

National Institute for Public Health
and the Environment
Ministry of Health, Welfare and Sport

NGS and TGS analysis of STEC O80:H2 and O45:H2 ST301 isolates

Angela van Hoek



Introduction

- Characterisation of 53 STEC strains with ExPEC-associated virulence genes (It; 30, NL; 23)
- Whole-Genome Sequencing using Ion Torrent (It) and Illumina (NL) platforms
- ST301 (41), O80:H2 (33), O45:H2 (4), *stx*_{2a} (34), *stx*_{2d} (6), *stx*_{2f} (4), *eae*-xi variant (41)
- Harbor mosaic pR444_A like plasmids, which combines several virulence genes (*iucC* (aerobactin), *iroN* (salmochelins), *iss* (serum resistance protein), *etsABC* (putative secretion system I), *ompT* (omptin), *hlyF* (hemolysin), *cia* and *cva* (bacteriocins) as well as antimicrobial resistance genes

11-10-2022

Genomic Characterization of *hlyF*-positive Shiga Toxin-Producing *Escherichia coli*, Italy and the Netherlands, 2000–2019

Federica Gigliucci, Angela H.A.M. van Hoek, Paola Chiani, Arnold Knijn, Fabio Minelli, Gaia Scavia, Eelco Franz, Stefano Morabito, Valeria Michelacci

Shiga toxin-producing *Escherichia coli* (STEC) O80:H2 has emerged in Europe as a cause of hemolytic uremic syndrome associated with bacteremia. STEC O80:H2 harbors the mosaic plasmid pR444_A, which combines several virulence genes, including *hlyF* and antimicrobial resistance genes. pR444_A is found in some extraintestinal pathogenic *E. coli* (ExPEC) strains. We identified and characterized 53 STEC strains with ExPEC-associated virulence genes isolated in Italy and the Netherlands during 2000–2019. The isolates belong to 2 major populations: 1 belongs to sequence type 301 and harbors diverse *stx*₂ subtypes, the intimin variant *eae-ξ*, and pO157-like and pR444_A plasmids; 1 consists of strains belonging to various sequence types, some of which lack the pO157 plasmid, the locus of enterocyte effacement, and the antimicrobial resistance-encoding region. Our results showed that STEC strains harboring ExPEC-associated virulence genes can include multiple serotypes and that the pR444_A plasmid can be acquired and mobilized by STEC strains.

Shiga toxin-producing *Escherichia coli* (STEC) is a group of enteric pathogens that cause foodborne disease ranging from uncomplicated diarrhea to hemorrhagic colitis (HC) or hemolytic uremic syndrome (HUS) (1). The most serious complication of STEC infection is HUS, which can be fatal.

STEC strains produce Shiga toxins (Stx), a family composed of 2 main types of cytotoxins: Stx1 and Stx2 (2). Stx1 is classified into subtypes a, c, and d; Stx2 is classified into subtypes a–k (3–6). Lysogenic bacteriophages harbor the genes for different types of Stx;

Author affiliations: Istituto Superiore di Sanità, Rome, Italy (F. Gigliucci, P. Chiani, A. Knijn, F. Minelli, G. Scavia, S. Morabito, V. Michelacci); National Institute for Public Health and the Environment, Bilthoven, the Netherlands (A.H.A.M. van Hoek, E. Franz)

DOI: <https://doi.org/10.3201/eid2703.203110>

infected bacteria then produce the protein (7). Although the production of Stx plays a central role in the pathogenesis of STEC-associated illness, the development of HC and HUS requires an efficient host colonization by the infecting STEC. Many HUS-associated STEC strains possess a chromosomal pathogenicity island, defined as the locus of enterocyte effacement (LEE), which is associated with the attaching and effacing lesions (8) described in enteropathogenic *E. coli* (9), or possess the genetic machinery conferring the enteroaggregative pattern of adhesion to the enterocyte described in enteroaggregative *E. coli* (10,11). Other STEC strains harbor colonization factors of enterotoxigenic *E. coli* (12,13) and genes encoding virulence features associated with extraintestinal pathogenic *E. coli* (ExPEC) (14,15). The ExPEC-associated virulence genes code for aerobactin (encoded by *iucC*), salmochelins (*iroN*), serum resistance protein (*iss*), a putative secretion system I (*etsABC*), omptin (*ompT*), hemolysin (*hlyF*), and bacteriocins (*cia* and *cva*) (14,15).

STEC strains belonging to the O157, O26, O103, O111, and O145 serogroups are considered critical public health concerns. Nevertheless, since 2015, other STEC serogroups have been increasingly associated with HUS and other infections in humans (16). Among these, O80 is emerging in Europe (17–21); since 2015, it has become a predominant serogroup associated with HUS in children in France (22). In addition to the typical clinical features of a STEC infection, bacteremia can develop in patients with STEC O80 (19,23).

Virulence genes increase the pathogenicity of STEC strains. For example, the strains associated with HUS are characterized by specific subtypes of the *stx*₂ gene, mainly *stx*_{2a}, *stx*_{2c}, and *stx*_{2d} (24). In 2020, experts proposed a new approach to categorizing STEC



Methods

- Nine ST301 STEC strains
 - O80:H2;
NL1700800, NL1700844, NL1800064, PT28_1, PT28_2, PT28_3, PT28_4
 - O45:H2;
NL1800418, NL1800740
- Long-read third-generation sequencing (TGS) using a MinION flow cell on a GridION (Nanopore)
- Demultiplexed Nanopore fastq files were concatenated, and quality filtered and trimmed (nanofilt (v2.8.0) and filtlong (v0.2.1))
- Hybrid assembly of the Nanopore TGS data together with the Illumina short-read next-generation sequencing (NGS) data using Unicycler (v0.4.8) with bold settings and pilon polishing
- Annotation using Prokka (v1.14.6) and RAST (<https://rast.nmpdr.org/>)



Results; Assembly details

NGS assemblies*

Hybrid assemblies (NGS & TGS)

Strain	# Contigs	N50	Base Count (Mb)	Max Length	Min Length	# Contigs	N50	Base Count (Mb)	Max Length	Min Length
NL1700800	260	107295	5.56	302752	514	7	4830233	5.81	4830233	16831
NL1700844	234	127121	5.50	301551	500	7	4035882	5.67	4035882	14649
NL1800064	255	120260	5.44	252848	502	7	4512517	5.62	4512517	10782
NL1800418	229	94080	5.54	247599	520	6	4129118	5.75	4129118	22447
NL1800740	215	115059	5.42	323670	508	4	5300570	5.65	5300570	96225
PT28_1	272	87768	5.45	303454	505	7	4009096	5.65	4009096	66095
PT28_2	301	70608	5.59	303453	503	56	810274	5.79	1728085	154
PT28_3	303	89882	5.44	302967	500	9	3651034	5.66	3651034	11929
PT28_4	328	74083	5.58	301863	500	8	4216161	5.70	4216161	25002

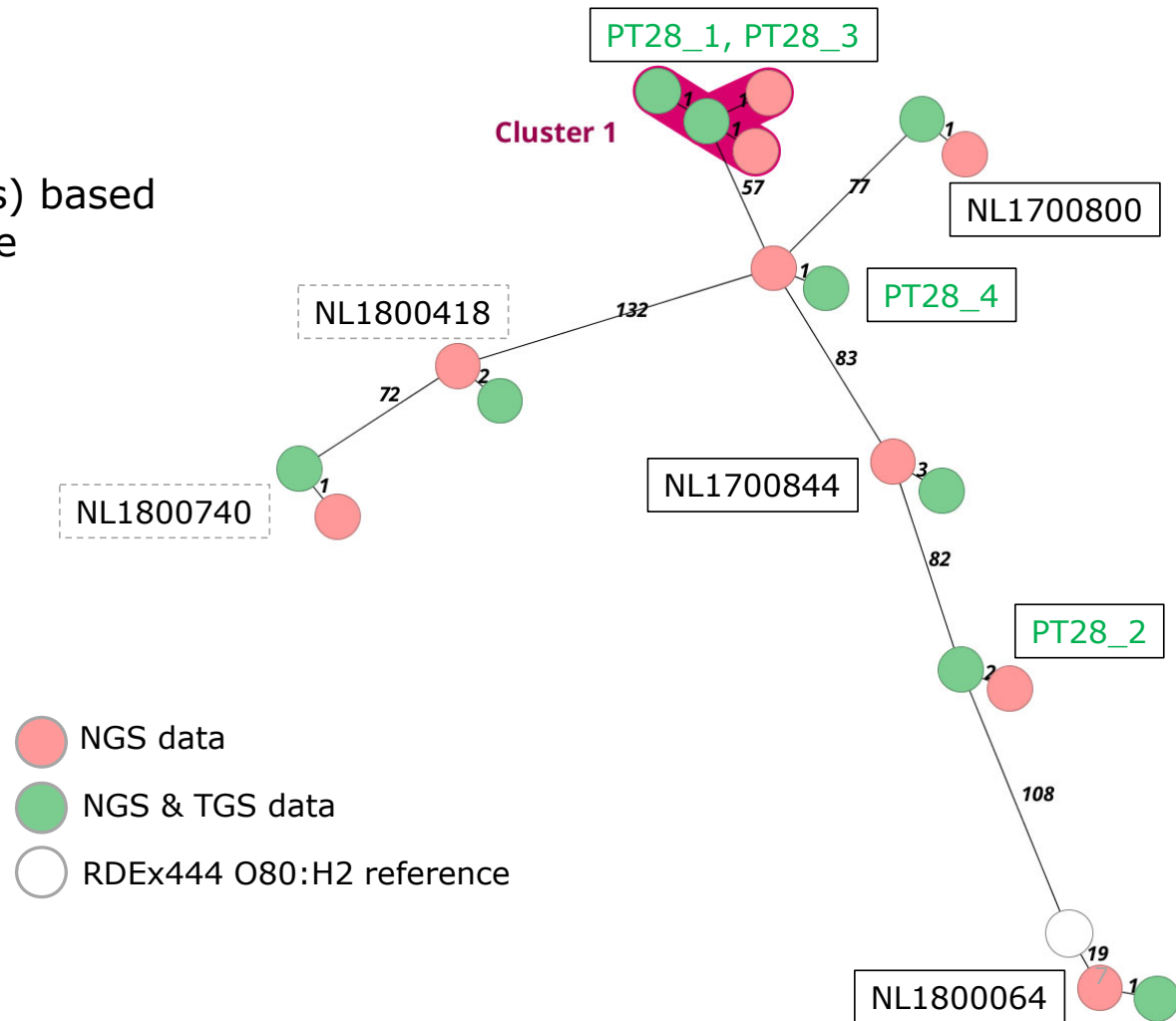
*in house assembly pipeline using SPAdes (3.15.3)



O80:H2 O45:H2

Results; MST

- cgMLST scheme (4203 gene targets) based on *E. coli* O80:H2 reference genome (pR444_A positive)





O80:H2

O45:H2

Results; Contig details

NL1700800

Contig	Length	
1	4830233	
2	336420	
3	284714	circular
4	144663	circular
5	96472	circular
6	96402	circular
7	16831	circular

NL1700844

Contig	Length	
1	4035882	
2	998309	
3	265061	
4	157864	circular
5	110008	circular
6	88484	
7	14649	

NL1800064

Contig	Length	
1	4512517	
2	730499	
3	141289	circular
4	117094	circular
5	94387	circular
6	11659	
7	10782	circular

NL1800418

Contig	Length	
1	4129118	
2	1320518	
3	103479	circular
4	96405	circular
5	73211	
6	22447	

NL1800740

Contig	Length	
1	5300570	
2	158978	circular
3	96405	circular
4	96225	circular

PT28_1

Contig	Length	
1	4009096	
2	593999	
3	521612	
4	222121	circular
5	142970	
6	96362	circular
7	66095	

PT28_3

Contig	Length	
1	3651034	
2	1099477	
3	295361	circular
4	240052	
5	214434	circular
6	95728	circular
7	30777	
8	24072	
9	11929	

PT28_4

Contig	Length	
1	4216161	
2	747581	
3	231054	
4	199507	
5	110012	circular
6	103671	
7	67228	
8	25002	

pR444_A
like plasmids

PT28_2

Contig	Length	
1	1728085	
2	1085037	
3	810274	
4	766147	
5	388571	
6	282080	
7	224251	
8	146623	
9	111221	
10	89874	
11	21330	
12	16457	
13	11527	
14	11340	etc



Located on pR444_A like contig

Results, VFs

Strain	Virulence genes
NL1700800	<i>afaABCDE, cia, cma, cvaC, eae, efa1, ehly2, ehxA, espA, espB, espF, espP, hlyF, hra, iha, iroN, iss, mchB, mchC, mchF, nleA, nleB, nleC, ompT, sitA, stx_{2a'}, terC, tir, traT</i>
NL1700844	<i>cia, cvaC, eae, efa1, ehxA, espA, espB, espF, espP, etsC, hlyF, hra, iha, iroN, iss, iucC, iutA, mchB, mchC, mchF, nleA, nleB, nleC, ompT, stx_{2a'}, terC, tir, traT</i>
NL1800064	<i>cia, cvaC, eae, efa1, ehxA, espA, espB, espF, espP, etsC, hlyF, hra, iroN, iss, iucC, iutA, mchB, mchC, mchF, nleA, nleB, nleC, ompT, stx_{2d'}, terC, tir, traT</i>
NL1800418	<i>cia, cma, cvaC, eae, efa1, espA, espB, espF, espP, hlyF, hra, iha, iroN, iss, mchB, mchC, mchF, nleA, nleB, nleC, ompT, sitA, stx_{2a'}, terC, tir</i>
NL1800740	<i>cia, cma, cvaC, eae, efa1, ehxA, espA, espB, espF, espP, hlyF, hra, iha, iroN, iss, mchB, mchC, mchF, nleA, nleB, nleC, ompT, sitA, stx_{2a'}, terC, tir, traT</i>
PT28_1	<i>cia, cma, cvaC, eae, efa1, ehxA, espA, espB, espF, espP, hlyF, hra, iha, iroN, iss, mchB, mchC, mchF, nleA, nleB, nleC, ompT, sitA, stx_{2a'}, terC, tir, traT</i>
PT28_2	<i>cma, cvaC, eae, efa1, espA, espB, espF, hlyF, hra, iha, iroN, iss, mchB, mchC, mchF, nleA, nleB, nleC, ompT, sitA, stx_{2a'}, terC, tir, traT</i>
PT28_3	<i>cia, cma, cvaC, eae, efa1, ehxA, espA, espB, espF, espP, hlyF, hra, iha, iroN, iss, mchB, mchC, mchF, nleA, nleB, nleC, ompT, sitA, stx_{2a'}, terC, tir, traT</i>
PT28_4	<i>cvaC, eae, efa1, espA, espB, espF, etsC, hlyF, iha, iroN, iss, iucC, iutA, mchB, mchC, mchF, nleA, nleB, nleC, ompT, sitA, stx_{2a'}, terC, tir, traT</i>



Located on pR444_A like contig

Results, AMR

Strain	Antimicrobial resistance genes
NL1700800	<i>aadA</i> , <i>aadB</i> , <i>aph(3')-Ia</i> , <i>catA</i> , <i>floR</i> , <i>mdfA</i> , <i>strA</i> , <i>strB</i> , <i>sul1</i> , <i>sul2</i> , <i>tet(A)</i>
NL1700844	<i>aadA</i> , <i>aadB</i> , <i>aph(3')-Ia</i> , <i>bla_{TEM-1}</i> , <i>dfrA5</i> , <i>floR</i> , <i>mdfA</i> , <i>strA</i> , <i>strB</i> , <i>sul1</i> , <i>sul2</i> , <i>tet(A)</i>
NL1800064	<i>aph(3')-Ia</i> , <i>bla_{TEM-1}</i> , <i>dfrA5</i> , <i>mdfA</i> , <i>strA</i> , <i>strB</i> , <i>sul2</i>
NL1800418	<i>aadA</i> , <i>aph(3')-Ia</i> , <i>bla_{OXA-1}</i> , <i>floR</i> , <i>mdfA</i> , <i>strA</i> , <i>strB</i> , <i>sul1</i> , <i>sul2</i> , <i>tet(A)</i>
NL1800740	<i>aadA</i> , <i>aph(3')-Ia</i> , <i>bla_{OXA-1}</i> , <i>bla_{TEM-1}</i> , <i>catA</i> , <i>floR</i> , <i>mdfA</i> , <i>strA</i> , <i>strB</i> , <i>sul2</i> , <i>tet(A)</i>
PT28_1	<i>aadB</i> , <i>aph(3')-Ia</i> , <i>floR</i> , <i>mdfA</i> , <i>strA</i> , <i>strB</i> , <i>sul1</i> , <i>sul2</i>
PT28_2	<i>aadA</i> , <i>aadB</i> , <i>aph(3')-Ia</i> , <i>bla_{TEM-1}</i> , <i>catA</i> , <i>floR</i> , <i>mdfA</i> , <i>strA</i> , <i>strB</i> , <i>sul1</i> , <i>sul2</i> , <i>tet(A)</i>
PT28_3	<i>aadB</i> , <i>aph(3')-Ia</i> , <i>floR</i> , <i>mdfA</i> , <i>strA</i> , <i>strB</i> , <i>sul1</i> , <i>sul2</i>
PT28_4	<i>aadA</i> , <i>aph(3')-Ia</i> , <i>bla_{CTX-M-1}</i> , <i>dfrA17</i> , <i>mdfA</i> , <i>mphA</i> , <i>qnrS1</i> , <i>strA</i> , <i>strB</i> , <i>sul2</i>

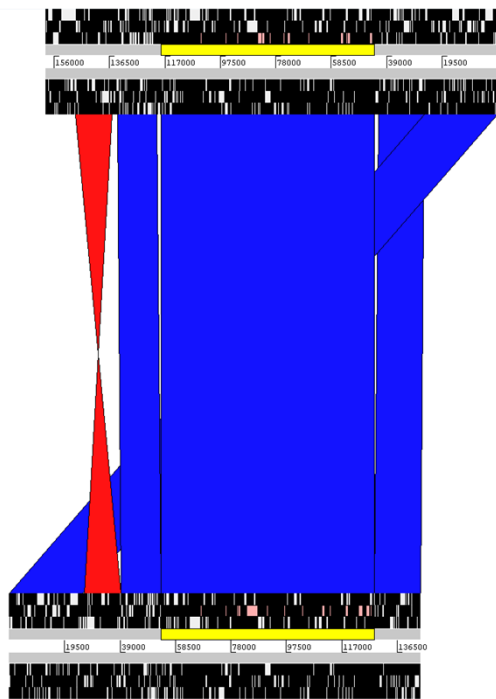


O80:H2

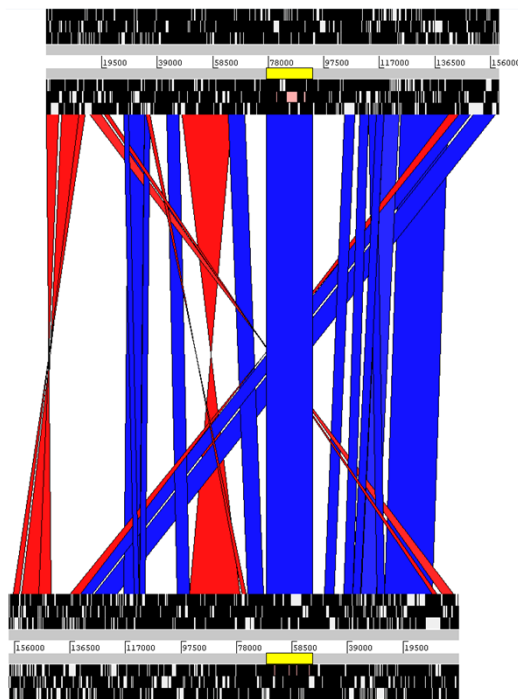
O45:H2

Results, blastn & ACT_1

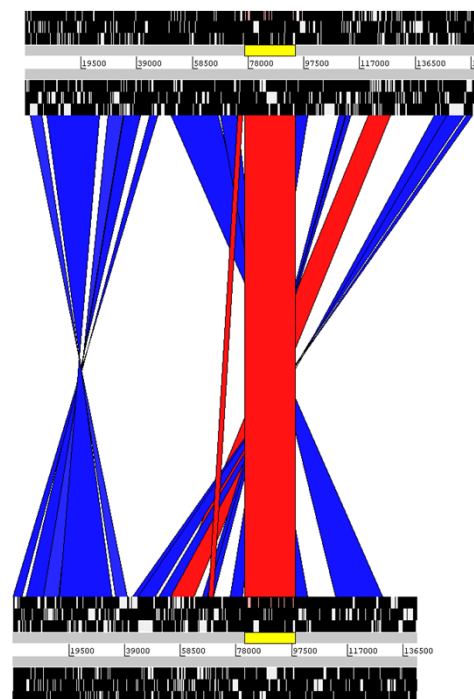
NL1800740, **contig 2** versus:



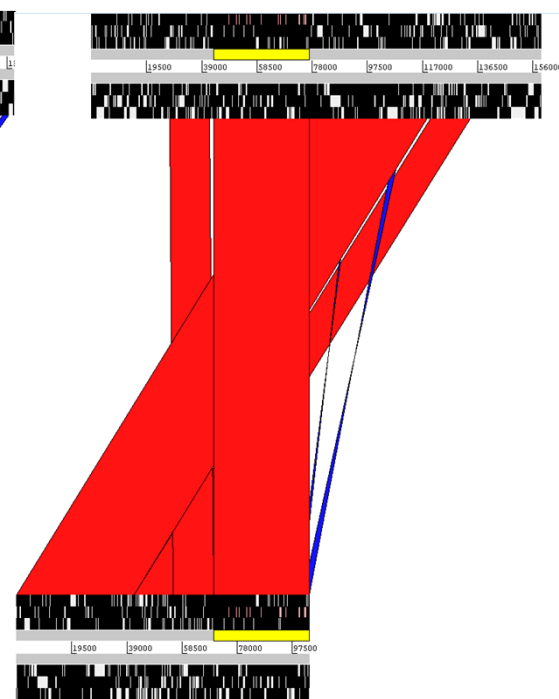
NL1700800, **contig 4**



NL1700844, **contig 4**



NL1800064, **contig 3**



NL1800418, **contig 3**



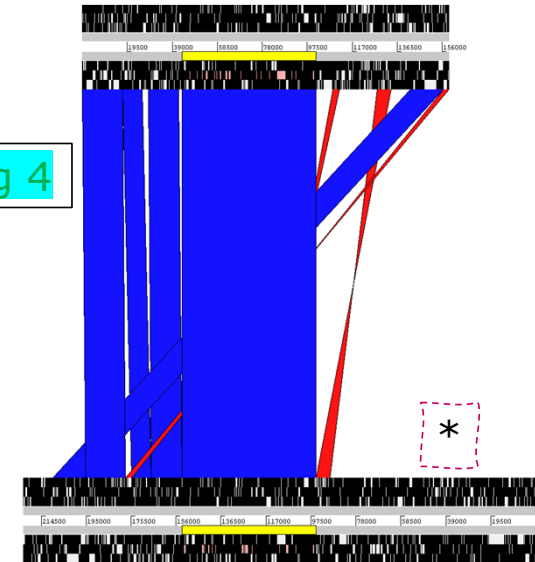
O80:H2

O45:H2

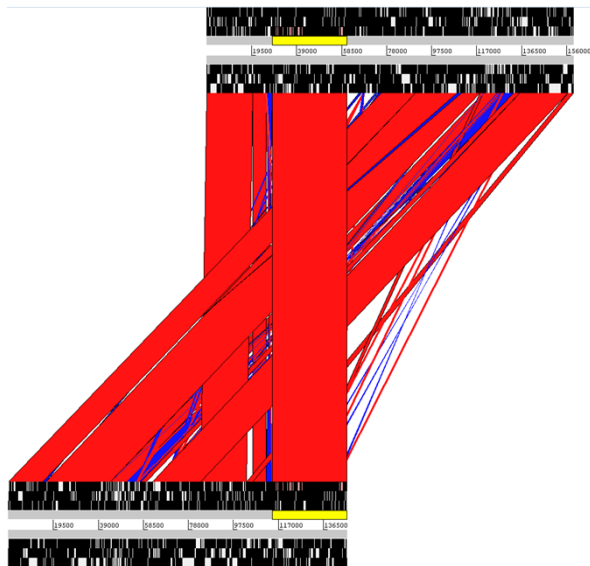
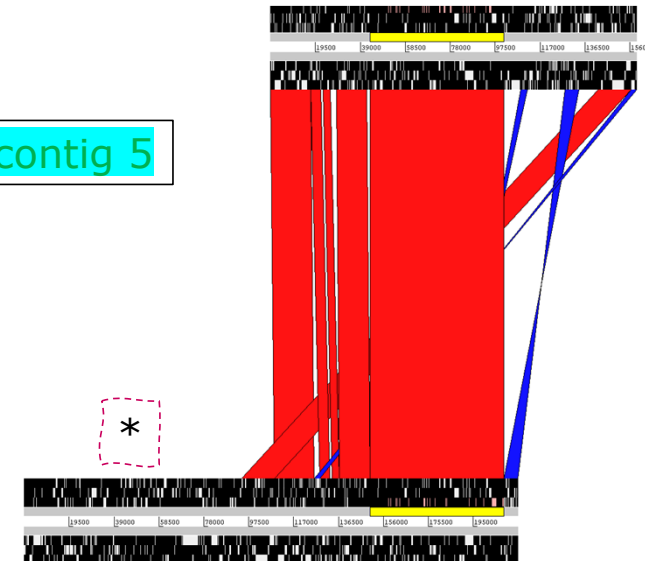
PT28_1, contig 4

Results, blastn & ACT_2

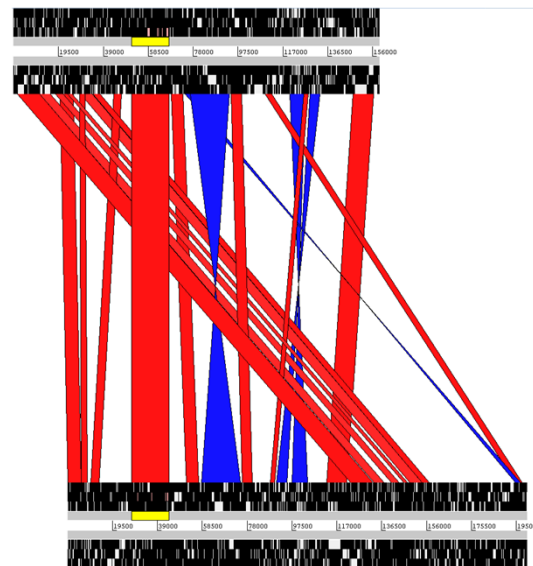
NL1800740, contig 2 versus:



PT28_3, contig 5



PT28_2, contig 8



PT28_4, contig 3



Discussion

- TGS helped characterise nine STEC/ExPEC O80:H2 and O45:H2 strains in more detail, more specifically their pR444_A like plasmid.
- Complete pR444_A like plasmids were obtained for seven strains
- The gene content of the ExPEC plasmids is diverse as well as their structure.
- The strains contain 25 or more VFs, 8-11 on the pR444_A like plasmids
- The strains contain 7-12 AMR genes, 1-9 on the pR444_A like plasmids



Acknowledgements

RIVM

- Amber Appel, Marta Rozwandowicz, Maaïke van den Beld, Eelco Franz and people from our bioinformatics team

ISS

- Federica Gigliucci, Valeria Michelacci, Stefano Morabito, and the rest of the EURL staff