: 980–988). Of note, the *ampC* gene was absent from the MGI in all sequenced *Salmonella* genomes.



Supplemental Figure S2. Genetic map of the 5' end of the MIR17-GI mobile genomic island of study isolate 1 up to the high frequency lysogeny operon. MIR17-GI is integrated in the Phe-GAA tRNA gene at the 5'end, which has a P4 prophage-associated attachment site in the 3'-5'direction. The positions of genes for AmpR, AmpD and AmpE orthologues and other identified genes are indicated.


A
B
C
D
E
F
Clade I
Clade IV
Clade V
Shigella
ECOR18
ECOR23
ECOR28
IAI1
ECOR68
ECOR26
ECOR27
ECOR29
ECOR32
ECOR33
ECOR34
HS
ECOR45
ECOR07
ECOR20
ECOR21
REL606
ECOR15
ECOR19
ECOR16
ECOR24
ECOR70
APEC O78
ECOR69
LB226692
ECOR67
ECOR06
EDL933
Sakai
69-3818
ECOR01
ECOR03
ECOR10
ECOR13
ECOR25
ECOR43
ECOR02
K-12
ECOR09
ECOR08
ECOR11
ECOR05
ECOR42
ATCC 9210
59-248
CDC 3083-94
2017C-4522
ATCC 49346
0439
1602
ATCC 29903
Shi06HN005
ATCC 29930
2015AM-1099
ECOR50
ECOR49
ECOR48
ECOR44
UMN026
SE15
IAI39
ECOR41
ECOR35
ECOR36
ECOR60
ECOR59
ECOR53
ECOR65
ECOR64
ECOR61
S88
NA114
ECOR56
CFT073
ECOR54
ECOR55
ECOR52
E2348/69
ED1a
H223
536
ECOR17
ECOR22
ECOR31
M863
H605
E1118
0.1

Supplemental Figure S3. Neighbor joining tree showing the diversity of *ampC* genes on MIR17-GI-like MGIs in *E. coli* and *Shigella*. The neighbor joining tree is based on genome sequences used to construct the phylogenetic tree in Supplemental Figure S1. The scale bar represents the % difference among 197 nucleotides of *ampC*, using the *ampC* gene of the *E. coli* K12 strain MG1655 as a reference.



Supplemental Figure S4. Numbers of expressed proteins in *E. cloacae* isolates 1 and 2.
The Venn diagram shows overlapping and uniquely expressed proteins of the *E. cloacae* isolates 1 and 2 in the exponential (Exp) and stationary (St) growth phases. The diagram was created using the Venn diagram web tool of the VIB and the University of Gent in Belgium (http://bioinformatics.psb.ugent.be/webtools/Venn/).