

**Table S5. *C. difficile* Cas locus analysis.** A) Occurrence of *cas* loci homologous to *CD2982-CD2975* and *CD2455-CD2451* from 630 strain in published genomes of *C. difficile* strains. Correlation between MLST of strains and the absence of *CD2982-2975 cas* locus (B) or the presence of the *CDR2998-2994* locus (C).

**A**

Accession	PMID	Nb. strains analysed	<i>CD2982-75</i>	<i>CD2455-51</i>	<i>CDR2998-94</i>
PRJEB4556	24066741	785	633	785	289
PRJEB3093	24336451	494	465	494	117
PRJEB2318	23222960	274	273	274	269
PRJEB1483	24336451	210	200	210	4
PRJEB2207	23222960	151	151	151	147
PRJEB1729	24336451	143	125	143	49
PRJEB2039	23222960	55	53	52	51
PRJEB3010	24336451	43	37	43	2
PRJEB190-216	24482682	27	23	27	1
PRJEB2134	23222960	22	22	22	21
PRJEB6600-2	24913157	3	0	3	0
		<b>2207</b>	<b>89.8%</b>	<b>99.9%</b>	<b>43.0%</b>

**B**

MLST <sup>a</sup>	Strains lacking <i>CD2982-75</i>	%
3 (027)	3 / 856	0%
5	3 / 3	100%
9	19 / 19	100%
12	9 / 10	90%
13	7 / 7	100%
18	2 / 2	100%
19	124 / 237	52%
22	2 / 2	100%
23	1 / 23	4%
25	5 / 10	50%
27	1 / 3	33%
45	37 / 37	100%
46	6 / 6	100%
51	2 / 2	100%
58	4 / 4	100%

<sup>a</sup> MLST: Multilocus sequence typing

**C**

MLST <sup>a</sup>	Strains with <i>CDR2998-94</i>	%
3 (027)	849 / 856	99%
4	2 / 4	50%
6	2 / 12	17%
9	19 / 19	100%
15	8 / 8	100%
17	20 / 23	87%
19	1 / 237	0%
21	12 / 12	100%
23	16 / 23	70%
24	19 / 20	95%
25	9 / 10	90%
33	1 / 151	1%
38	7 / 7	100%
39	3 / 180	2%
49	1 / 61	2%

<sup>a</sup> MLST: Multilocus sequence typing