

Supplementary data

Figure S1. Consensus sequence of the 43-bp *sopC* repeats. Alignment of the *E. coli* plasmid F centromeric sequence (Mori *et al.*, 1986) every 43 bases allow the determination of the consensus sequence, presented at the bottom. Inverted arrows indicate SopB binding sites, present in each repeats, and composed of two 7-bp inverted repeat sequences (in bold), separated by 2 central bases. Highlighted bases correspond to differences from the consensus. Repeats are numbered from 1 to twelve on the left.

| Repeat # | |
|------------------|---|
| 1 | 1- G CCTGA AT AT T CTC TCTGGG C CCAC T GT T CCACT T GTATCGTC |
| 2 | 44- GGTCTGAT A AT C AG A CTGGGACCACGGTCCCACTCGTATCGTC |
| 3 | 87- GGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTC |
| 4 | 130- GGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTC |
| 5 | 173- GGTCTGAT A AT C AG A CTGGGACCACGGTCCCACTCGTATCGTC |
| 6 | 216- GGTCTGATTATTAGTCTGGGACCA TGGTCCCACTCGTATCGTC |
| 7 | 259- GGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTC |
| 8 | 302- GGTCTGATTATTAGTCTGG A ACCACGGTCCCACTCGTATCGTC |
| 9 | 345- GGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTC |
| 10 | 388- GGTCTGATTATTAGTCTGGGACCACG A TCCCACTCGT C T TGTCT |
| 11 | 431- GGTCTGATTAT C GGTCTGGGACCACGGTCCCACT TGTAT TGTCT |
| 12 | 474- G A TC A GA C TAT C AG C G TGA GAC TACGA T TCCA TCAATGCC TGT |
| | |
| consensus | GGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTC |

Figure S2. SopB-*sopC* binding efficiency measured by EMSA. Increased SopB concentrations (0, 3, 10, 30, 100 and 300 nM) were incubated with 0.3 nM of 32 P-labelled DNA fragments. Various *sopC* sequences were used as probes (E1, E2, E19, E23, E39, E40 and E46 are described in the main text; Fig. 6). All reaction mixtures contained unlabeled sonicated salmon sperm DNA ($100 \mu\text{g}\cdot\text{ml}^{-1}$) as non-specific competitor DNA. Position of discrete SopB-*sopC* complexes (B1) and free DNA duplexes (*sopC*) are indicated on the left. * indicates a secondary complex formed when unpurified oligonucleotides, but not when HPLC purified oligonucleotides (bottom panel, E2), were used for duplexes formation.

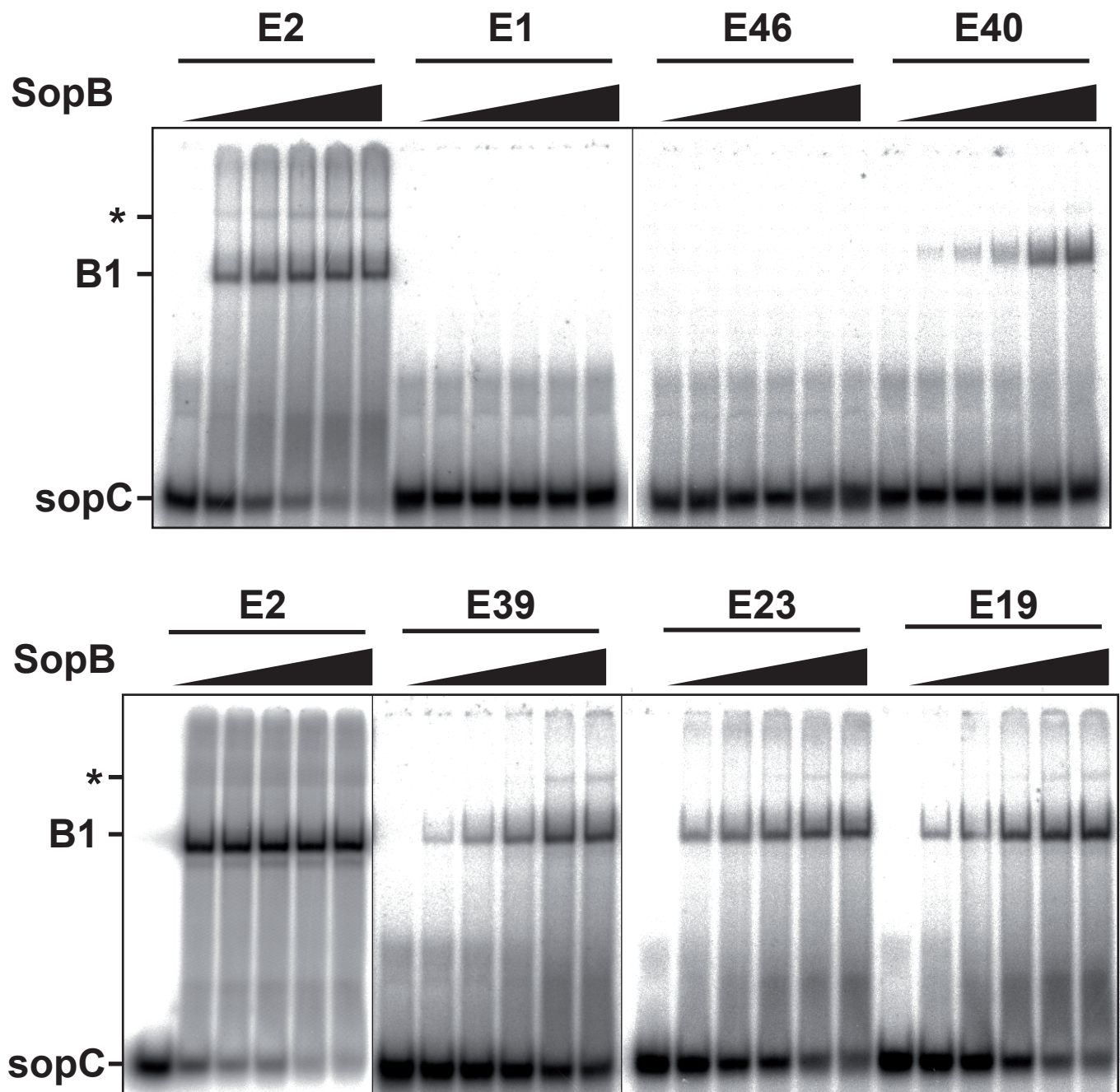


Figure S3: Alignments of centromeric sequences homologous to *F sopC*. Homologous sequences of the SopB (F) binding site consensus were identified in the nr database using Blastn (NCBI), with various combinations of query sequences. Sequences with at least 2 IR spaced by 43-, 44-, and 45- bp were retained for further analysis. Sequences were aligned to display repeats, and classified according to the number of repeats. Sequences in bold correspond to the SopB binding site. Nucleotides in blue correspond to the inverted repeat, while red ones are different from the consensus binding site.

- 17 repeats of 43 bp

Klebsiella pneumoniae plasmid pKPN3

>gi|150958208|gb|CP000648.1| *Klebsiella pneumoniae* subsp. *pneumoniae* MGH 78578 plasmid pKPN3

```
CCTTTACTTCATTTCTTTCTCACAGGCCAGAAAACATAACTG
GCCTGAATATTCTCTCTTGGGACCACGGTCCCACCTGCATCGTC
GGTCAGGTTATCTGCCTGGGACCACGGTCTCACCTGCGTCATC
GGTCAGGTTATCTGCCTGGGACCACGGTCCCACCTGCGTCGTC
GGTCAGGTTATCTGTCTTGGGACCACGGTCCCACCTGCGTCATC
TGTCAGGTAATCTGTCTTGGGACCACGGTCCCACCTGCGTCATC
TGTCAGGTAATCTGTCTTGGGACCACGGTCCCACCTGCGTCATC
GGTCAGGTTATCTGACTTGGGACCACGGGCCCACCTGCGTCGTC
GGTCAGGTTATCTGCCTGGGACCACGGTCCCACCTGCGTCATC
GGTCAGCTTATCTGTCTTGGGACCACGGTCTCACCTGCGTCATC
GGTCAGCTTATCTGTCTTGGGACCACGGTCCCACCTGCGTCGTC
GGTCAGGTTATCTGTCTTGGGACCACGGTCCCACCTGCGTCATC
GGTCAGGTAATCTGCCTGGGACCACGGTCCCACCTGTATCGCC
GGTCAGGTTATCTCCCTGGGACCACGGTCCCACCTGCGTCATC
GGTCAGGTTATCTGCCTGGGACCACGGTCCCACCTGCGTCGTC
GGTCAGCTTATCTGCCTGGGACCACGGTCCCACCTGCGTCATC
TGTCAGCTTATCTGCCTGGGACCACGGTCCCACCTGCATCGTC
GG CACATTATTAGTTAGGGATAACGGTTCAATCAATGCCTGT
CACGAACAGGTATTGGCTTGTGCCAGCAACCCGATGAGCGGA
```

(note that a gap was introduced in the 17th repeat for alignment)

- 16 repeats of 43 bp

Escherichia coli plasmid p1658/97

>gi|28629230|gb|AF550679.1| *Escherichia coli* plasmid p1658/97

```
TACTTAATTTCTTTAGCACAGGCTAGAAAACATAACTAACTG
GCCTGAATATTATATCTTGGGACCACGGTCCCAGTTGTATTGTC
GGTCTGATAATCAGTCTTGGGACCACGGTCCCACTCATATTGTC
GGTCTGATAATCAGTCTTGGGACCACGGTCCCACTCATATTGTC
GGTCTGATAATCAGTCTTGGGACCACGGTCCCAGTTGTATTGTC
GGTCTGATAATCAGTCTTGGGACCACGGTCCCACTCATATTGTC
GGTCTGATAATCAGTCTTGGGACCACGGTCCCACTCATATTGTC
GGTCTGATTATCGGTCTGGGACCACGGGCCCACTCATATTGTT
AGTCTGATAATCAGTCTTGGGACCACGGTCCCACTCGTATTGTT
GGTCTGATAATCAGTCTTGGGACCACGGTCCCACTCATATTGTC
GGTCTGATAATCAGTCTTGGGACCACGGTCCCACTCATATTGTC
GGTCTGATAATCAGTCTTGGGACCACGGTCCCACTCATATTGTC
GGTCTGATAATCAGTCTTGGGACCACGGTCCCACTCATATTGTT
GGTCTGATTATCGGTCTGGGACCACGGTCCCACTCATATTGTT
AGTCTGATAATCAGTCTTGGGACCACGGTCCCACTCGTATTGTT
GGTCTGATAATCAGTCTTGGGACCACGGTCTCACTTGTATTGTC
GATCTGATTATCAGCGTGAGACTACGATTCCATTAATGCCTGC
CAGTGAGTGTGTGGAACATATCTTGGCCAGTGGCTTGGTCTGT
```

- 13 repeats of 43-bp

Escherichia coli plasmid pVM01

>gi|168830962|gb|EU330199.1| Escherichia coli plasmid pVM01.

```
TACTTAATTTCTTTAGCACAGGCTAGAAAACATAACTAACTG
GCCTGAATATTATATCTGGGACCACGGTCCCACTTGTATTGTC
GGTCTGATAATCAGTCTGGGACCACGGTCCCACTCATATTGTC
GGTCTGATAATCAGTCTGGGACCACGGTCCCACTCATATTGTC
GGTCTGATTATCGGTCTGGGACCACGGGCCCACTCATATTGTT
AGTCTGCTAATCAGTCTGGGACCACGGTCCCACTCGTATTGTT
GGTCTGCTAATCAGTCTGGGACCACGGTCCCACTCATATTGTC
GGTCTGCTAATCAGTCTGGGACCACGGTCCCACTCATATTGTC
GGTCTGCTAATCAGTCTGGGACCACGGTCCCACTCATATTGTC
GGTCTGCTAATCAGTCTGGGACCACGGTCCCACTCATATTGTT
GGTCTGATTATCGGTCTGGGACCACGGTCCCACTCATATTGTT
AGTCTGATAATCAGTCTGGGACCACGGTCCCACTCGTATTGTT
GGTCTGCTAATCAGTCTGGGACCACGGTCTCACTTGTATTGTC
GATCTGATTATCAGCGTGGAGACTACGATTCCAATTAATGCCTGC
CAGTGAGTGTGTGGAACATATCTTGGCCAGTGGCTTGGTCTGT
```

(*S. enterica* pCVM29188_146 and *E. coli* pSMS35_130 have the same centromeric sequence as pVM01)

- 12 repeats of 43-bp

Salmonella enterica plasmid pCVM19633_110

>gi|194709275|gb|CP001125.1| Salmonella enterica subsp. enterica str. CVM19633 plasmid pCVM19633_110.

```
CCTTTACTTCATTTCTTTCTCACAGGCCAGAAAACATAACTG
GCCTGAATATTCTCTCTGGGACCACGGTCCCACTTGTATCGCC
GCTCAGGTTATCTCCCTGGGACCACGGTCCCACTTGTATCGCC
GGTCAGCTTATCTGCCCTGGGACCACGGTCCCACTTGTATCGCC
GGTCAGGTTATCTGCCCTGGGACCACGGTCCCACTTGCATCGCC
GGTCAGCTTATCTGCCCTGGGACCACGGTCCCACTTGTATCGCC
GGTCAGGTTATCTGCCCTGGGACCACGGTCCCACTTGCATCGTC
TGTCAGCTTATCTGCCCTGGGACCACGGTCCCACTTGTATCGCC
GGTCAGGTTATCTGCCCTGGGACCACGGTCCCACTTGCATCGTC
TGTCAGCTTATCTGCCCTGGGACCACGGTCCCACTTGCCTTCGTC
GGTCAGCTTATCTGCCCTGGGACCACGGTCCCACTTGTATCGCC
GGTCAGGTTATCTGTCCTGGGACCACGGTGCCACTTGTATCGCC
GGTCAGGTTATCTGCCCTGGGACCACGGTCCCACTTGCATCGTC
GGCACATTATTAGTTAGGGACAACGGTTCAATCAATGCCTGTC
```

Klebsiella pneumoniae plasmid pLVPK

>gi|38016624|gb|AY378100.1| Klebsiella pneumoniae plasmid pLVPK.

```
CCTTTACTTCATTTCTTTCTCACAGGCCAGAAAAATAACTG
GCCTGAATACTCTCTATGGGACCACGGTCCCACCTGCATCGTC
GTTTAGGTTTTCCAGCCTGGGGCCACGGTCCCACCTGCATCGTC
GTTTAGGCTTCCAGCCTGGGACCACGGTCCCACCTGTATCGCC
GCTTAGGCTCCCAGCCTGGGACCACGGTCCCACCTGCATCGTC
GTTTAGGCTCCCAGCCTGGGACCACGGTCCCACCTGCATCGTC
GTTTAGGCTACCAGCCTGGGACCACGGTCCCACCTGCATCGTC
GTTTAGGTTTTCCAGCCTGGGACCACGGTCCCACCTGCATCGTC
GTTTAGGTTTTC AGCCTGGGACCACGGTCCCACCTGTATCGTC
GTTTAGGTTTTCCAGCCTGGGACCACGGTCCCACCTGCATCGTC
GTTTAGGCTTCCAGCCTGGGACCACGGTCCCACCTGTATCGTC
GTTTAGGTTTTCCAGCCTGGGACCACGGTCCCACCTGTATCGTC
GTTTAGGTTTTCCAGCCTGGGACCACGGTCCCACCTGCATCGTC
GTTTACATTATTAGTTTGGGATCGCAGTTCATCAATGCCTGTC
```

(note that a gap was introduced in the 8th repeat for alignment)

Klebsiella pneumoniae NTUH-K2044 plasmid pK2044

>gi|57158258|dbj|AP006726.1| Klebsiella pneumoniae NTUH-K2044 plasmid pK2044 DNA.

```
CCTTTACTTCATTTCTTTCTCACAGGCCAGAAAAATAACTG
GCCTGAATACTCTCTATGGGACCACGGTCCCACCTGCATCGTC
GTTTAGGCTTCCAGCCTGGGACCACGGTCCCACCTGTATCGCC
GCTTAGGCTCCCAGCCTGGGACCACGGTCCCACCTGCATCGTC
GTTTAGGCTCCCAGCCTGGGACCACGGTCCCACCTGCATCGTC
GTTTAGGCTTCCAGCCTGGGACCACGGTCCCACCTGCATCGTC
GTTTAGGTTTTCCAGCCTGGGACCACGGTCCCACCTGCATCGTC
GTTTAGGTTTTC AGCCTGGGACCACGGTCCCACCTGTATCGTC
GTTTAGGTTTTCCAGCCTGGGACCACGGTCCCACCTGCATCGTC
GTTTAGGCTTCCAGCCTGGGACCACGGTCCCACCTGTATCGTC
GTTTAGGTTTTCCAGCCTGGGACCACGGTCCCACCTGTATCGTC
GTTTAGGTTTTCCAGCCTGGGACCACGGTCCCACCTGCATCGTC
GTTTACATTATT AGTTTGGGATCGCAGTTCCAATGCCTGT
CACGAACAGGTATTGGCATGGCGCCGGTGACCCGTTGAGCGGA
```

(note that gaps were introduced in the 7th and 12th repeats for alignment)

Salmonella enterica plasmid pSC138

>gi|45758072|gb|AY509004.1| Salmonella enterica subsp. enterica serovar Choleraesuis str. SC-B67.

```
CCTTTACTTCATTTCTTTCTCACAGGCCAGAAAACATAACTG
GCCTGAATATTCTCTCTGGGACCACGGTCCCACCTGTATCGCC
GCTCAGGTTATCTCCCTGGGACCACGGTCCCACCTGCGTCGTC
GGTCAGCTTATCTGCCTGGGACCACGGTCCCACCTGTATCGCC
GGTCAGGTTATCTGCCTGGGACCACGGTCCCACCTGCATCGCC
GGTCAGGTTATCTGCCTGGGACCACGGTCCCACCTGCATCGTC
TGTCAGCTTATCTGCCTGGGACCACGGTCCCACCTGTATCGCC
GGTCAGGTTATCTGCCTGGGACCACGGTCCCACCTGCATCGTC
TGTCAGCTTATCTGCCTGGGACCACGGTCCCACCTGCTTCGTC
GGTCAGCTTATCTGCCTGGGACCACGGTCCCACCTGTATCGCC
GGTCAGGTTATCTGTCTGGGACCACGGTCCCACCTGTATCGCC
GGTCAGGTTATCTGCCTGGGACCACGGTCCCACCTGCATCGTC
GGCACATTATTAGTTAGGGACAACGGTTCATCAATGCCTGTC
```

(note that the 4th repeat contains 44 nucleotides for alignment)

Escherichia coli plasmid pAPEC-1

>gi|221589217|gb|CP000836.1| Escherichia coli strain chi7122 plasmid pAPEC-1.

TACTTAATTTTCCTTTAGCACAGGCTAGAAAACATAACTAACTG
GCCTGAATATTATATCTGGGACCACGGTCCCA GTTGTATTGTC
GGTCTGATAATCAGTCTGGGACCACGGTCCCA CTCATATTGTC
GGTCTGATAATCAGTCTGGGACCACGGTCCCA CTCATATTGTC
GGTCTGATTATCGGTCTGGGACCACGGGCCCA CTCATATTGTT
AGTCTGCTAATCAGTCTGGGACCACGGTCCCA CTCGTATTGTT
GGTCTGCTAATCAGTCTGGGACCACGGTCCCA CTCATATTGTC
GGTCTGCTAATCAGTCTGGGACCACGGTCCCA CTCATATTGTC
GGTCTGCTAATCAGTCTGGGACCACGGTCCCA CTCATATTGTT
GGTCTGATTATCGGTCTGGGACCACGGTCCCA CTCATATTGTT
AGTCTGCTAATCAGTCTGGGACCACGGTCCCA CTCGTATTGTT
GGTCTGCTAATCAGTCTGGGACCACGGTCCCA CTCGTATTGTT
GATCTGATTATCAGCGTGAGACTACGATTCCA TTAATGCCTGC
CAGTGAGTGTGTGGAACATATCTTGGCCAGTGGCTTGGTCTGT

Escherichia coli plasmid pAPEC-O1-ColBM

>gi|88770133|gb|DQ381420.1| Escherichia coli APEC O1 plasmid pAPEC-O1-ColBM.

TACTTAATTTTCCTTTAGCACAGGCTAGAAAACATAACTAACTG
GCCTGAATATTATATCTGGGACCACGGTCCCA GTTGTATTGTC
GGTCTGATAATCAGTCTGGGACCACGGTCCCA CTCATATTGTC
GGTCTGATTATCGGTCTGGGACCACGGGCCCA CTCATATTGTT
AGTCTGCTAATCAGTCTGGGACCACGGTCCCA CTCGTATTGTT
GGTCTGCTAATCAGTCTGGGACCACGGTCCCA CTCATATTGTC
GGTCTGCTAATCAGTCTGGGACCACGGTCCCA CTCATATTGTC
GGTCTGCTAATCAGTCTGGGACCACGGTCCCA CTCATATTGTT
GGTCTGATTATCGGTCTGGGACCACGGTCCCA CTCATATTGTT
AGTCTGATAATCAGTCTGGGACCACGGTCCCA CTCGTATTGTT
GGTCTGCTAATCAGTCTGGGACCACGGTCCCA CTCGTATTGTT
GATCTGATTATCAGCGTGAGACTACGATTCCA TTAATGCCTGC
CAGTGAGTGTGTGGAACATATCTTGGCCAGTGGCTTGGTCTGT

Escherichia coli plasmid F

TCTTTACTTAATGTCCTTTGTTACAGGCCAGAAAGCATAACTG
GCCTGAATATTCTCTCTGGGCCCACTGTTCCA CTTGTATCGTC
GGTCTGATAATCAGACTGGGACCACGGTCCCA CTCGTATCGTC
GGTCTGATTATTAGTCTGGGACCACGGTCCCA CTCGTATCGTC
GGTCTGATTATTAGTCTGGGACCACGGTCCCA CTCGTATCGTC
GGTCTGATAATCAGACTGGGACCACGGTCCCA CTCGTATCGTC
GGTCTGATTATTAGTCTGGGACCACGGTCCCA CTCGTATCGTC
GGTCTGATTATTAGTCTGGGACCACGGTCCCA CTCGTATCGTC
GGTCTGATTATTAGTCTGGGACCACGGTCCCA CTCGTATCGTC
GGTCTGATTATTAGTCTGGGACCACGGTCCCA CTCGTATCGTC
GGTCTGATTATTAGTCTGGGACCACGGTCCCA CTCGTATCGTC
GGTCTGATTATTAGTCTGGGACCACGGTCCCA CTCGTATCGTC
GGTCTGATTATTAGTCTGGGACCACGGTCCCA CTCGTATCGTC
GGTCTGATTATTAGTCTGGGACCACGGTCCCA CTCGTATCGTC
GGTCTGATTATTAGTCTGGGACCACGGTCCCA CTCGTATCGTC
GATCAGACTATCAGCGTGAGACTACGATTCCA TCAATGCCTGT
CAAGGGCAAGTATTGACATGTCGTGTAACCTGTAGAACGGAG

- 11 repeats of 45-bp

***Yersinia pseudotuberculosis* PB1/+ plasmid pYPTS01**

>gi|186700521|gb|CP001049.1| *Yersinia pseudotuberculosis* PB1/+ plasmid pYPTS01

```
CTAGTTTAAATCTCCCTAAAGCTATCACTAAAAAATGCTCATCGA
CTTAATCGTTTTCAATTCGAATCACATTTCCACTAGGATGTTCAA
TGCCAGTGTTAGTAATTGGGACCACGGTCCACCTTGATATCCAA
TGCCAACGTTGGTAATTGGGACCACGGTCCCACTTGATATCCAA
TGCCAACGTTAGTAATTGGGACCACGGTCCCACTTGATATCCAA
TGCCAACGTTGGTAATTGGGACCACGGTCCCACTTGATATCCAA
TGCCAACGTTGGTAATTGGGACCACGGTCCCACTTGATATCCAA
TGCCAACGTTGGTAATTGGGACCACGGTCCCACTTGATATCCAA
TGCCAACGTTGGTAATTGGGACCACGGTCCCACTTGATATCCAA
TGCCAACGTTGGTAATTGGGACCACGGTCCCACTTGATATCCAA
TGCCAACGTTGGTAATTGGGACCACGGTCCCACTTGATATCCAA
TGCCAACGTTGGTAATTGGGACCACGGTCCCACTTGATATCCAA
TGCCAACGTTGGTAATTGGGACCACGGTCCCACTTGATATCCAA
TGCTAACGTTGGTATGATGGGACCACAGTTCAGCTTAGGAAATGAG
TTGTAAAGAAGGAGATCTTCTACGAAGGATGAGCTAAAAACAATAG
```

- 11 repeats of 43 bp

***Escherichia coli* pECOS88**

>gi|218349681|emb|CU928146.1| *Escherichia coli* str. S88 plasmid pECOS88.

```
TACTTAATTTCTTTTGCACAGGCTAGAAAACATAACTAACTG
GCCTGAATATTATATCTGGGACCACGGTCCCACTTGATATTGTC
GGTCTGATAATCAGTCTGGGACCACGGTCCCACTCATATTGTC
GGTCTGATAATCAGTCTGGGACCACGGTCCCACTCATATTGTC
GGTCTGATTATCGGTCTGGGACCACGGTCCCACTCATATTGTT
AGTCTGCTAATCAGTCTGGGACCACGGTCCCACTCATATTGTC
GGTCTGCTAATCAGTCTGGGACCACGGTCCCACTCATATTGTC
GGTCTGCTAATCAGTCTGGGACCACGGTCCCACTCATATTGTT
GGTCTGATTATCGGTCTGGGACCACGGTCCCACTCATATTGTT
AGTCTGATAATCAGTCTGGGACCACGGTCCCACTCATATTGTT
GGTCTGCTAATCAGTCTGGGACCACGGTCTCACTTGATATTGTC
GATCTGATTATCAGCGTGAGACTACGATTCCAATTAATGCCTGC
CAGTGAGTGTGTGGAACATATCTTGCCAGTGGCTTGGTCTGT
```

- 10 repeats of 43 bp

***Salmonella typhimurium* plasmid pU302L**

>gi|37962716|gb|AY333434.1| *Salmonella typhimurium* plasmid pU302L.

```
TCTTTACTTAATGTCTTTTGTACAGGCCAGAAAGCATAACTG
GCCTGAATATTCTCTCTGGGCCACTGTTCCACTTGATATCGTC
GGTCTGATAATCAGACTGGGACCACGGTCCCACTCGTATCGTC
GGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTC
GGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTC
GGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTC
GGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTC
GGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTC
GGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTC
GGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTC
GGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTC
GGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTC
GGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTC
GGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTC
GATCAGACTATCAGCGTGAGACTACGATTCCAATCAATGCCTGT
CAAGGGCAAGTATTGACATGTCGTTCGTAACCTGTAGAACGGAG
```

(*K. pneumoniae* pKF3-140, *E. coli* pIP1206 and *E. coli* pEK499 have same centromeric sequences as pU302L)

- 8 repeats of 45-bp

***Yersinia pseudotuberculosis* IP32953 pYV plasmid**

>gi|108860311|emb|BX936399.2| *Yersinia pseudotuberculosis* IP32953 pYV plasmid

TAGTTTTAATCTCCCTAAAAGCTATCACTAAAAAATGCTCATCGA
CTTAATCGCTTTCAAT**TGGAATCACATTTCCA**CTAGGATGTTCAA
TGCCAGTGTTAGTAAT**TGGGACCACGGTCCCA**CTTGGATATCCAA
TGCCAACGTTGGTAAT**TGGGACCACGGTCCCA**CCTGGATATCCAA
TGTCACGTTGGTAAT**TGGGACCACGGTCCCA**CTTGGATATCCAA
TGCCAACGTTAGTAAT**TGGGACCACGGTCCCA**CCTGGATATCCAA
TGCCAACGTTGGTAAT**TGGGACCACGGTCCCA**CCTGGATATCCAA
TGCCAACGTTAGTAAT**TGGGACCACGGTCCCA**CCTGGATATCCAA
TGCTAACGTTAGTAAT**TGGGACCACGGTCCCA**CCTGGATATCCAA
TGCTAACGTTGGTGAT**TGGGGCCACAGTTCAG**CTTAGGAAATGAG
TTGTAAGAAGGAGATCTTCTACGAAGGATGAGCTAAAACAATAG

- 8 repeats of 44-bp

***Klebsiella pneumoniae* 342 plasmid pKP187**

>gi|206570325|gb|CP000965.1| *Klebsiella pneumoniae* 342 plasmid pKP187

CCGCCTCTCTAGTTTTTTTTCTCAGGCCAGTTTTGATCACCTGGC
CTTCATCTTGT CAGCC**TGGGACCACGGTCCCA**CCTTCACCGTCA
GTACCGGA ACTCTGCC**TGAGAC ACGGTCCCA**CCTTCACCGTCA
GCGTCGGAATTCTGCC**TGGGACCACGGTCCCA**CCTGCGCCGTCA
GTGCCGGAATTCTGCC**TGGGACCACGGTCCCA**CCTGCGCCGTCA
GCGCCGGA ACTCTGCC**TGGGACCACGGTCCCA**CCTGCGCCGTCA
GTGCCGGAATTCTGCC**TGGGACCACGGTCCCA**CCTATGCCGTCA
GTACTGGA ACTCTTCC**TGGGACCATGGTCCCA**CCTGCGCCGTCA
GCGCCGGA ATGCTGCC**TGGGACCACGGTCCCA**CCTGTACCCGCA
GCGCCGGA ATAGAGGTCAATGTCTGTCATCGAATTGCTAATCGC

(note that a gap was introduced in the 7th repeat for alignment)

- 8 repeats of 43 bp

***Escherichia coli* plasmid pVir68**

>gi|253721152|gb|CP001162.1| *Escherichia coli* Vir68 plasmid pVir68.

TCTTTACTTAATGTCCTTTGTTACAGGCCAGAAAGCATAACTG
GCCTGAATATTCTCTCT**TGGGCCCACTGTTCCA**CTTGTATCGTC
GGTCTGATAATCAGACT**TGGGACCACGGTCCCA**CTCGTATCGTC
GGTCTGATTATTAGTCT**TGGGACCACGGTCCCA**CTCGTATCGTC
GGTCTGATTATTAGTCT**TGGGACCACGGTCCCA**CTCGTATCGTC
GGTCTGATTATTAGTCT**TGGGACCACGGTCCCA**CTCGTATCGTC
GGTCTGATTATTAGTCT**TGGGACCACGGTCCCA**CTCGTATTGTC
GGTCTGATTATCGGTCT**TGGGACCACGGTCCCA**CTTGTATTATC
GATCAGACTATCAGCG**TGAGACTACGATTC**CAATCAATGCCTGT
CAAGGGCAAGTATTGACATGTCGTTCGTAACCTGTAGAACGGAG

(*E. coli* pSFO157 has the same centromeric sequence as pVIR68)

- 7 repeats of 45-bp

Yersinia pestis Z176003 plasmid pCD1

>gi|294352395|gb|CP001594.1| Yersinia pestis Z176003 plasmid pCD1

```
CTAGTTTAATCTCCCTAAAGCTATCACTAAAAAAATGCTCATCGA
CTTAATCGCTTTCAATTCGAATCACATTTCCACTAGGATGTTCAA
TGCCAGTGTAGTAATTGGGACCACGGTCCCACTTGGATATCCAA
TGCCAACGTTGGTAATTGGGACCACGGTCCCACTTGGATATCCAA
TGTCACGTTGGTAATTGGGACCACGGTCCCACTTGGATATCCAA
TGCCAACGTTGGTAATTGGGACCACGGTCCCACTTGGATATCCAA
TGCCAACGTTAGTAATTGGGACCACGGTCCCACTTGGATATCCAA
TGCTAACGTTGGTGATTGGGACCACAGTTCAGCTTAGGAAATGAG
TTGTAAAGAAGGAGATCTTCTACGAAGGATGAGCTAAAAACAATAG
```

(*Yersinia enterocolitica* pYVe8081 has the same centromeric sequence as pCD1)

- 6 repeats of 45-bp

Yersinia enterocolitica strain A127/90 plasmid pYVa127/90

>gi|28302111|gb|AY150843.2| Yersinia enterocolitica strain A127/90 plasmid pYVa127/90

```
TAGTTTAATCTCCCTAAAGCTATCACTAAAAAAATGCTCATCGA
CTTAATCGCTTTCAATTCGAATCACATTTCCACTTGGATATCCAA
TGCCAACGTTGGTAATTGGGACCACGGTCCCACTTGGATATCCAA
TGCCAACGTTGGTAATTGGGACCACGGTCCCACTTGGATATCCAA
TGCCAACGTTGGTAATTGGGACCACGGTCCCACTTGGATATCCAA
TGCCAACGTTGGTAATTGGGACCACGGTCCCACTTGGATATCCAA
TGCCAACGTTGGTGATTGGGATCACAGTTCAGATTAGGAAATGAG
TTGTAAAGAAGGTGATCTTCTATGAAGGATGAGCTAAAAACAATAG
```

- 6 repeats of 43-bp

Escherichia coli plasmid pO157

>gi|4589740|dbj|AB011549.2| Escherichia coli O157:H7 str. Sakai plasmid pO157 DNA.

```
TCTTTACTTAATGTCCTTTGTTACAGGCCAGAAAGCATAACTG
GCCTGAATATTCTCTCTGGGCCACTGTTCCACTTGTATCGTC
GGTCTGATAATCAGACTGGGACCACGGTCCCACTCGTATCGTC
GGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTGTGTC
GGTCTGATTATCGGTCGGGACCACGGTCCCACTCGTGTGTC
GGTCTGATTATCGGTCGGGACCACGGTCCCACTTGTATTGTC
GATCAGACTATCAGCGTGAGACTACGATTCTATCAATGCCTGT
CAAGGGCAAGTATTGACATGTCGTCGTAACCTGTAGAACGGAG
```

- 5 repeats of 45-bp

Yersinia pestis Pestoides F plasmid CD

>gi|145212871|gb|CP000669.1| Yersinia pestis Pestoides F plasmid CD

```
CTAGTTTAAATCTCCCTAAAGCTATCACTAAAAAATGCTCATCGA
CTTAATCGCTTTCAATTCGAATCACATTTCCACTAGGATGTTCAA
TGCCAGTGTTAGTAATTGGGACCACGGTCCCACTTGGATATCCAA
TGCCAACGTTGGTAATTGGGACCACGGTCCCACCTGGATATCCAA
TGCCAACGTTAGTAATTGGGACCACGGTCCCACCTGGATATCCAA
TGCTAACGTTGGTGATTGGGGCCACAGTTCAGCCTTAGGAAATGAG
TTGTAAAGAAGGAGATCTTCTACGAAGGATGAGCTAAAAACAATAG
```

- 4 repeats of 43-bp

Escherichia coli plasmid pETEC_74

>gi|157076671|gb|CP000799.1| Escherichia coli E24377A plasmid pETEC_74.

```
TCTTTACTTAAATGTCCCTTTGTTACAGGCCAGAAAGCATAACTG
GCCTGAATATTCTCTCTGGGCCCACTGTTCCACTTGTATCGTC
GGTCTGATTATTAGTCTTGGGACCACGGTCCCACTCGTGTGTC
GGTCTGATTATCGGTCTTGGGACCACGGTCCCACTTGTATTGTC
GATCAGACTATCAGCGTGAGACTACGATTCCAATCAATGCCTGT
CAAGGGCAAGTATTGACATGTGTCGTAACCTGTAGAACGGAG
```

(*E. coli* p557 has the same centromeric sequence as pETEC_74)

- 3 repeats of 43-bp

Yersinia enterocolitica plasmid pYVe227

>gi|4324323|gb|AF102990.1| Yersinia enterocolitica plasmid pYVe227.

```
TAATCGCTTTCAATTCGAATCACATTTCCACTTGGATATCCAA
TGCCAACGTTGGTAATTGGGACCACGGTCACCTGGATATCCAA
TACCAGCGTTAGTAATTGGGACCACGGTCACCTGGATATCCAA
TGCCAACGTTGGTGATTGGGACCACAGTTCAGCCTTAGGAAATC
AATTGTAAAGAAGGACATCTTCAATGAAGGATGAGCTAAAAATA
```

Figure S4. Sequence conservation of the SopB specificity domain involved in SopB-*sopC* interaction. **A.** Alignment of SopB sequences, corresponding to cognate *sopC* sequences listed in Fig. S3, from residues 176 to 221. Residues conserved amongst all SopB members are highlighted in black. Residues conserved in subgroups belonging to 43-, 44- and 45-bp repeats carrying centromeres are highlighted in blue, green and red, respectively. Amino acid residues are numbered according to the F SopB sequence; R190, K191, R195 and R219 responsible for the specific binding to *sopC* 16-bp IR are indicated above in red. **B.** Weblogo of HTH motifs from all SopB sequences listed in A. SopB amino-acid sequences from residues 176 to 200 were aligned with ClustalW software and computed on the weblogo.berkeley.edu website to generate the sequence logo.

