

Supplementary Table 1. Functional analysis of proteins encoded by the prophage phIS3501.

phage gene	size , aa	best match	best match strain	identity	best function guess*	matches > 80%	best phage match
phIS3501p01	353	228912160	<i>B. thuringiensis</i> IBL 200	349/353 (99%)	phage integrase	2	<i>Cl. perfringens</i> phage phi3626 (38%)
phIS3501p02	57		not found		phage protein	0	not found
phIS3501p03	84	228956232	<i>B. thuringiensis</i> var kurstaki T03a001	76/84 (90%)	excisionase	4	not found
phIS3501p04	384	228981667	<i>B. thuringiensis</i> Bt407	377/384 (98%)	H-T-H protein	6	<i>B. cereus</i> phage phBC6A51 (22%)
phIS3501p05	109	228912155	<i>B. thuringiensis</i> IBL 200	105/109 (96%)	lysogenic repressor	3	Listeria phage B054 (41%)
phIS3501p06	92		not found		phage protein	0	not found
phIS3501p07	259	228962250	<i>B. thuringiensis</i> var pakistani TI3001	254/259 (98%)	antirepressor	1	<i>B. anthracis</i> phage Fah (74%)
phIS3501p08	112	228955973	<i>B. thuringiensis</i> var kurstaki T03a001	94/103 (91%)	conserved phage protein	2	Bacillus sp. 6k512 phage bv1 (29%)
phIS3501p09	215	228967547	<i>B. thuringiensis</i> var sotto str. T04001	168/215 (78%)	RNA polymerase sigma-factor	0	not found
phIS3501p10	334	196042515	<i>B. thuringiensis</i> NVH0597-99	258/338 (76%)	phage replication protein DnaD	0	<i>Lactobacillus casei</i> phage phiAT3 (36%)
phIS3501p11	270	228956030	<i>B. thuringiensis</i> var kurstaki T03a001	241/247 (98%)	phage replication protein DnaC	2	Bacillus sp. 6k512 phage bv1 (27%)
phIS3501p12	88	228956031	<i>B. thuringiensis</i> var kurstaki T03a001	83/88 (94%)	group-specific phage protein	2	not found
phIS3501p13	71	75761347	<i>B. thuringiensis</i> var israelensis ATCC 35646	49/71 (69%)	conserved phage protein	0	<i>Staphylococcus aureus</i> phage phiNM3 (41%)
phIS3501p14	99	228956033	<i>B. thuringiensis</i> var kurstaki T03a001	84/86 (98%)	phage protein	1	not found
phIS3501p15	63	152976662	" <i>B. cytotoxicus</i> " NVH 391-98	58/63 (92%)	phage protein	1	not found
phIS3501p16	102	217958104	<i>B. cereus</i> AH187	89/102 (87%)	phage protein	5	not found
phIS3501p17	130	67078316	<i>B. cereus</i> E33-L	90/105 (86%)	conserved phage protein	2	Geobacillus virus E2 (50%)
phIS3501p18	87	229187784	<i>B. cereus</i> BGSC 6E1	81/87 (93%)	phage protein	2	not found
phIS3501p19	52		not found		phage protein	0	not found
phIS3501p20	111	228955992	<i>B. thuringiensis</i> var kurstaki T03a001	103/104 (99%)	group-specific phage protein	1	not found
phIS3501p21	161	228962266	<i>B. thuringiensis</i> var pakistani TI3001	159/161 (99%)	ArpU-family transcriptional regulator	42	Geobacillus phage GBSV1 (37%)
phIS3501p22	180	229017458	<i>Bacillus cereus</i> AH1273	165/180 (92%)	site-specific recombinase	55	Clostridium phage phi3626 (51%)
phIS3501p23	74	324328843	<i>B. thuringiensis</i> var finitimus YBT-020	69/74 (93%)	group-specific phage protein	4	not found

phis3501p24	99	228962269	<i>B. thuringiensis</i> var pakistani TI3001	87/95 (92%)	phage protein	2	not found
phis3501p25	73	228962270	<i>B. thuringiensis</i> var pakistani TI3001	64/73 (88%)	group-specific phage protein	4	<i>B. cereus</i> phage phBC6A51 (65%)
phis3501p26	92	228962317	<i>B. thuringiensis</i> var pakistani TI3001	65/87 (75%)	group-specific phage protein	0	<i>B. cereus</i> phage Wβ (61%)
phis3501p27	70	228961949	<i>B. thuringiensis</i> var pakistani TI3001	69/70 (99%)	group-specific phage protein	25	<i>B. cereus</i> phage phBC6A51 (97%)
phis3501p28	84	228966909	<i>B. thuringiensis</i> var sotto str. T04001	76/84 (90%)	phage protein	12	<i>B. cereus</i> phage Wβ (82%)
phis3501p29	125	218232061	<i>B. cereus</i> B4264	119/125 (95%)	endonuclease	11	<i>B. cereus</i> phage Wβ (51%)
phis3501p30	167	218234873	<i>B. cereus</i> B4264	164/167 (98%)	terminase small subunit	8	Geobacillus virus E2 (68%)
phis3501p31	564	218234096	<i>B. cereus</i> B4264	553/564 (98%)	terminase large subunit	9	Geobacillus virus E2 (58%)
phis3501p32	417	228936951	<i>B. thuringiensis</i> var andalousiensis BGSC 4AW1	402/417 (96%)	portal protein	9	<i>B. subtilis</i> phage phi105 (41%)
phis3501p33	236	228987127	<i>B. thuringiensis</i> var tochigiensis BGSC 4Y1	182/197 (92%)	prohead protease	6	<i>Staphylococcus aureus</i> phage phiSLT (41%)
phis3501p34	388	218231720	<i>B. cereus</i> B4264	371/387 (96%)	major capsid protein	9	<i>B. subtilis</i> phage phi105 (43%)
phis3501p35	95	218230975	<i>B. cereus</i> B4264	93/94 (99%)	DNA packaging protein	6	Geobacillus virus E2 (51%)
phis3501p36	107	218232936	<i>B. cereus</i> B4264	102/107 (95%)	head-tail adaptor	6	<i>B. cereus</i> phage Wβ (85%)
phis3501p37	144	218232189	<i>B. cereus</i> B4264	137/144 (95%)	conserved phage protein	9	<i>B. cereus</i> phage Wβ (80%)
phis3501p38	119	218232693	<i>B. cereus</i> B4264	112/119 (94%)	conserved phage protein	10	<i>B. cereus</i> phage Wβ (85%)
phis3501p39	202	228956048	<i>B. thuringiensis</i> var kurstaki T03a001	197/202 (98%)	major tail protein	15	<i>B. cereus</i> phage Wβ (84%)
phis3501p40	105	228956406	<i>B. thuringiensis</i> var kurstaki T03a001	100/104 (96%)	conserved phage protein	15	<i>B. cereus</i> phage Wβ (83%)
phis3501p41	1343	218231322	<i>B. cereus</i> B4264	1105/1345 (82%)	tape measure protein	6	<i>B. cereus</i> phage Wβ (87%)
phis3501p42	494	218231887	<i>B. cereus</i> B4264	462/493 (94%)	phage tail protein	5	<i>B. cereus</i> phage Wβ (66%)
phis3501p43	1341	206975170	<i>B. cereus</i> H3081.97	1187/1337 (89%)	phage minor structural protein protein	4	<i>B. cereus</i> phage Wβ (70%)
phis3501p44	74	228976467	<i>B. thuringiensis</i> var thuringiensis str. T01001	69/74 (93%)	group-specific phage protein	23	<i>B. cereus</i> phage phBC6A51 (50%)
phis3501p45	141	228955776	<i>B. thuringiensis</i> var kurstaki T03a001	137/141 (97%)	holin	42	<i>B. cereus</i> phage Wβ (95%)
phis3501p46	229	228955777	<i>B. thuringiensis</i> var kurstaki T03a001	221/229 (97%)	phage lysin	8	<i>B. cereus</i> phage Wβ (77%)
phis3501p47	193		not found		phage protein	0	not found
phis3501p48	192	30263934	<i>Bacillus anthracis</i> str. Ames	178/191 (93%)	phage protein	7	not found
phis3501p49	105	324328816	<i>B. thuringiensis</i> var finitimus YBT-020	103/105 (98%)	group-specific phage protein	2	<i>B. cereus</i> phage Wβ (73%)
phis3501p50	393	228912737	<i>B. thuringiensis</i> IBL 200	384/393 (98%)	FtsK-SpoIIIE family phage protein	13	<i>B. cereus</i> phage Wβ (76%)
phis3501p51	206	74488608	<i>B. thuringiensis</i> var israelensis ATCC 35646	197/206 (96%)	conserved phage protein	22	<i>B. cereus</i> phage phBC6A51 (80%)
phis3501p52	110	228912735	<i>B. thuringiensis</i> IBL 200	95/99 (96%)	phage protein	5	not found

\*defined biochemical functions labeled by rose color; grey indicates proteins without similarity in the used databases.

Supplementary Table 2.

Identity between contigs of *Bacillus thuringiensis* IBL 4222 and pHIS3501 (query).

Contigs in IBL4222	Query coverage	E-value	% of identity	Other info
contig00563	39%	0.0	99%	Identities = 17536/17537 (99%), Gaps = 0/17537 (0%)
contig00026	35%	0.0	100%	Identities = 15759/15761 (99%), Gaps = 0/15761 (0%) Strand=Plus/Plus
contig00668	13%	0.0	99%	Identities = 6213/6213 (100%), Gaps = 0/6213 (0%) Strand=Plus/Plus
contig00040	4%	0.0	100%	Identities = 1774/1779 (99%), Gaps = 5/1779 (0%) Strand=Plus/Plus
contig00092	3%	0.0	100%	Identities = 1560/1560 (100%), Gaps = 0/1560 (0%) Strand=Plus/Minus
contig00598	6%	0.0	91%	Identities = 1661/2050 (81%), Gaps = 14/2050 (1%) Strand=Plus/Minus

Supplementary Table 3.

Primers used for sequence assembly verification and attachment site sequencing

Primer names	Sequences	Position relative to the phIS3501 genome
PHISB8	GGTATGGCATCCGATGGCTTAC	30766-30787
PHISE8*	CCCTTGCGTAGCATCAATAGTTTTC	+ (115-90)
PHISE5	AGCGGTGAAAAGGTAACGGTTGAAT	21517-21541
PHISC3	CGATAGCGGATCATCCCCATTA	30840-30819
PHISG8	CTACACGAATTCCGACATTCTCATGC	5782-5808
PHISD6	TTCATCTCCCCATCTCCCTTCTCAA	15046-15021
PHISA1*	GTGAGATACACGTCAAAGATTTAGC	- (5907-5932)
PHISH2	CCTAATAGCGCAACGCCATGACGCTTG	6211-6185
PHISI2	AAATCGATTTAGTTCACCTGTTGCCC	27813-27788
PHISI3	AAAACATTATCTAATAAGGGGGAAGATGA	27147-27175
PHISK5	AATACGTTTTTTCATAACGGTTTGTCTCC	44075-44103
PHISK6	CAATAGTATATCTTCGAGTACCTTACG	323-297
PHISS8	TTTACCTGGTCAGCTCAGGT	44201-44220
PHISC7*	CTACGGAAGATGGCAGACAAATTGC	- (939-964)
PHISA2*	GGCCATCAGCTTATCATACTGAAGC	-(246-271)
PHISC8	CAATTCGAAATTAGCAAGTCTACAAC	247-222
PHISC5*	CTTCGTAGCCACGCTCTCCAAA	+ (336-314)
PHIST6**	GTGAGAAAAATGAAAGAACCAAAGATTGGT	1557-1586
PHIST7**	TTAAACATTATGTTTCTCCTTCCACTGCTCA	786-816

\*primer outside of the phage sequence

\*\*contig 751 (position given on this contig).