

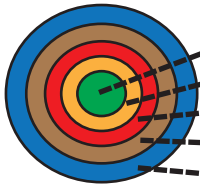
**B**

*Rickettsia buchneri* REIS\_1424 N- [ ] -C 263

Fig. S1

*Rickettsia*-specific insertion

REIS_1424	112	[38]	KIP	[8]	SEKGTGLLY	[4]	NGVRIMKGRS	[6]	QKVDYAVIVDKGCRIDKYG	[23]	[14]	HINLEEWIKKWITHNKP	[0]	263
NF27_IC00050	2067	[29]	KIP	[8]	NREGVGIRI	[7]	SVIRIDKGNP	[6]	QQVDHVRINYNQVIGRDG	[0]	[18]	HIPLSEW-LQWKEWNKP	[0]	2192
Sphch_1491	1275	[38]	RRR	[8]	NRKGVGSRF	[7]	NRVRVDRGKP	[6]	QRPDHVEQRNGVTVDGNG	[0]	[14]	HVPLRDW-LR-----	[2]	1400
J0H87_08305	650	[10]	KVP	[8]	A-KGEGWRW	[6]	DGVRIMKGDG	[6]	QHVDYVKINTDGKLVGLDG	[0]	[14]	HIPLSEW-LTWKEWN--	[2]	750
C4J97_0166	2557	[9]	KFP	[7]	NKKGVGFRW	[7]	NGVRIDKGEF	[6]	QQVDHVI VRSNGRVIGRDG	[0]	[14]	HIPLSEW-KKWKWNNSP	[0]	2658
EDF81_3004	454	[11]	KFP	[8]	NKKGVGFRW	[7]	NGVRIDQGNP	[6]	QQVDHVI VRNNGKVIIGRDG	[0]	[14]	HIPLSEW-KKWKWNNSP	[0]	558
EKH77_08965	1030	[7]	KVP	[9]	NRKGVGVRW	[5]	NGVRIDKGEA	[6]	QQVDHVVINYGGKVIIGRDG	[0]	[14]	HIPLEEW-RKWRSWNQP	[0]	1128
BJ981_004090	2803	[9]	KIP	[8]	NKKGVGTRW	[7]	SGVRIDKGDG	[6]	QQRDHVI VRDHGKVLGRNG	[0]	[14]	HIPFSEW-SKWSQWNKP	[0]	2904



Taxon database

	NF27_IC00050 (as in Fig. 2A)			REIS_1424 (with insertion)			REIS_1424 (w/o insertion)		
	Subj.	$\bar{x}$ Sm	top 10 Sm	Subj.	$\bar{x}$ Sm	top 10 Sm	Subj.	$\bar{x}$ Sm	top 10 Sm
Rickettsiales	5	4330.26	na	1	na	na	1	na	na
Alphaproteobacteria (- Rickettsiales)	10	2149.50	2149.50	0	na	na	1	na	na
Proteobacteria (- Bacteria)	44	3160.01	3820.23	0	na	na	3	950.56	na
Bacteria (- Proteobacteria)	96	3048.58	4635.10	2	na	na	21	1385.68	1579.64
- Bacteria	1	na	na	0	na	na	0	na	na

removing the insertion results in greatest similarity to non-proteobacterial proteins