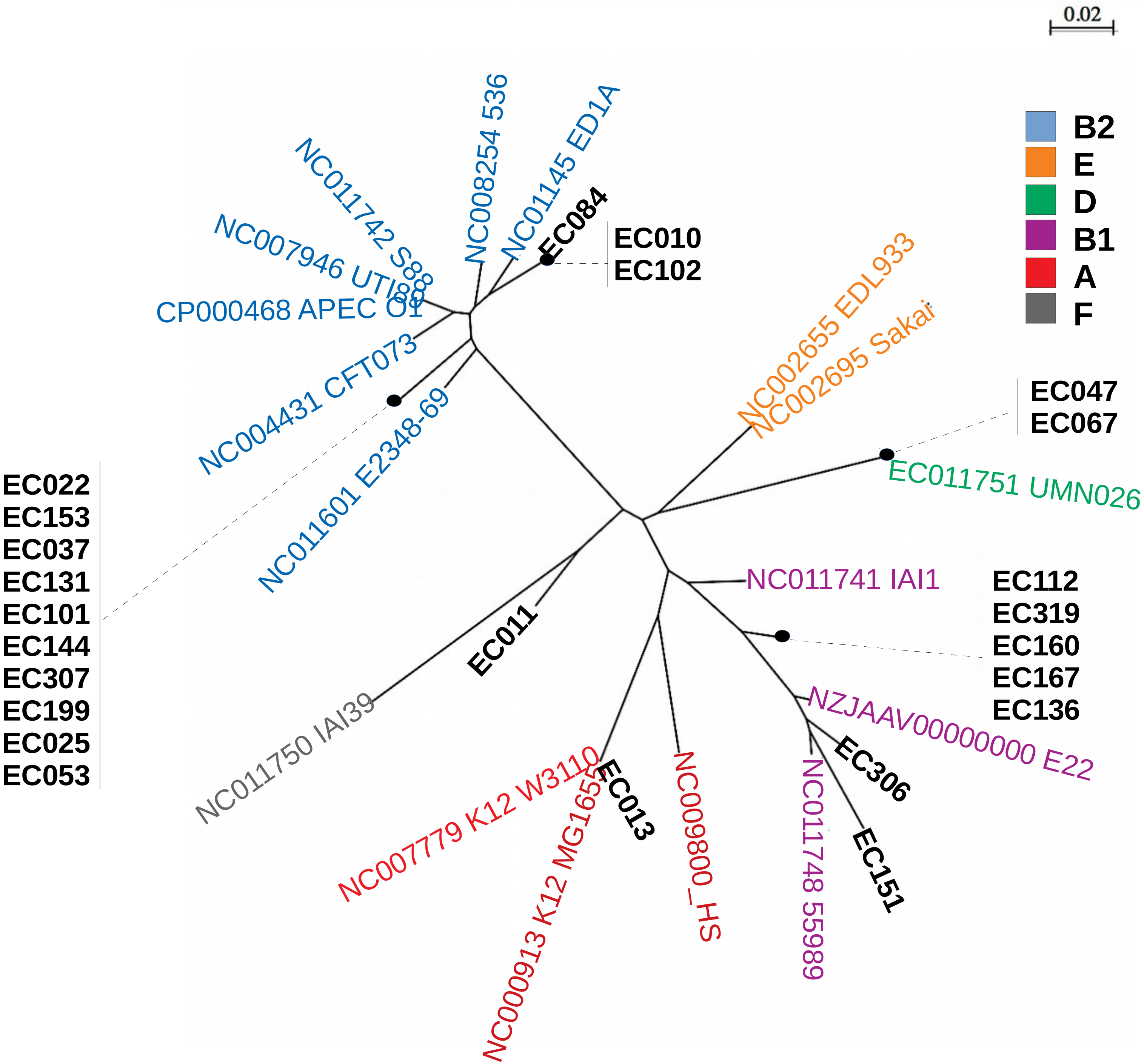
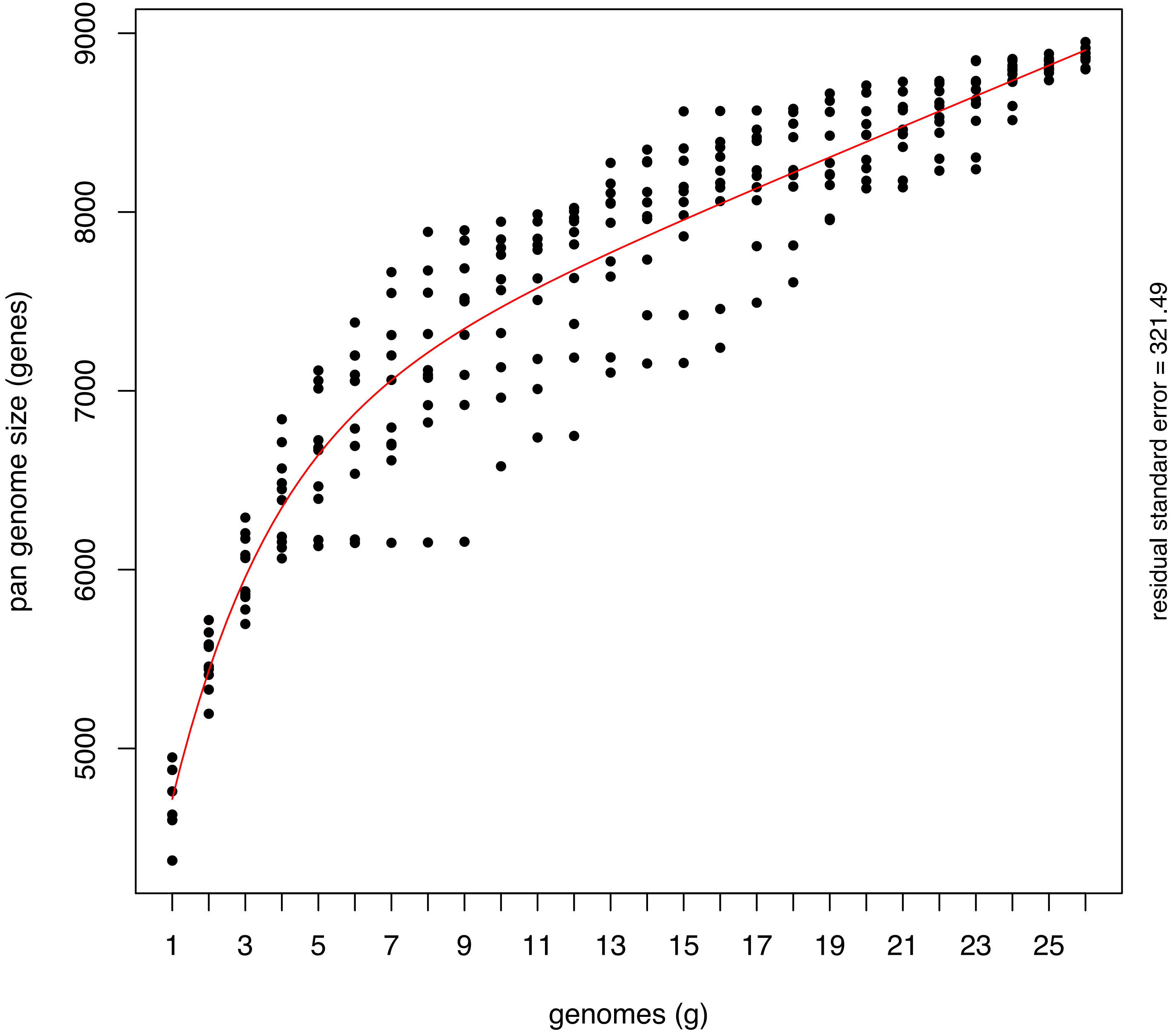

0.02
🟦 B2
🟧 E
🟩 D
🟪 B1
🟥 A
⬛ F
NC011742 S88
NC008254 536
NC011145 ED1A
EC084
EC010
EC102
NC007946 UTI89
CP000468 APEC O1
NC004431 CFT073
NC011601 E2348-69
EC022
EC153
EC037
EC131
EC101
EC144
EC307
EC199
EC025
EC053
NC002655 EDL933
NC002695 Sakai
EC047
EC067
EC011751 UMN026
NC011741 IAI1
EC112
EC319
EC160
EC167
EC136
NZJAAV00000000 E22
EC306
EC151
NC011748 55989
NC009800_HS
EC013
NC000913 K12 MG1655
NC007779 K12 W3110
EC011
NC011750 IAI39

Supp. Fig. 1

$$\text{pangenes}(g) = 4717 + 85.2(g-1) + 1317\,\exp\left(\frac{-2}{2.72}\right)\frac{1-\exp\left(\frac{-(g-1)}{2.72}\right)}{1-\exp\left(\frac{-1}{2.72}\right)}$$



Supp. Fig. 2

Supplementary File 1: Quality evaluation for 24 UPEC de novo draft genome assemblies.

We calculated the total assembly length and the N50 value (the length of the smallest contig which when added to the other larger contigs of the assembly yields at least 50% of the genome assembly) of our 24 *de novo* draft genome assemblies. **Table S1-A** shows the obtained values for these statistics. **Figure S1-1** shows the total Assembly length of our assemblies in comparison with genome length of *E. coli* K.12 MG1655 and UPEC CFT073. **Figure S1-2** shows the N50 values of our assemblies in comparison with the N50 value obtained by the SPAdes developers in a *de novo* genome assembly with 26 million reads of an Illumina 2x100 paired end library of *E. coli* K-12 MG1655 (http://cab.spbu.ru/software/spades/).

To further examine the completeness of our assemblies we aligned the forward reads of each of our 24 UPEC libraries (included in the assembly input) against the corresponding genome assemblies with bowtie. and measured the percentage of aligned reads. We expect that if an assembly is nearly complete (represents most of the genome of a strain) most of the library reads will align to it, as the reads are expected to represent the complete genome in fragments. **Table S1-B** shows the percentage of each library reads that align to the corresponding *de novo* draft genome assembly.

Table S1-A. N50 and Assembly Length values for our de-novo draft genome assemblies.

Serotype-ST-Phylogroup	Sample	N50	Assembly length
025:H4 (ST 131) B2	EC022	227017	5290048
025:H4 (ST 131) B2	EC153	188599	5266057
025:H4 (ST 131) B2	EC037	173941	5268188
025:H4 (ST 131) B2	EC131	190541	5388762
025:H4 (ST 131) B2	EC101	191226	5256127
025:H4 (ST 131) B2	EC144	165794	5343505
025:H4 (ST 131)B2	EC307	189333	5344076
025:H4 (ST 131) B2	EC199	173944	5303066
025:H4 (ST 131) B2	EC025	191227	5174605
016:H5 (ST-131) B2	EC053	260028	4990640
075:H5 (ST-1193) B2	EC010	134538	5141539
075:H5 (ST-1193) B2	EC102	178414	5075017
075:H5 (ST-14) B2	EC084	147108	5147666
08:H9 (ST-423) B1	EC112	187031	5065990
08:H9 (ST-423) B1	EC319	172754	5066920
08:H9 (ST-423) B1	EC136	172754	5060390
08:H9 (ST-423) B1	EC167	231780	5080520
08:H9 (ST-423) B1	EC160	231544	5066067
045:H11 (ST-297) B1	EC306	217310	4817787
0116:H48 (ST-351) B1	EC151	126473	4717866
016:H4 (ST10) A	EC013	94645	5203374
NA (ST-69) D	EC047	47191	5362429
NA (ST-69) D	EC067	267993	5201365
01:H6 (ST-648) -	EC011	97802	5374532

Supplementary File 1: Quality evaluation for 24 UPEC de novo draft genome assemblies.

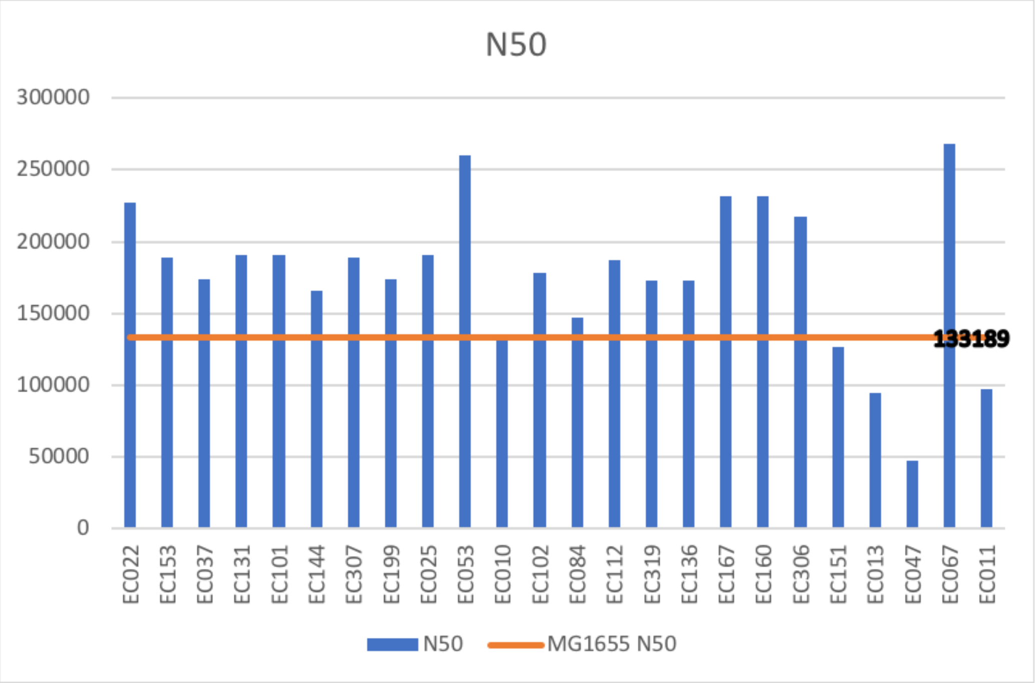


Figure S1-1. N50 values for our 24 de novo genome assemblies. The orange line indicates de N50 value reported by the SPAdes developers for a Illumina paired-end library of *E. coli* K-12 MG1655.

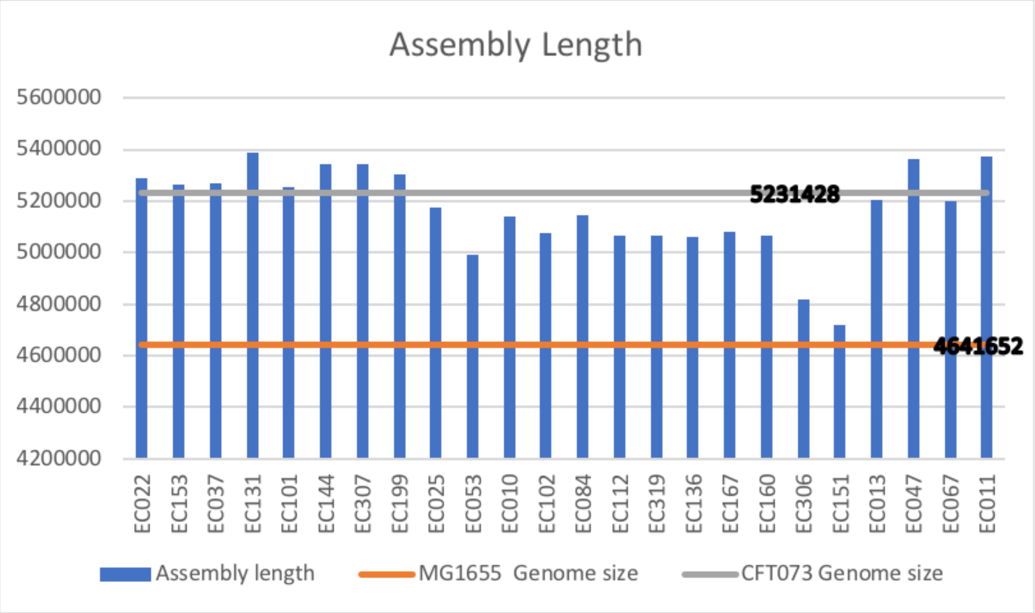
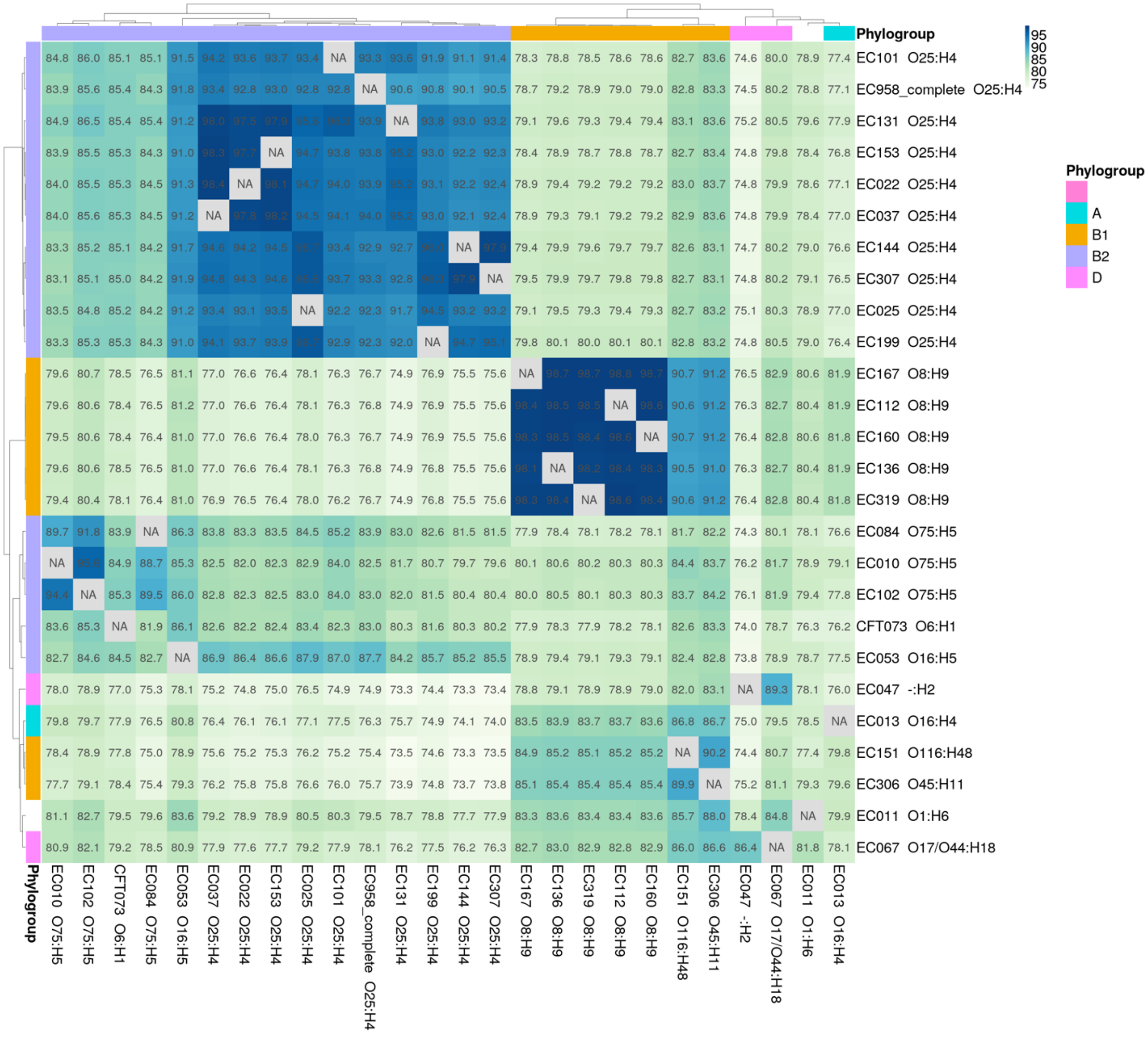


Figure S1-2. Assembly length values for our 24 de novo genome assemblies. The orange and gray lines indicate de genome size of *E. coli* K-12 MG1655 and UPEC CFT073, respectively.

Supplementary File 2: Additional synteny assessment statistics for the 24 UPEC strains and the model strains CFT073 and EC958 ST-131.

We identified all of the regions comprising at least 5 contiguous genes with a maximum gap of 1 gene present in only one of the strains (see Materials and Methods). **Figures S2-1, S2-2 and S2-3** show different statistics obtained from the i-ADHoRe results.

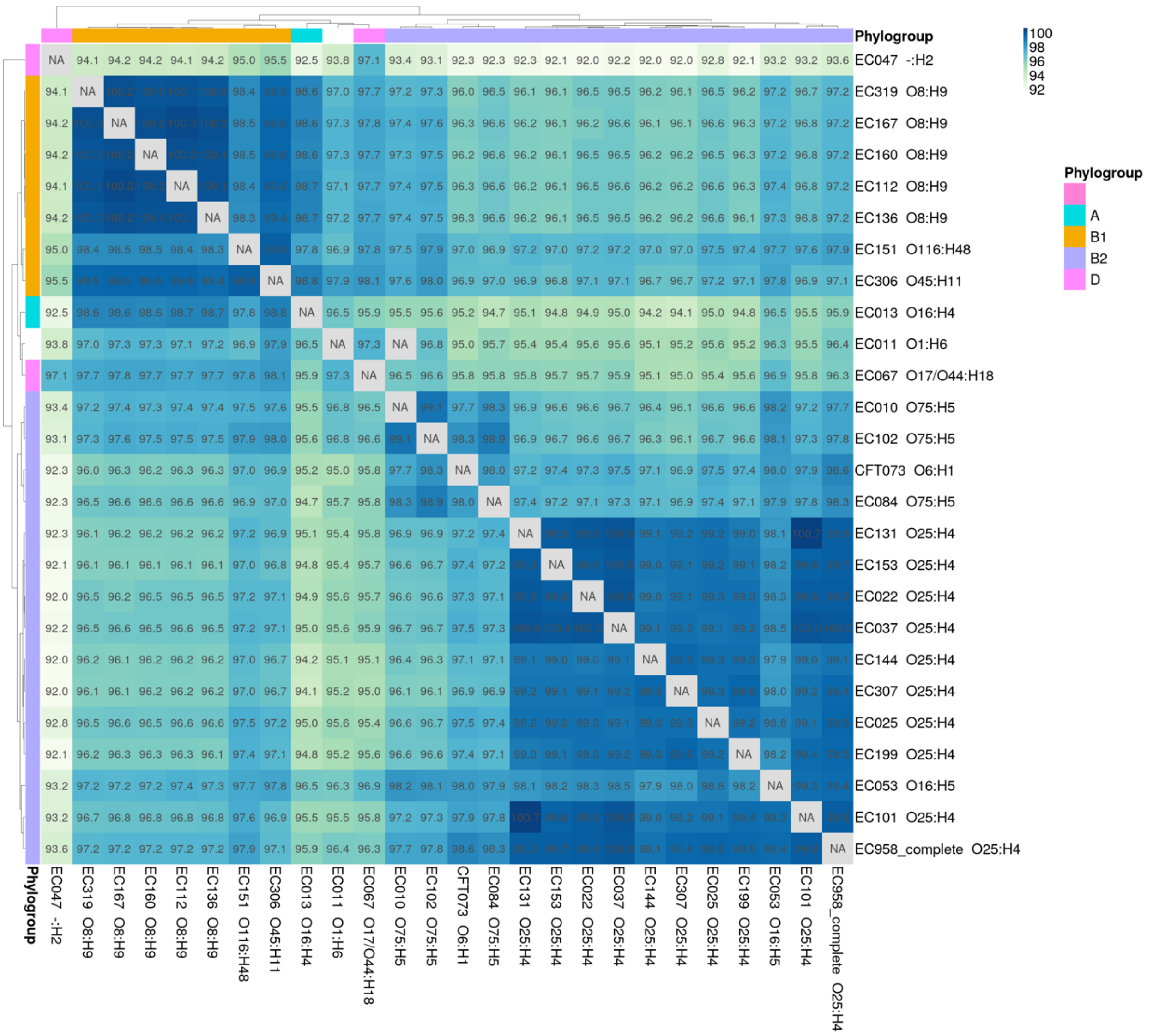
Figure S2-1. Synteny percentage for each strain comparison. The percentage of the total genes of each strain a (column labels) that appear in syntenic regions in each strain b (row labels) is shown. The strains were hierarchically clustered according to these values. The phylogroups and serotypes of the strains are shown at the heatmap margins and in the row and column labels, respectively.



	EC010 O75:H5	EC102 O75:H5	CFT073 O6:H1	EC084 O75:H5	EC053 O16:H5	EC037 O25:H4	EC022 O25:H4	EC153 O25:H4	EC025 O25:H4	EC101 O25:H4	EC958_complete O25:H4	EC131 O25:H4	EC199 O25:H4	EC144 O25:H4	EC307 O25:H4	EC167 O8:H9	EC136 O8:H9	EC319 O8:H9	EC112 O8:H9	EC160 O8:H9	EC151 O116:H48	EC306 O45:H11	EC047 -:H2	EC067 O17/O44:H18	EC011 O1:H6	EC013 O16:H4
EC101 O25:H4	84.8	86.0	85.1	85.1	91.5	94.2	93.6	93.7	93.4	NA	93.3	93.6	91.9	91.1	91.4	78.3	78.8	78.5	78.6	78.6	82.7	83.6	74.6	80.0	78.9	77.4
EC958_complete O25:H4	83.9	85.6	85.4	84.3	91.8	93.4	92.8	93.0	92.8	92.8	NA	90.6	90.8	90.1	90.5	78.7	79.2	78.9	79.0	79.0	82.8	83.3	74.5	80.2	78.8	77.1
EC131 O25:H4	84.9	86.5	85.4	85.4	91.2	98.0	97.5	97.9	95.6	96.3	93.9	NA	93.8	93.0	93.2	79.1	79.6	79.3	79.4	79.4	83.1	83.6	75.2	80.5	79.6	77.9
EC153 O25:H4	83.9	85.5	85.3	84.3	91.0	98.3	97.7	NA	94.7	93.8	93.8	95.2	93.0	92.2	92.3	78.4	78.9	78.7	78.8	78.7	82.7	83.4	74.8	79.8	78.4	76.8
EC022 O25:H4	84.0	85.5	85.3	84.5	91.3	98.4	NA	98.1	94.7	94.0	93.9	95.2	93.1	92.2	92.4	78.9	79.4	79.2	79.2	79.2	83.0	83.7	74.8	79.9	78.6	77.1
EC037 O25:H4	84.0	85.6	85.3	84.5	91.2	NA	97.8	98.2	94.5	94.1	94.0	95.2	93.0	92.1	92.4	78.9	79.3	79.1	79.2	79.2	82.9	83.6	74.8	79.9	78.4	77.0
EC144 O25:H4	83.3	85.2	85.1	84.2	91.7	94.6	94.2	94.5	96.7	93.4	92.9	92.7	96.0	NA	97.9	79.4	79.9	79.6	79.7	79.7	82.6	83.1	74.7	80.2	79.0	76.6
EC307 O25:H4	83.1	85.1	85.0	84.2	91.9	94.8	94.3	94.6	96.6	93.7	93.3	92.8	96.3	97.9	NA	79.5	79.9	79.7	79.8	79.8	82.7	83.1	74.8	80.2	79.1	76.5
EC025 O25:H4	83.5	84.8	85.2	84.2	91.2	93.4	93.1	93.5	NA	92.2	92.3	91.7	94.5	93.2	93.2	79.1	79.5	79.3	79.4	79.3	82.7	83.2	75.1	80.3	78.9	77.0
EC199 O25:H4	83.3	85.3	85.3	84.3	91.0	94.1	93.7	93.9	96.7	92.9	92.3	92.0	NA	94.7	95.1	79.8	80.1	80.0	80.1	80.1	82.8	83.2	74.8	80.5	79.0	76.4
EC167 O8:H9	79.6	80.7	78.5	76.5	81.1	77.0	76.6	76.4	78.1	76.3	76.7	74.9	76.9	75.5	75.6	NA	98.7	98.7	98.8	98.7	90.7	91.2	76.5	82.9	80.6	81.9
EC112 O8:H9	79.6	80.6	78.4	76.5	81.2	77.0	76.6	76.4	78.1	76.3	76.8	74.9	76.9	75.5	75.6	98.4	98.5	98.5	NA	98.6	90.6	91.2	76.3	82.7	80.4	81.9
EC160 O8:H9	79.5	80.6	78.4	76.4	81.0	77.0	76.6	76.4	78.0	76.3	76.7	74.9	76.9	75.5	75.6	98.3	98.5	98.4	98.6	NA	90.7	91.2	76.4	82.8	80.6	81.8
EC136 O8:H9	79.6	80.6	78.5	76.5	81.0	77.0	76.6	76.4	78.1	76.3	76.8	74.9	76.8	75.5	75.6	98.1	NA	98.2	98.4	98.3	90.5	91.0	76.3	82.7	80.4	81.9
EC319 O8:H9	79.4	80.4	78.1	76.4	81.0	76.9	76.5	76.4	78.0	76.2	76.7	74.9	76.8	75.5	75.6	98.3	98.4	NA	98.6	98.4	90.6	91.2	76.4	82.8	80.4	81.8
EC084 O75:H5	89.7	91.8	83.9	NA	86.3	83.8	83.3	83.5	84.5	85.2	83.9	83.0	82.6	81.5	81.5	77.9	78.4	78.1	78.2	78.1	81.7	82.2	74.3	80.1	78.1	76.6
EC010 O75:H5	NA	95.6	84.9	88.7	85.3	82.5	82.0	82.3	82.9	84.0	82.5	81.7	80.7	79.7	79.6	80.1	80.6	80.2	80.3	80.3	84.4	83.7	76.2	81.7	78.9	79.1
EC102 O75:H5	94.4	NA	85.3	89.5	86.0	82.8	82.3	82.5	83.0	84.0	83.0	82.0	81.5	80.4	80.4	80.0	80.5	80.1	80.3	80.3	83.7	84.2	76.1	81.9	79.4	77.8
CFT073 O6:H1	83.6	85.3	NA	81.9	86.1	82.6	82.2	82.4	83.4	82.3	83.0	80.3	81.6	80.3	80.2	77.9	78.3	77.9	78.2	78.1	82.6	83.3	74.0	78.7	76.3	76.2
EC053 O16:H5	82.7	84.6	84.5	82.7	NA	86.9	86.4	86.6	87.9	87.0	87.7	84.2	85.7	85.2	85.5	78.9	79.4	79.1	79.3	79.1	82.4	82.8	73.8	78.9	78.7	77.5
EC047 -:H2	78.0	78.9	77.0	75.3	78.1	75.2	74.8	75.0	76.5	74.9	74.9	73.3	74.4	73.3	73.4	78.8	79.1	78.9	78.9	79.0	82.0	83.1	NA	89.3	78.1	76.0
EC013 O16:H4	79.8	79.7	77.9	76.5	80.8	76.4	76.1	76.1	77.1	77.5	76.3	75.7	74.9	74.1	74.0	83.5	83.9	83.7	83.7	83.6	86.8	86.7	75.0	79.5	78.5	NA
EC151 O116:H48	78.4	78.9	77.8	75.0	78.9	75.6	75.2	75.3	76.2	75.2	75.4	73.5	74.6	73.3	73.5	84.9	85.2	85.1	85.2	85.2	NA	90.2	74.4	80.7	77.4	79.8
EC306 O45:H11	77.7	79.1	78.4	75.4	79.3	76.2	75.8	75.8	76.6	76.0	75.7	73.9	74.8	73.7	73.8	85.1	85.4	85.4	85.4	85.4	89.9	NA	75.2	81.1	79.3	79.6
EC011 O1:H6	81.1	82.7	79.5	79.6	83.6	79.2	78.9	78.9	80.5	80.3	79.5	78.7	78.8	77.7	77.9	83.3	83.6	83.4	83.4	83.6	85.7	88.0	78.4	84.8	NA	79.9
EC067 O17/O44:H18	80.9	82.1	79.2	78.5	80.9	77.9	77.6	77.7	79.2	77.9	78.1	76.2	77.5	76.2	76.3	82.7	83.0	82.9	82.8	82.9	86.0	86.6	86.4	NA	81.8	78.1

Supplementary File 2: Additional synteny assessment statistics for the 24 UPEC strains and the model strains CFT073 and EC958 ST-131.

Figure S2-2. Synteny percentage for the genes shared between each pair of strains. The percentage of the shared genes that appear in regions with the same synteny in both strains is shown. The strains were hierarchically clustered according to these values. The phylogroups and serotypes of the strains are shown at the heatmap margins and in the row and column labels, respectively.



	EC047 -:H2	EC319 O8:H9	EC167 O8:H9	EC160 O8:H9	EC112 O8:H9	EC136 O8:H9	EC151 O116:H48	EC306 O45:H11	EC013 O16:H4	EC011 O1:H6	EC067 O17/O44:H18	EC010 O75:H5	EC102 O75:H5	CFT073 O6:H1	EC084 O75:H5	EC131 O25:H4	EC153 O25:H4	EC022 O25:H4	EC037 O25:H4	EC144 O25:H4	EC307 O25:H4	EC025 O25:H4	EC199 O25:H4	EC053 O16:H5	EC101 O25:H4	EC958_complete O25:H4
EC047 -:H2	NA	94.1	94.2	94.2	94.1	94.2	95.0	95.5	92.5	93.8	97.1	93.4	93.1	92.3	92.3	92.3	92.1	92.0	92.2	92.0	92.0	92.8	92.1	93.2	93.2	93.6
EC319 O8:H9	94.1	NA	100.2	100.0	100.1	100.0	98.4	99.5	98.6	97.0	97.7	97.2	97.3	96.0	96.5	96.1	96.1	96.5	96.5	96.2	96.1	96.5	96.2	97.2	96.7	97.2
EC167 O8:H9	94.2	100.2	NA	100.2	100.3	100.2	98.5	99.6	98.6	97.3	97.8	97.4	97.6	96.3	96.6	96.2	96.1	96.2	96.6	96.1	96.1	96.6	96.3	97.2	96.8	97.2
EC160 O8:H9	94.2	100.0	100.2	NA	100.2	100.1	98.5	99.6	98.6	97.3	97.7	97.3	97.5	96.2	96.6	96.2	96.1	96.5	96.5	96.2	96.2	96.5	96.3	97.2	96.8	97.2
EC112 O8:H9	94.1	100.1	100.3	100.2	NA	100.1	98.4	99.6	98.7	97.1	97.7	97.4	97.5	96.3	96.6	96.2	96.1	96.5	96.6	96.2	96.2	96.6	96.3	97.4	96.8	97.2
EC136 O8:H9	94.2	100.0	100.2	100.1	100.1	NA	98.3	99.4	98.7	97.2	97.7	97.4	97.5	96.3	96.6	96.2	96.1	96.5	96.5	96.2	96.2	96.6	96.1	97.3	96.8	97.2
EC151 O116:H48	95.0	98.4	98.5	98.5	98.4	98.3	NA	99.8	97.8	96.9	97.8	97.5	97.9	97.0	96.9	97.2	97.0	97.2	97.2	97.0	97.0	97.5	97.4	97.7	97.6	97.9
EC306 O45:H11	95.5	99.5	99.6	99.6	99.6	99.4	99.8	NA	98.8	97.9	98.1	97.6	98.0	96.9	97.0	96.9	96.8	97.1	97.1	96.7	96.7	97.2	97.1	97.8	96.9	97.1
EC013 O16:H4	92.5	98.6	98.6	98.6	98.7	98.7	97.8	98.8	NA	96.5	95.9	95.5	95.6	95.2	94.7	95.1	94.8	94.9	95.0	94.2	94.1	95.0	94.8	96.5	95.5	95.9
EC011 O1:H6	93.8	97.0	97.3	97.3	97.1	97.2	96.9	97.9	96.5	NA	97.3	NA	96.8	95.0	95.7	95.4	95.4	95.6	95.6	95.1	95.2	95.6	95.2	96.3	95.5	96.4
EC067 O17/O44:H18	97.1	97.7	97.8	97.7	97.7	97.7	97.8	98.1	95.9	97.3	NA	96.5	96.6	95.8	95.8	95.8	95.7	95.7	95.9	95.1	95.0	95.4	95.6	96.9	95.8	96.3
EC010 O75:H5	93.4	97.2	97.4	97.3	97.4	97.4	97.5	97.6	95.5	96.8	96.5	NA	99.1	97.7	98.3	96.9	96.6	96.6	96.7	96.4	96.1	96.6	96.6	98.2	97.2	97.7
EC102 O75:H5	93.1	97.3	97.6	97.5	97.5	97.5	97.9	98.0	95.6	96.8	96.6	99.1	NA	98.3	98.9	96.9	96.7	96.6	96.7	96.3	96.1	96.7	96.6	98.1	97.3	97.8
CFT073 O6:H1	92.3	96.0	96.3	96.2	96.3	96.3	97.0	96.9	95.2	95.0	95.8	97.7	98.3	NA	98.0	97.2	97.4	97.3	97.5	97.1	96.9	97.5	97.4	98.0	97.9	98.6
EC084 O75:H5	92.3	96.5	96.6	96.6	96.6	96.6	96.9	97.0	94.7	95.7	95.8	98.3	98.9	98.0	NA	97.4	97.2	97.1	97.3	97.1	96.9	97.4	97.1	97.9	97.8	98.3
EC131 O25:H4	92.3	96.1	96.2	96.2	96.2	96.2	97.2	96.9	95.1	95.4	95.8	96.9	96.9	97.2	97.4	NA	99.9	99.9	100.0	99.1	99.2	99.2	99.0	98.1	100.7	99.8
EC153 O25:H4	92.1	96.1	96.1	96.1	96.1	96.1	97.0	96.8	94.8	95.4	95.7	96.6	96.7	97.4	97.2	99.9	NA	99.8	100.0	99.0	99.1	99.2	99.1	98.2	99.6	99.7
EC022 O25:H4	92.0	96.5	96.2	96.5	96.5	96.5	97.2	97.1	94.9	95.6	95.7	96.6	96.6	97.3	97.1	99.9	99.8	NA	100.0	99.0	99.1	99.3	99.3	98.3	99.8	99.9
EC037 O25:H4	92.2	96.5	96.6	96.5	96.6	96.5	97.2	97.1	95.0	95.6	95.9	96.7	96.7	97.5	97.3	100.0	100.0	100.0	NA	99.1	99.2	99.1	99.2	98.5	100.0	100.0
EC144 O25:H4	92.0	96.2	96.1	96.2	96.2	96.2	97.0	96.7	94.2	95.1	95.1	96.4	96.3	97.1	97.1	99.1	99.0	99.0	99.1	NA	99.5	99.3	99.3	97.9	99.0	99.1
EC307 O25:H4	92.0	96.1	96.1	96.2	96.2	96.2	97.0	96.7	94.1	95.2	95.0	96.1	96.1	96.9	96.9	99.2	99.1	99.1	99.2	99.5	NA	99.3	99.6	98.0	99.2	99.4
EC025 O25:H4	92.8	96.5	96.6	96.5	96.6	96.6	97.5	97.2	95.0	95.6	95.4	96.6	96.7	97.5	97.4	99.2	99.2	99.3	99.1	99.3	99.3	NA	99.2	98.8	99.1	99.5
EC199 O25:H4	92.1	96.2	96.3	96.3	96.3	96.1	97.4	97.1	94.8	95.2	95.6	96.6	96.6	97.4	97.1	99.0	99.1	99.3	99.2	99.3	99.6	99.2	NA	98.2	99.4	99.5
EC053 O16:H5	93.2	97.2	97.2	97.2	97.4	97.3	97.7	97.8	96.5	96.3	96.9	98.2	98.1	98.0	97.9	98.1	98.2	98.3	98.5	97.9	98.0	98.8	98.2	NA	99.3	99.4
EC101 O25:H4	93.2	96.7	96.8	96.8	96.8	96.8	97.6	96.9	95.5	95.5	95.8	97.2	97.3	97.9	97.8	100.7	99.6	99.8	100.0	99.0	99.2	99.1	99.4	99.3	NA	99.8
EC958_complete O25:H4	93.6	97.2	97.2	97.2	97.2	97.2	97.9	97.1	95.9	96.4	96.3	97.7	97.8	98.6	98.3	99.8	99.7	99.9	100.0	99.1	99.4	99.5	99.5	99.4	99.8	NA