Supplemental Material

Supplemental material and methods

Whole genome sequencing

For next generation sequencing with the Illumina MiSeq (Illumina Netherlands BV, Eindhoven, The Netherlands), whole genome DNA was isolated using the Purelink® Genome DNA Mini kit (Thermo Fischer Scientific, Dreieich, Germany). An Illumina Nextera XT library (Illumina Netherlands BV, Eindhoven, The Netherlands) was produced and sequencing was performed with 2x300 bp paired-end reads. For the Illumina sequencing, 4,597,732 paired-end reads with a mean read length of 175 nt (mean coverage 137x) were obtained. For long-read single-molecule real-time (SMRT) sequencing (Pacific Biosciences, MenloPark, CA), DNA was isolated using the method described by Pitcher et al (1). A SMRTbell™ template library was prepared according to the instructions from Pacific Biosciences, Menlo Park, CA, USA, following the Procedure & Checklist - 10 kb Template Preparation Using BluePippin™ Size-Selection System. Briefly, for the preparation of 15kb libraries 8µg genomic DNA was sheared using g-tubes™ from Covaris, Woburn, MA, USA according to the manufacturer´s instructions. DNA was end-repaired and ligated overnight to hairpin adapters applying components from the DNA/Polymerase Binding Kit P6 from Pacific Biosciences, Menlo Park, CA, USA. Reactions were carried out according to the manufacturer´s instructions. BluePippin™ Size-Selection to 4000 kb and 15000 kb was performed according to the manufacturer´s instructions (Sage Science, Beverly, MA, USA). Conditions for annealing of sequencing primers and binding of polymerase to purified SMRTbell™ template were assessed with the Calculator in RS Remote, Pacific Biosciences, Menlo Park, CA, USA. SMRT sequencing was carried out on the PacBio *RSII* (Pacific Biosciences, Menlo Park, CA, USA) taking one 240-minutes movie for each SMRT cell. In total 3 SMRT cells were run. A total of 198,383 reads with a mean read length of 9,430 nt were generated. Assembly of PacBio reads was performed using

RS_HGAP_Assembly.3 included in the SMRT Portal 2.3.0. To obtain a high quality genome sequence, mapping of paired-end reads from Illumina sequencing was performed using Burrows–Wheeler Aligner [BWA, (2)].

In-silico analyses

The multi-locus sequence type of *E. coli* NRZ14408 was identified using MLST 1.8 implementing the scheme of Wirth et al. (3, 4). Serotypes were identified using SeroTypeFinder (9). The insertion sequences were investigated using ISFinder (10) and blastn (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Bacteriophages were analyzed using PHAST (11). Plasmidfinder was used to investigate the incompatibility groups of the plasmids (12). Database search for *mcr-1* encoding IncHI2 plasmids was performed using blastn with p14408_M as a reference (http://blast.ncbi.nlm.nih.gov/Blast.cgi; date of database search: 15th August 2016). The presence of acquired antibiotic resistance genes and virulence genes was assessed using ResFinder and VirulenceFinder (7, 8). Analysis of operons involved in the iron acquisition was performed using blastn (http://blast.ncbi.nlm.nih.gov/Blast.cgi) with the sequences presented in supplementary table S5. For comparative analyses, EasyFig and BRIG were used (5, 6).

E. coli of sequence type ST362 were identified using Enterobase (https://enterobase.warwick.ac.uk/). The contig sequences of all isolates for which the isolation date was known, were downloaded. In addition, the data from NCTC 3000 Project (Public Health England, https://www.phe-culturecollections.org.uk/collections/nctc-3000-project.aspx) was analyzed using the MLST package provided on https://github.com/tseemann/mlst for the presence of *E. coli* ST362. One *E. coli* ST362 was identified (*E. coli* NCTC9014; accession number ERS701895; ftp://ftp.sanger.ac.uk/pub/project/pathogens/NCTC3000/datalinks_manual/ERS701895.gff).

Supplementary Tables

Supplementary Table S1: Properties of the plasmids and phages. [a] not 100% identity on DNA sequence level, [b]two copies are present on different plasmids, [c]two copies present on the same plasmid, n.a. not applicable (no plasmid), n.t.: not typeable using conventional methods (no incompatibility group), MLS= Macrolide, Lincosamide and Streptogramin B. Genes which are marked red are present in two copies on different plasmids.

	Name	p14408_M	p14408_1	p14408_2	phage 1	phage 2	p14408_3
	Size [bp]	238,073	137,186	88,196	41,305	19,226	11,988
	Incompatibility group	IncHI2	IncFII, IncFIB	IncN	n.a.	n.a.	n.t.
Genes encoding resistance to:	Aminoglycosides	*aac(3)-IIa*, *aadA1*, *aadA2*, *aph(3')-Ic*[a], *strA*[b], *strB*[b]	*aac(3)-IId*[b]	*aac(3)-IId*[b], *strA*[b], *strB*[b]			
	Beta-lactams	$bla_{\text{TEM-1A}}$	$bla_{\text{KPC-2}}$[b], $bla_{\text{OXA-1}}$[b], $bla_{\text{TEM-1B}}$[b]	$bla_{\text{KPC-2}}$[b], $bla_{\text{OXA-1}}$[b], $bla_{\text{TEM-1B}}$[b]			
	Colistin	*mcr-1*					
	Fluoroquinolones		*aac(6')Ib-cr*[b]	*aac(6')Ib-cr*[b], *qnrB2*			
	MLS		*mph(A)*[b]	*mph(A)*[b]			
	Phenicols	*cmlA1*	*catA1*, *catB3*[b]	*catB3*[b]			
	Rifampicin		*arr-3*[b]	*arr-3*[b]			
	Sulphonamides	*sul3*	*sul1*[b]	*sul1*[b,c]			
	Tetracyclines	*tet(A)*					
	Trimethoprim			*dfrA18*			

Supplementary Table S2: Distribution of antibiotic resistance genes among the plasmids p14408_M, pHNSHP45-2, pSA26-MCR-1 and pS38. All genes depict an identity of >95% and >98% coverage, unless stated differently. [a]1x 100% coverage, second truncated (75%) MLS= Macrolide, Lincosamide and Streptogramin B

	Aminoglycoside										**Beta-lactam**				**Colistin**	**Fluoroquinolone**		**Fosfomycin**	**MLS**		**Phenicol**			**Sulphonamide**			**Tetracycline**	**Trimethoprim**		
plasmid	*aadA2*	*aadA1a*	*aadA1b*	*strA*	*strB*	*aph(3')-Ia*	*aph(3')-Ic*	*aac(3)-IIa*	*aph(4)-Ia*	*aac(3)-IVa*	*bla*$_{TEM-1B}$	*bla*$_{TEM-1A}$	*bla*$_{CTX-M-1}$	*bla*$_{CTX-M-14}$	*mcr-1*	*oqxB*	*oqxA*	*fosA*	*mph(A)*	*mef(B)*	*cmlA1*	*floR*	*cml*	*sul3*	*sul1*	*sul2*	*tet(A)*	*dfrA1*	dfrA12	dfrA14
p14408_M1	+	+		+	+		+	+				+			+						+			+		+	+		+	
pSA26-MCR-1	+	+		+	+	+					+				+				+			+	+	+		+	+			+
pHNSHP45-2	+[a]	+				+			+	+				+	+	+	+	+			+	+		+	+	+			+	
pS38	+	+	+										+							+				+	+		+	+		

Supplementary Table S3: Virulence genes present in the chromosome of *E. coli* NRZ14408. [a]*astA* was present in twelve copies, [b] present in two copies. n.a.: not applicable. The iron metabolism operons were identified using blastn and the references presented in Supplementary Table S3, the virulence genes were assessed using VirulenceFinder (8).

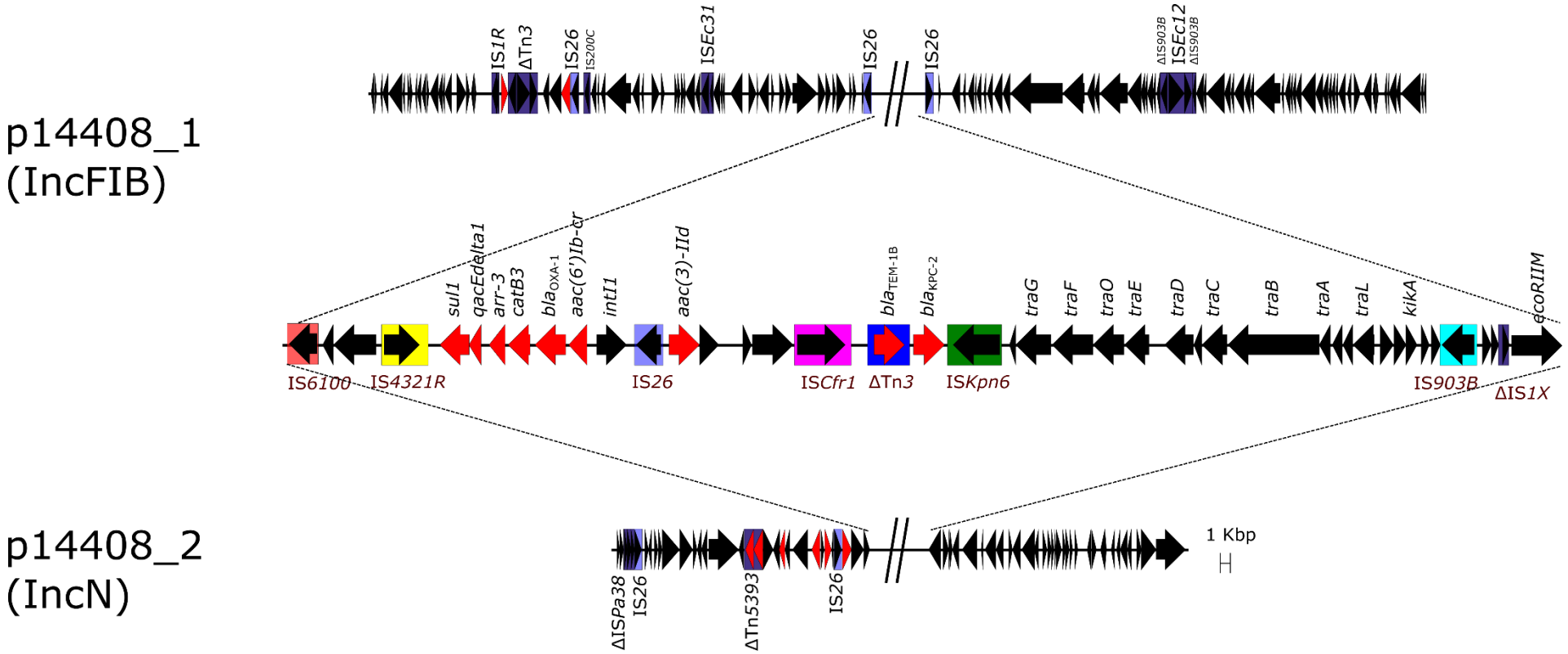
Genes	Operon	Function	category	NRZ14408
astA	n.a.	Toxin	not known	+[a]
iha	n.a.	Adherence	Adherence	+
gad	n.a.	Acid resistance	Acid resistance	+[b]
air, eilA	*selC*-B pathogenicity island	Adherence	Adherence	+
iss	n.a.	Serum resistance	Serum resistance	+
chuA, chuS, chuT, chuU, chuW, chuX, chuY	chu	Hemin uptake system	Iron metabolism	+
iucA, iucB, iucC, iucD, iutA	aerobactin	Aerobactin	Iron metabolism	+
fhuA, fhuB, fhuC, fhuD	fhu	ferrichrome uptake DH10B	Iron metabolism	+
feoA, feoB, feoC	feo	ferrous iron uptake	Iron metabolism	+

Supplementary Table S4: Properties of additional *E. coli* ST362 included into the study. AB, antibiotic

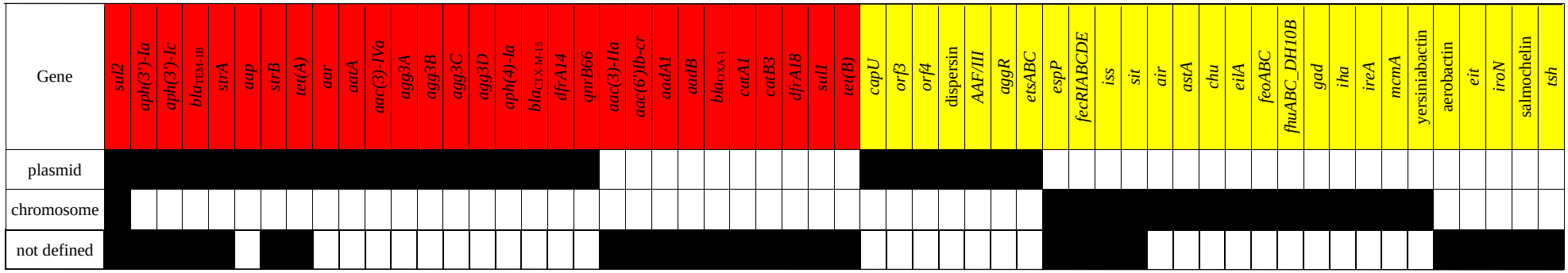
Isolate	Accession number	AB-resistance genes?	Source	Collection Year	Country	Serotype	No. of virulence genes	Beta-Lactamase Genes
NCTC9014	ERS701895	no	Human	1952	Denmark	:H6	6	None
MOD1-EC6899	SRS1529503	yes	Livestock	1975	unknown	O7:H6	10	No beta-lactamase
MOD1-EC5512	SRS1594124	yes	Livestock	1976	United States	O7:H6	10	$bla_{\text{TEM-1B}}$
MOD1-EC5501	SRS1594118	yes	Human	1992	India	O7:H6	11	$bla_{\text{OXA-1}}$
MOD1-EC5465	SRS1594016	no	Poultry	1993	United States	O7:H6	6	None
MOD1-EC6950	SRS1590224	yes	Livestock	1997	United States	O15:H1	9	$bla_{\text{TEM-1B}}$
HVH 6 (3-8296502)	SRS402207	yes	Human	2004	Denmark	O86:H28	12	$bla_{\text{TEM-1B}}$
HICF113	SRS1221335	yes	Human	2010	United Kingdom	O7:H6	5	$bla_{\text{CTX-M-15}}$, $bla_{\text{OXA-1}}$, $bla_{\text{TEM-1B}}$
KTE78	SRS410376	no	Human	2010	Denmark	O7:H6	4	None
KTE79	SRS410371	no	Human	2010	Denmark	O7:H6	4	None
AZ_TG77018	SRS876455	yes	Poultry	2014	United States	O7:H6	13	No beta-lactamase
AZ_TG77102	SRS893459	yes	Poultry	2014	United States	O15:H1	14	$bla_{\text{TEM-1B}}$
sam_103239	ERS1164728	yes	Livestock	2014	Denmark	O11:H9	9	$bla_{\text{CTX-M-1}}$, $bla_{\text{TEM-1B}}$

Supplementary Table S5: Iron acquisition operons investigated in the study

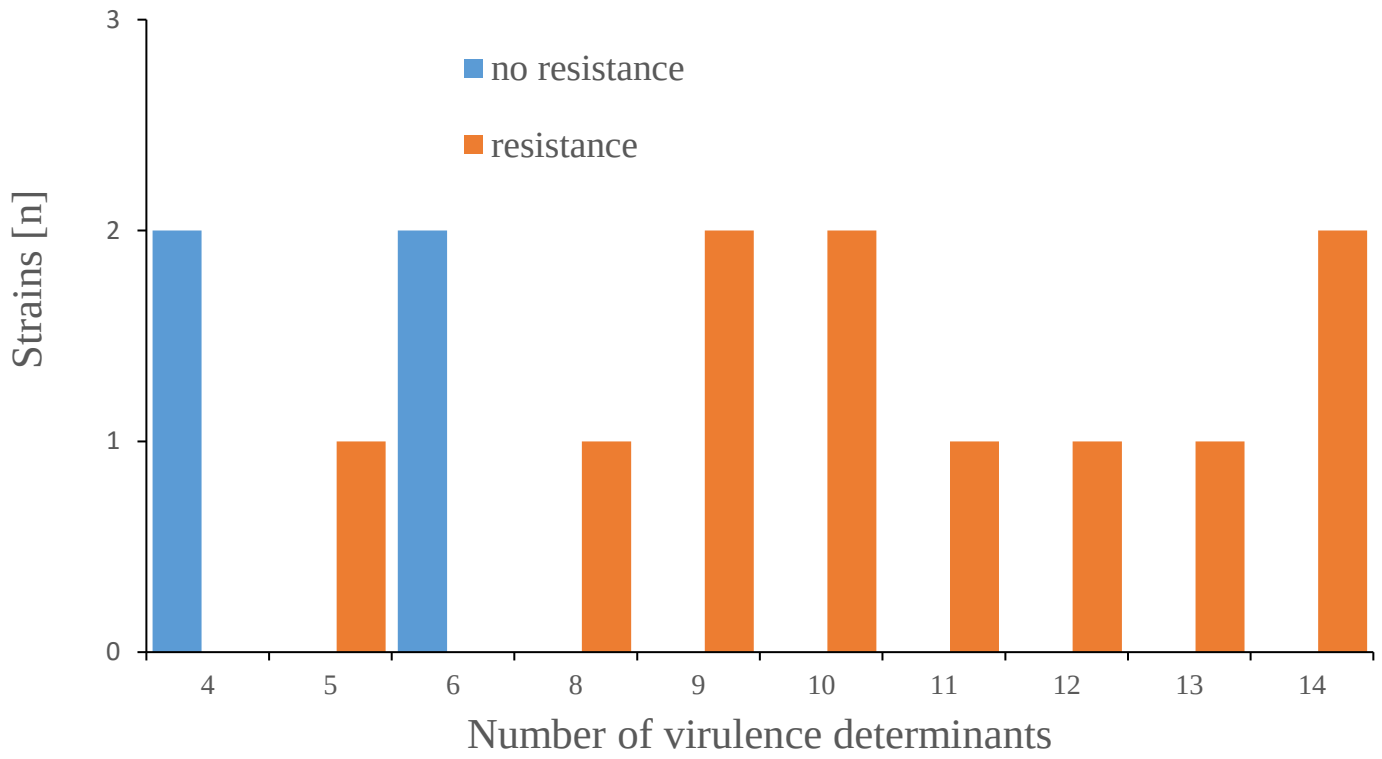
Operon	function	Source	Database accession number	position	length
fhu	ferrichrome uptake	*Escherichia coli* str. K12 substr. DH10B	CP000948.1	141588-147548	5961 bp
fhu	ferrichrome uptake	*Escherichia coli* APEC O78	CP004009.1	869099-875005	5907 bp
feo	Ferrous iron uptake	*Escherichia coli* CFT073	AE014075.1	3967027-3969592	2566 bp
chu	heme iron-utilization	*Escherichia coli* CFT073	AE014075.1	4084996-4093181	8186 bp
aerobactin	iron acquisition	Plasmid pAPEC-O2-ColV	NC_007675.1	98089-106117	8029 bp
salmochelin	iron acquisition	Plasmid pAPEC-O2-ColV	NC_007675.1	141336-150894	9559 bp
yersiniabactin	iron acquisition	*Yersinia pestis* CO92 chromosome	NC_003143.1	2140840-2169669	28830 bp
sit	iron ABC transport system (Fe2+)	Plasmid pAPEC-O1-ColBM	DQ381420.1	152749-156198	3450bp
fec	iron(III) dicitrate ABC transporter	Plasmid pKPN_CZ	NC_019390.1	22157-29701	7545 bp
eit	putative ABC iron transport system	Plasmid pAPEC-O2-ColV	NC_007675.1	176416-180456	4041 bp
ets	putative ABC iron transport system	Plasmid pAPEC-O2-ColV	NC_007675.1	119488-123176	3689 bp
no name	putative ABC iron transport system	Plasmid pEC14_114	GQ398086.1	6106-13149	7044 bp



Supplementary Figure S1: Comparison of the *bla*$_{KPC-2}$ encoding plasmids of *E. coli* NRZ14408.



Supplementary Figure S2: Depiction of the location of antibiotic resistance genes (red) and virulence genes (yellow) in the investigated ST362 isolates. Black = presence of genes, white = absence of genes.



Supplementary Figure S3: Number of virulence determinants per isolate subdivided according to isolates harboring antibiotic resistance genes and those not harboring those.

Supplementary References

1. **Pitcher DG**, **Saunders NA**, **Owen RJ**. 1989. Rapid extraction of bacterial genomic DNA with guanidium thiocyanate. Lett Appl Microbiol **8**:151–156.

2. **Li H**, **Durbin R**. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics **25**:1754–1760.

3. **Wirth T**, **Falush D**, **Lan R**, **Colles F**, **Mensa P**, **Wieler LH**, **Karch H**, **Reeves PR**, **Maiden MCJ**, **Ochman H**, **Achtman M**. 2006. Sex and virulence in *Escherichia coli*: an evolutionary perspective. Mol Microbiol **60**:1136–51.

4. **Larsen M V.**, **Cosentino S**, **Rasmussen S**, **Friis C**, **Hasman H**, **Marvig RL**, **Jelsbak L**, **Sicheritz-Pontén T**, **Ussery DW**, **Aarestrup FM**, **Lund O**. 2012. Multilocus sequence typing of total-genome-sequenced bacteria. J Clin Microbiol **50**:1355–1361.

5. **Sullivan MJ**, **Petty NK**, **Beatson SA**. 2011. Easyfig: a genome comparison visualizer. Bioinformatics **27**:1009–10.

6. **Alikhan N-F**, **Petty NK**, **Ben Zakour NL**, **Beatson SA**. 2011. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. BMC Genomics **12**:402.

7. **Zankari E**, **Hasman H**, **Cosentino S**, **Vestergaard M**, **Rasmussen S**, **Lund O**, **Aarestrup FM**, **Larsen MV**. 2012. Identification of acquired antimicrobial resistance genes. J Antimicrob Chemother **67**:2640–2644.

8. **Joensen KG**, **Scheutz F**, **Lund O**, **Hasman H**, **Kaas RS**, **Nielsen EM**, **Aarestrup FM**. 2014. Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic *Escherichia coli*. J Clin Microbiol **52**:1501–10.

9. **Joensen KG**, **Tetzschner AMM**, **Iguchi A**, **Aarestrup FM**, **Scheutz F**. 2015. Rapid and easy in silico serotyping of *Escherichia coli* isolates by use of whole-genome sequencing data. J Clin Microbiol **53**:2410–26.

10. **Siguier P**, **Perochon J**, **Lestrade L**, **Mahillon J**, **Chandler M**. 2006. ISfinder: the reference centre for bacterial insertion sequences. Nucleic Acids Res **34**:D32-6.

11. **Zhou Y**, **Liang Y**, **Lynch KH**, **Dennis JJ**, **Wishart DS**. 2011. PHAST: a fast phage search tool. Nucleic Acids Res **39**:W347-52.

12. **Carattoli A**, **Zankari E**, **García-Fernández A**, **Voldby Larsen M**, **Lund O**, **Villa L**, **Møller Aarestrup F**, **Hasman H**. 2014. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. Antimicrob Agents Chemother **58**:3895–903.