

**Table S1.** Pherotype distribution among serotypes.

<b>Serotype</b>	<b>CSP-1</b>	<b>CSP-2</b>	<b>OR (95%CI)</b>	<b>FDR<sup>a</sup></b>
1 <sup>b</sup>	48	2	11.43 (2.93;98.52)	<10 <sup>-4</sup>
3 <sup>b</sup>	23	23	0.38 (0.19;0.73)	0.017
4	16	13	0.49 (0.21;1.14)	NS <sup>c</sup>
5	9	4	0.94 (0.26;4.23)	NS <sup>c</sup>
6A <sup>b</sup>	2	11	0.08 (0.007;0.33)	<10 <sup>-4</sup>
6B	9	5	0.74 (0.22;2.88)	NS <sup>c</sup>
7C	0	1	0 (0;16.24)	NS <sup>c</sup>
7F	14	4	1.48 (0.45;6.27)	NS <sup>c</sup>
8	20	8	1.04 (0.43;2.81)	NS <sup>c</sup>
9L	0	1	0 (0; 16.24)	NS <sup>c</sup>
9N <sup>b</sup>	2	8	0.10 (0.01;0.51)	0.013
9V	20	2	4.35 (1.03;38.92)	NS <sup>c</sup>
10A	4	5	0.33 (0.064;1.54)	NS <sup>c</sup>
10F	2	1	0.83 (0.04;49.42)	NS <sup>c</sup>
11A	7	0	Inf <sup>d</sup> (0.60;Inf <sup>d</sup> )	NS <sup>c</sup>
11B	0	1	0 (0;16.24)	NS <sup>c</sup>
12B	13	3	1.83 (0.49;10.19)	NS <sup>c</sup>
13	2	1	0.83 (0.04;49.42)	NS <sup>c</sup>
14 <sup>b</sup>	61	4	7.50 (2.70;28.99)	<10 <sup>-4</sup>
15B	2	0	Inf <sup>d</sup> (0.08;Inf <sup>d</sup> )	NS <sup>c</sup>
16F	2	2	0.41 (0.00;5.76)	NS <sup>c</sup>
17F	0	2	0 (0;2.21)	NS <sup>c</sup>
18A	3	2	0.62 (0.07;7.52)	NS <sup>c</sup>
18C	8	4	0.83 (0.22;3.83)	NS <sup>c</sup>
19A	12	5	1.00 (0.32;3.69)	NS <sup>c</sup>
19F	5	6	0.34 (0.08;1.35)	NS <sup>c</sup>
20	3	2	0.62 (0.07;7.52)	NS <sup>c</sup>
22F	12	0	Inf <sup>d</sup> (1.18;Inf <sup>d</sup> )	NS <sup>c</sup>
23A	1	0	Inf <sup>d</sup> (0.01; Inf <sup>d</sup> )	NS <sup>c</sup>
23B	2	0	Inf <sup>d</sup> (0.08; Inf <sup>d</sup> )	NS <sup>c</sup>
23F	18	6	1.26 (0.47;3.97)	NS <sup>c</sup>
24F	0	2	0 (0;2.21)	NS <sup>c</sup>
25A	0	1	0 (0;16.24)	NS <sup>c</sup>
25F	0	1	0 (0;16.24)	NS <sup>c</sup>
31	0	3	0 (0; 1)	NS <sup>c</sup>
33A	1	0	Inf <sup>d</sup> (0.01;Inf <sup>d</sup> )	NS <sup>c</sup>
33F	7	0	Inf <sup>d</sup> (0.60;Inf <sup>d</sup> )	NS <sup>c</sup>
35B	1	0	Inf <sup>d</sup> (0.01;Inf <sup>d</sup> )	NS <sup>c</sup>
35F	3	0	Inf <sup>d</sup> (0.17;Inf <sup>d</sup> )	NS <sup>c</sup>
36	0	1	0 (0;16.24)	NS <sup>c</sup>
38	1	1	0.42 (0.005;32.78)	NS <sup>c</sup>
NT <sup>e</sup>	8	7	0.42 (0.14;1.54)	NS <sup>c</sup>

<sup>a</sup> Correction for multiple testing performed by the false discovery rate method (FDR)

<sup>b</sup> p<0.05 after FDR correction

<sup>c</sup> NS, not significant

<sup>d</sup> Inf, infinite value.

<sup>e</sup> NT, non-typable strains